Behavioral and neural signatures of information sampling to guide economic choice

Paula Kaanders

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the University of Oxford

Wadham College

Trinity Term 2022

Behavioral and neural signatures of information sampling to guide economic choice

Paula Kaanders

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the University of Oxford.

Wadham College, Oxford **Trinity Term 2022**

Abstract

Humans have an extraordinary capacity to seek out novel information relevant for current and future goals. However, the use of simple choice tasks in decision-making research in which participants are not able to sample information freely has limited our ability to investigate the drivers of information sampling and their neural signatures. In this thesis, I present findings using active information sampling tasks combined with functional magnetic resonance imaging (fMRI) data and large-scale online participant samples.

We find that activity in areas of medial frontal cortex (MFC) and parietal cortex predict upcoming decisions to sample in a multi-step decision-making task with active sampling. Furthermore, a large network of brain regions is shown to represent key features of sampled information.

I then describe the development of a novel information sampling task, in which hypothesized drivers of information sampling can be manipulated. Running many pilot task versions on online participants reveals sampling behavior to be heterogeneous and the existence of many sampling strategies.

A version of this task was used in combination with psychiatric questionnaires to investigate the link between compulsivity and information sampling in a large-scale online study. We find evidence for an effect of compulsivity on the integration of task features to drive sampling, but this is not replicated in a confirmation sample.

Finally, another version of the developed task was used in combination with fMRI to investigate the neural signatures of different drivers of sampling. One driver of sampling was represented in brain activity: goal proximity. We find a similar pattern of brain activity predictive of upcoming decisions to sample, but this effect does not reach significance in MFC.

The work in this thesis shows that areas known to frequently activate in decision-making tasks are likely to also drive information sampling to satisfy upcoming goals.

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

Acknowledgements

I am very grateful to have been supervised by Jill O'Reilly and Laurence Hunt during this PhD, without whose support this work would not have been possible. Laurence has always generously given a lot of his time to help me out when I got stuck, which has enabled me to make steady progress over the years. Jill would frequently solve problems we'd been struggling with for days with a single suggestion, which is a skill I hope to someday acquire. They're a great team, and I'm grateful to have been allowed to benefit and learn from their unique skills and expertise.

I would like to thank all the other people who have worked on the projects described here. Hamed Nili taught me how to do RSA on our fMRI data. Keno Jüchems has contributed to the literature review with his expertise on machine learning. The radiographers and scan operators at OHBA, in particular Juliet Semple, Nicky Aikin and Jemma Pitt, helped collect the fMRI data discussed in this thesis. I especially want to thank them for going out of their way to get my data collection finished when personal circumstances combined with the Covid-19 pandemic made me very short on time towards the end. Andrea Colon-Perez and Yuting Zhang were two MSc students who were great to supervise and have contributed to some of the data collection and analyses described here. Specifically, Andrea collected the discovery sample for our large-scale online study and Yuting investigated the effect of information revelation on brain activity in my second fMRI study.

I am also grateful to Benedetto De Martino, who allowed me to come work with him as an MSc student and then continued to support me by taking me on as an RA. I also want to thank Eric Schulz, who has hired me as a postdoc, and has already proven to be a very considerate and creative supervisor.

Finally, I would like to thank all the lab members I've worked with over the years: all projects in this thesis have greatly benefitted from everybody's input.

Contents

Short abstract 2
Acknowledgements 3

Chapter 1: Introduction 7
1.1 Sampling, searching, curiosity and exploration 9
1.1.1 Sampling vs searching, or curiosity? 9
1.1.2 Exploration 10
1.1.3 Instrumental vs non-instrumental 11
1.1.4 Experimental work in this thesis 11
1.2 Information sampling in economic choice 12
1.2.1 Evidence accumulation 13
1.2.2 How does uncertainty affect information seeking? 15
1.3 Dissociable mechanisms of information sampling in the brain 16
1.3.1 Is information inherently rewarding? 17
1.3.2 Exploration vs exploitation 20
1.3.3 The value of information vs choosing to sample 22
1.3.4 Individual differences in information search 28
1.4 New approaches and challenges to studying information sampling for decision-making 29
1.4.1 Need for more complex tasks 29
1.4.2 Use of priors in sampling and decision-making 29
1.4.3 Future directions 31
1.5 Conclusion 32

Chapter 2: Methods for analysing fMRI data 34
2.1 Model-based fMRI analysis 35
2.1.1 General linear model 35
2.1.2 Higher-level analyses 39
2.1.3 Significance testing 41
2.2 Representational similarity analysis 43
2.2.1 Representational geometry 44
2.2.2 Similarity measures 44
2.2.3 Analysis 46
2.2.4 What are the advantages of this approach? 47
2.2.5 How is RSA used in neuroscience? 49
2.2.6 How successful is RSA performance in PFC? 53
2.3 Summary 58

Chapter 3: Medial frontal cortex activity predicts information sampling in economic choice **59**
3.1 Introduction 61
3.2 Methods 63
3.2.1 Experimental Design and Statistical Analyses 64
3.2.2 fMRI Data Collection 66
3.2.3 fMRI Data Analysis 67
3.2.4 Behavioral Data Analysis 67
3.2.5 Representational Similarity Analysis (RSA) 68
3.2.6 Mass Univariate Analysis 68
3.2.7 Code Accessibility 69
3.3 Results 70
3.4 Discussion 85

Chapter 4: Developing a novel information sampling task **91**
4.1 Ecological information search 91
4.1.1 Active sampling 92
4.1.2 Separating information sampling from choice 92
4.1.3 Manipulating volatility and goal proximity 93
4.1.4 Overview of task development 94
4.2 Task development 96
4.2.1 Task description 96
4.2.2 Task version 1: Dual goal manipulation 98
4.2.3 Task version 2: Information obtained from sampling effectively used to guide choices 101
4.2.4 Task version 3: No strong effect of rate of value change on sampling 109
4.2.5 Task version 4: Manipulating variance of rate of value change affects sampling behavior, but is difficult for online participants to learn 111
4.2.6 Task version 5: Goal proximity not a strong predictor of sampling 113
4.2.7 Task development conclusions 117
4.3 Unlimited sampling – task version 6 118
4.4 Values change unpredictably – task version 7 120
4.4.1 Task description 120
4.4.2 Pilot results 121
4.5 Summary 124
4.6 Supplementary materials 126

Chapter 5: Compulsivity and information sampling **154**
5.1 Introduction 155
5.2 Methods 157
5.2.1 Participants 157
5.2.2 Task 158
5.2.3 Questionnaires 160
5.2.4 Participant exclusion 161
5.2.5 Statistical analyses 162
5.2.6 Transdiagnostic psychiatric dimension scores 162
5.2.7 Ethics 162
5.3 Results 162

5.3.1 Discovery sample 165
5.3.2 Confirmation sample 173
5.4 Discussion 178
5.5 Supplementary materials 183

Chapter 6: Goal proximity mediates activity in a widespread brain network involved in information sampling to guide economic choice **189**
6.1 Introduction 189
6.2 Results 193
6.3 Discussion 205
6.4 Methods 209
6.4.1 Experimental design 209
6.4.2 fMRI data collection 211
6.4.3 fMRI data analysis 211
6.4.4 Behavioral data analysis 212
6.4.5 Mass univariate analysis 212

Chapter 7: General discussion **213**
7.1 What have we learned? 213
7.2 What should be addressed in future work? 215
7.3 What are the implications of this work? 217
7.4 Conclusion 218

References 219

Chapter 1: Introduction

This chapter is an extended version of this publication:

Kaanders, P., Juechems, K., O'Reilly, JX, Hunt, LT (2021). Dissociable mechanisms of information sampling in prefrontal cortex and the dopaminergic system. *Current Opinion in Behavioral Sciences, Vol. 41*, 63-70.

Abstract

This thesis presents novel behavioral and neural findings on the role of information sampling in economic choice. Animals have an extraordinary capacity for sparsely and effectively sampling environments rich in information, but until recently this capacity has only been studied in simplified tasks that often do not allow the agent to actively sample information. The drift diffusion model (DDM) makes implicit predictions about how information is sampled leading up to choice, which are difficult to test in these tasks. Recent information sampling studies that address some of the shortcomings of previously used simplified tasks suggest there may be functional dissociations between several regions in prefrontal (PFC) and parietal cortex. The work presented in this thesis investigates some of the open questions sparked by this prior literature. In this chapter, I introduce some of the terminology used in this literature, predictions leading on from the DDM, dissociable neural mechanisms involved in information sampling, and some current challenges this thesis hopes to address.

The large effect of information on people's decisions has become abundantly clear in the past two years. Vaccine hesitancy and the use of unproven medical treatments have been linked to the spread of misinformation in the media and online (Broniatowski et al., 2018; Kluger, 2020). Modern media have substantially increased the already enormous amount of information available in the world, and it seems that using different strategies to sample this large information space can lead to extremely divergent world views, which in turn can lead to strongly differing life decisions. Yet, while much scientific work has been done to investigate how humans deploy attention to gather information, we do not yet fully understand what features in the environment drive this attention and how sampling choices guide future decision-making.

In this thesis, these interactions are explored through the use of information sampling tasks in which sampling is under the participant's control. In doing so, we have used established frameworks of economic choice to view information sampling as adaptive sequential decisions leading up to a final choice. The combined use of these tasks with behavioral and neuroimaging data collection has allowed us to investigate the neural signatures of dissociable aspects of information sampling, and how these affect choice.

In this chapter, I will summarize some previous literature that inspired and informed the studies presented in this thesis. Firstly, I will go through the wide-ranging terminology used when discussing information seeking (**section 1.1**). Then I will discuss the role of information sampling in economic choice by outlining a decision-making framework that implicitly includes hypotheses about information sampling as a sequential decision-making process in itself (**section 1.2**). Next I will relate some of the different types of information seeking discussed in section 1.1 to findings from the

machine learning and brain imaging literatures, which will lead to some specific hypotheses about dissociable neural mechanisms supporting different types of information seeking (**section 1.3**). Finally, I will outline what the current challenges are in studying information seeking that have led me to set up the studies discussed in the rest of this thesis (**section 1.4**).

1.1 Sampling, searching, curiosity and exploration

When we study how animals look for information, we usually severely restrict the stimuli they are exposed to, the information that can be gathered, and the types of reward that can be maximized (Gottlieb & Oudeyer, 2018). This does not effectively probe the remarkable capacity intelligent animals have to sparsely sample the vast amount of information available to them in the world (Gottlieb et al., 1998). Increasingly, this is being addressed by studying curiosity within a wide range of more complex behaviors, such as exploration and learning (Gruber et al., 2014; Jirout & Klahr, 2012; Kang et al., 2009; Lau et al., 2020).

1.1.1 Sampling vs searching, or curiosity?

Partly as a result of this diversity in study settings, researchers have struggled to clearly define the different types of information gathering that intelligent agents are capable of (Kidd & Hayden, 2015). Gottlieb & Oudeyer (2018) make a distinction between *information sampling*, the gathering of information to guide a specific task, and *information search*, the exploration of an environment without the aim of completing a particular task or goal. However, there are many other terms used to describe similar behaviors. Daniel Berlyne, one of the most prominent psychologists

to study curiosity, made distinctions between *perceptual* and *epistemic curiosity*, where *perceptual curiosity* refers to animals' preference for novel stimuli (Berlyne, 1955; Butler, 1953, 1957a, 1957b; Ennaceur & Delacour, 1988; Hughes, 1965, 1968, 2007; Nicki, 1970; Reed et al., 1996; Stolba & Wood-Gush, 1980; Wood-Gush & Vestergaard, 1991), while *epistemic curiosity* refers to the desire to acquire knowledge, more commonly seen in humans (Berlyne, 1966). He also distinguished *specific* and *diversive* curiosity, where *specific curiosity* refers to the need for a specific piece of information, while *diversive* curiosity refers to the need to explore an environment in the absence of a specific task (Berlyne, 1954), which is similar to Gottlieb & Oudeyer's (2018) *sampling* vs *search* distinction.

1.1.2 Exploration

In neuroscience, *exploration* is often defined as foregoing the current best option in order to sample information about an alternative, in multi-alternative bandit or foraging tasks (Blanchard & Gershman, 2018; Gottlieb & Oudeyer, 2018). However, to maximize reward effectively an agent must also learn the structure of its environment through exploration (Schwartenbeck et al., 2019). This may be a main reason why children are highly exploratory and take delight in pursuing seemingly arbitrary goals – they have yet to learn the full structure of their surroundings (Singh et al., 2010). This type of exploration is also frequently referred to as curiosity and is most similar to Berlyne's definition of diversive curiosity (Berlyne, 1954; Kim et al., 2020).

1.1.3 Instrumental vs non-instrumental

Finally, a distinction is often made between sampling of information that is directly useful to guide upcoming decisions and of information that is not. This is commonly referred to as *instrumental* versus *non-instrumental information seeking*. Recent behavioral data suggests that a combination of both non-instrumental and instrumental factors drive human information search (Golman et al., 2022; Golman & Loewenstein, 2018; Kobayashi et al., 2019), but as of yet it is unclear what determines how each factor is weighted during sequential decisions to look for information. It is possible that what is deemed as non-instrumental information sampling in decision-making tasks, is in fact a form of exploration to reduce uncertainty about the structure of the environment rather than maximizing reward.

1.1.4 Experimental work in this thesis

Kidd & Hayden (2020) argue that the lack of standard definitions within the study of information seeking has harmed the field. Studies that are directly related to each other often use entirely separate vocabulary, creating the illusion that separate phenomena are being investigated. Kidd & Hayden (2020) therefore advocate for a definition of curiosity that is as broad as possible. In this spirit, I will include a diverse range of studies in this chapter, while also using some of the definitions of distinct sub-categories of information seeking behavior, as I will hypothesize that these behaviors may be driven by different neural processes. Firstly, in **chapter 3** I present an fMRI study investigating the role of medial frontal cortex (MFC) in information sampling to guide economic choice in an active information sampling task. **Section 1.2** in this chapter discusses the decision-making framework and the studies using it that led us

to predict this role of MFC. **Chapter 4** describes the development of a novel information sampling task used in two experiments in **chapters 5** and **6**, which aims to dissociate some of the neural signatures associated with information sampling, as described in **section 1.3** and attempts to address some of the shortcomings of the previous literature on information sampling discussed in **section 1.4**. In **chapter 5** I describe the results of a large-scale online study using this task to investigate the effects of psychiatric symptoms, in particular compulsivity, on the frequency of information sampling, as well as on the use of task features to guide sampling decisions. Finally, **chapter 6** discusses the results of an fMRI study using another version of the same task showing that activity in a widespread network in human frontal and parietal cortex represents goal proximity while making sampling decisions.

1.2 Information sampling in economic choice

Cognitive psychologists have studied information sampling in economic choice since the 70s (Newell & Simon, 1972; Payne, 1976). Padoa-Schioppa (2011) describes economic choice as finding the best option among choice alternatives that are different on at least one dimension of interest. For example, choosing between different apartments that are for rent when moving to a new city is an economic choice task. These apartments differ on multiple dimensions, such as price, size, and location, and the agent has to find a way to choose the best apartment among the options. In this section I will describe a decision-making framework that has been successful in explaining decisions in such scenarios, and in **chapter 3** I demonstrate that such a framework can also explain information sampling to guide these decisions.

1.2.1 Evidence accumulation

Evidence accumulation frameworks have been successful in explaining human decision-making behavior in simple binary economic choice tasks (Basten et al., 2010; Krajbich et al., 2010; Milosavljevic et al., 2010; Philiastides et al., 2010). In these frameworks, an agent continues to accumulate information from two choice alternatives until the evidence in favor of one of the options is high enough to exceed a pre-defined decision threshold (Figure 1.1; Ratcliff & McKoon, 2008). These frameworks account well for the fact that humans take longer to make difficult decisions, ie. decisions where the two choice alternatives are similar to each other (Busemeyer & Townsend, 1993; Hunt et al., 2012; Jamieson & Petrusic, 1977; Milosavljevic et al., 2010). Implicit in evidence accumulation frameworks is that these difficult decisions take longer to make, because more evidence needs to be sampled before a decision threshold is reached (Ratcliff & McKoon, 2008; Shadlen & Shohamy, 2016; Usher & McClelland, 2001). In other words, a choice is preceded by many sequential decisions between sampling or not sampling more information.

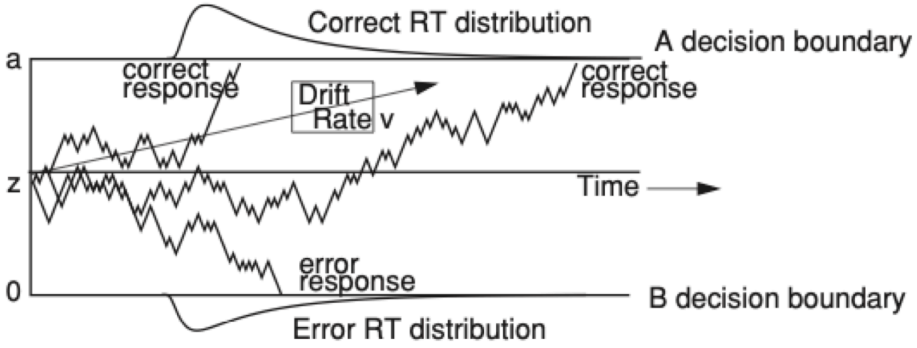


Figure 1.1. Evidence accumulation framework. Time is presented on the x-axis and evidence on the y-axis. The agent starts at point *z* on the y-axis, which is the mid-way point between two choice alternatives A and B. Over time, the agent collects evidence. Evidence in favor of option A moves the agent closer to the decision boundary for option A, while evidence in favor of option B moves the agent closer to the decision boundary for option B. When a decision boundary is reached, the corresponding choice alternative is selected by the agent. If A and B are similar, there will be much evidence in favor of each option, and it will thus take the agent a longer time to reach one of the decision boundaries. This can explain the well-established behavioral result that difficult decisions take longer to make (Busemeyer & Townsend, 1993; Hunt et al., 2012; Jamieson & Petrusic, 1977; Milosavljevic et al., 2010). This framework implicitly suggests that sequential sampling decisions precede economic choice.

1.2.1.1 Could information sampling be driven by evidence accumulation?

This is in line with the research on attention and choice. Eye gaze is the main way in which humans collect information about their environment and can therefore be a good measure of information seeking. Krajbich et al. (2010) found that evidence accumulation was biased towards the fixated option in a binary forced choice task, indicating that fixations on choice options likely correspond to information sampling of that option. Gaze predicts which option humans will choose in other studies as well (Cavanagh et al., 2014; Folke et al., 2017; Krajbich et al., 2010; Shimojo et al., 2003; Smith & Krajbich, 2019). Furthermore, this attention is driven by a process that prioritizes attending to goal-relevant information (Sepulveda et al., 2020).

However, this prediction that information sampling is an implicit prediction in the drift diffusion model and rarely an explicit feature of tasks used to study decision-making, as most economic choice tasks used in the experiments described present all

relevant information to the participants passively and all at the same time, rather than allowing them to sample sequentially, as evidence accumulation frameworks would predict.

In this thesis I have used two different economic choice tasks to address this problem. If it is the case that information sampling in economic choice arises from an evidence accumulation process, this evidence accumulation should be reflected in brain activity. Indeed, several studies have attempted to locate brain networks where activity correlates with the activity of evidence accumulators as predicted by models. In these studies, medial frontal cortex (MFC) has been identified as a potential 'evidence accumulator' in the brain (Gluth et al., 2012; Hare et al., 2011; Pisauro et al., 2017; Rodriguez et al., 2015; Venkatraman et al., 2009; Zhang et al., 2012). Such an evidence accumulator would be expected to drive future information sampling as long as a decision threshold is not reached. I have directly tested this prediction using an economic choice task where participants were able to actively sample information in **chapter 3**.

1.2.2 How does uncertainty affect information seeking?

Within this framework, the higher uncertainty is, the more the agent should sample preceding a choice. However, different relationships between information seeking and uncertainty have been reported. For example, Litman et al. (2005) showed that human participants' curiosity in a trivia questions task was highest when there was a small 'knowledge gap', meaning participants were close to knowing the answer, but were not completely certain about it. In contrast, curiosity was lower when the answer was either known or completely uncertain. Besides this, it is known that

high levels of novelty can induce avoidance and fear in animals, while they show the most curiosity for moderately novel stimuli (Bardo et al., 1996; Barnett, 1958; Berlyne, 1950; Hughes, 2007). This varying relationship between information seeking and uncertainty could be one way to dissociate different information seeking behaviors.

Besides this, people show biases that deviate from sampling from options with high uncertainty. Confirmation bias, for example, is the phenomenon whereby people sample evidence that confirms their beliefs (Hunt et al., 2016; Nickerson, 1998). Human participants have been shown to display confirmation bias even in simple perceptual decision-making tasks (Kaanders et al., 2022). More generally, humans have been shown to display Pavlovian approach, the enhanced learning of approach actions that lead to reward, as opposed to avoidance actions (Guitart-Masip et al., 2012). In information seeking, this manifests as a tendency to continue to sample from the currently preferred option (Hunt et al., 2016). Some of these biases may in fact be adaptive heuristics, as the environment we sample from usually contains many more choice alternatives than are available in the decision-making tasks we use to probe information seeking.

In the novel information sampling task I developed in **chapter 4**, I explicitly manipulate option uncertainty in an economic choice task with multiple choice options to investigate how this affects information sampling leading up to choice in **chapters 5** and **6**.

1.3 Dissociable mechanisms of information sampling in the brain

It is unclear to what extent curiosity in its broadest definition is driven by a single general brain network, or whether different subtypes of information seeking can be

dissociated in behavior and the brain. In this section I will summarize recent work aiming to better understand different types of information seeking: (i) through the development of reinforcement learning algorithms that display similar exploratory behavior to that seen in humans; (ii) the study of human participants' sampling behavior using experimental designs that dissociate information sampling from reward-maximizing choices; and (iii) the study of the neural mechanisms underlying these behaviors.

1.3.1 Is information inherently rewarding?

'Hard exploration' problems in machine learning are environments in which rewards are sparse and can only be obtained by performing a specific sequence of actions (Bellemare et al., 2016). Traditional reinforcement learning agents struggle to obtain any reward in this type of environment, as they fail to identify and learn the long action sequences necessary to discover these sparse rewards. Introducing *intrinsic* rewards for gaining information about the environment, that augment the externally delivered reward signal has improved performance of these agents on such problems, to the points where they outperform humans on tasks with sparse rewards, such as the Atari game 'Montezuma's Revenge' (Alet et al., 2020; Badia et al., 2020; Brafman & Tennenholtz, 2001; Burda et al., 2018; Ecoffet et al., 2019; Ermolov & Sebe, 2020; Ng et al., 1999; Raileanu & Rocktäschel, 2020; Sekar et al., 2020; Twomey & Westermann, 2018; Zheng et al., 2020).

1.3.1.1 Intrinsic curiosity

Humans and other animals seem to be inherently curious (Berlyne, 1954; Loewenstein, 1994). This idea of including an intrinsic reward for exploring suggests that we may have some innate preference for information seeking, which is in line with Berlyne's observation that animals exhibit perceptual curiosity and even consider novelty rewarding (Berlyne, 1955; Butler, 1953, 1957a, 1957b; Ennaceur & Delacour, 1988; Hughes, 1965, 1968, 2007; Nicki, 1970; Reed et al., 1996; Stolba & Wood-Gush, 1980; Wood-Gush & Vestergaard, 1991). For example, people are willing to forgo other rewards to gain information (Kang et al., 2009).

It has been shown that reinforcement learning agents can learn intrinsic rewards through experience, but that these can also be passed down to new agents via their reward function in an evolutionary manner (Singh et al., 2010; Zheng et al., 2020). Furthermore, many findings within developmental psychology suggest younger human learners show more exploratory behavior than older learners, who are typically more goal-directed (Gopnik, 2020). For this "explore-first/exploit-later" strategy to be present, some innate motivation in young agents to seek new information is to be expected. Indeed, infants have been shown to prefer stimuli of medium complexity (Kidd et al., 2012). This behavior is in line with Loewenstein's information-gap hypothesis, which argues that curiosity is at its highest when the agent already possesses a moderate amount of knowledge (Loewenstein, 1994).

1.3.1.2 Brain responses to novelty

That information is considered rewarding is also reflected in the brain. Dopamine cells, which are known to respond to a prediction error of future reward

(Schultz et al., 1997), have been shown to activate in response to novel stimuli in primates, regardless of their association with reward (Kakade & Dayan, 2002). Preferences for novel stimuli and curiosity have also been related to activity in brain regions associated with expected reward (Bardo et al., 1996; Hooks & Kalivas, 1995; Kang et al., 2009; Pierce et al., 1990; Taghzouti et al., 1985; Wittmann et al., 2008). Furthermore, variation in the type 4 dopamine receptor (DRD4) gene has in the past been related to individual differences in novelty seeking (Ekelund et al., 1999; Paterson, 1999).

1.3.1.3 Intrinsic rewards in animals

Recent discoveries in machine learning on hard exploration problems beg the question whether humans seek information in a similar way – that is to say, by an intrinsic motivation to seek out novelty. Recent studies have indeed shown that both non-human primates and human participants are willing to pay for information to reveal information that does not affect the outcome (Rodriguez Cabrero et al., 2019; Wang & Hayden, 2019), building on an established literature documenting preferences for early "temporal resolution of uncertainty" in such problems (Kreps & Porteus, 1978). By contrast, information sampling could also be guided in a goal-driven manner in order to optimize future decisions (Foley et al., 2017). Again, the goal here can either be to identify new sources of reward or to improve the agent's model of the environment, but most behavioral tasks used to investigate this type of information seeking focus on the former (Gottlieb & Oudeyer, 2018).

What is the benefit of treating information as inherently rewarding? As mentioned before, reinforcement learning agents show improved performance on

tasks where long action sequences are required to find reward when intrinsic rewards are introduced. This could also be an advantage in exploration by animals. Animals appear to learn cognitive maps of their environment, which are then used to simulate future outcomes (Behrens et al., 2018; Buzsáki, 2015; Diba & Buzsáki, 2007; Johnson & Redish, 2007; Kay et al., 2020; Tolman, 1948). If the brain is indeed constantly simulating the future, excluding future potential outcomes early on through information seeking may be rewarding, as this saves time and energy (Hunt et al., 2021).

Furthermore, curiosity has been shown to enhance memory such that the larger curiosity reduction (i.e. moving from a high curiosity state to curiosity relief by information revelation) the more incidental memory is enhanced and hippocampal activity increases (Jepma et al., 2012). Besides this, hippocampus has been shown to be a driver of novelty seeking as mice with lesioned hippocampi show less novelty seeking behavior (Misslin et al., 1981). So memory processes appear to be closely involved in driving information seeking behavior and in turn are modulated by uncertainty reduction as a result of information seeking. This interaction could be an important factor enabling learning. While the experimental tasks used in this thesis do not directly test the role of memory in information search, the novel information sampling task presented in **chapter 4** has allowed us to assess how the value of sampling completely new choice options compares to that of sampling options that have already been sampled previously.

1.3.2 Exploration vs exploitation

One of the best-described types of information seeking behavior is that shown in explore-exploit tasks (Blanchard & Gershman, 2018; Cohen et al., 2007; Kolling et

al., 2012; Navarro et al., 2016). In such studies, prefrontal cortex (PFC) activity has been found to predict exploratory choices of uncertain options (Fig. 1; Blanchard & Gershman, 2018; Daw et al., 2006; Domenech et al., 2020; Kolling et al., 2012; Tomov et al., 2020; Trudel et al., 2020). More specifically, Trudel et al. (2020) found that prefrontal subregions play distinct roles in resolving the exploitation-exploration dilemma: while dorsal anterior cingulate cortex (dACC) encoded uncertainty only when the agent was in an exploratory phase, ventromedial prefrontal cortex (vmPFC) encoded uncertainty throughout the task, but the polarity of its activity changed as a function of whether the agent was in an exploitative or explorative phase (Figure 1.2B). In our novel information sampling task (**chapter 4**), we have directly dissociated information sampling from reward-maximizing choice.

A challenge in these tasks is that it is hard to dissociate decisions to sample information, or 'exploration', from decisions to maximize expected reward, or 'exploitation'. Blanchard and Gershman (2018) addressed this by separating actions the agent can take into 'observation' and 'betting' actions and found larger dACC activity on observation trials (Figure 1.2A). This still presents information sampling and decision-making as a trade-off, though, while in reality this trade-off may not exist as agents continuously sample information to guide upcoming choices.

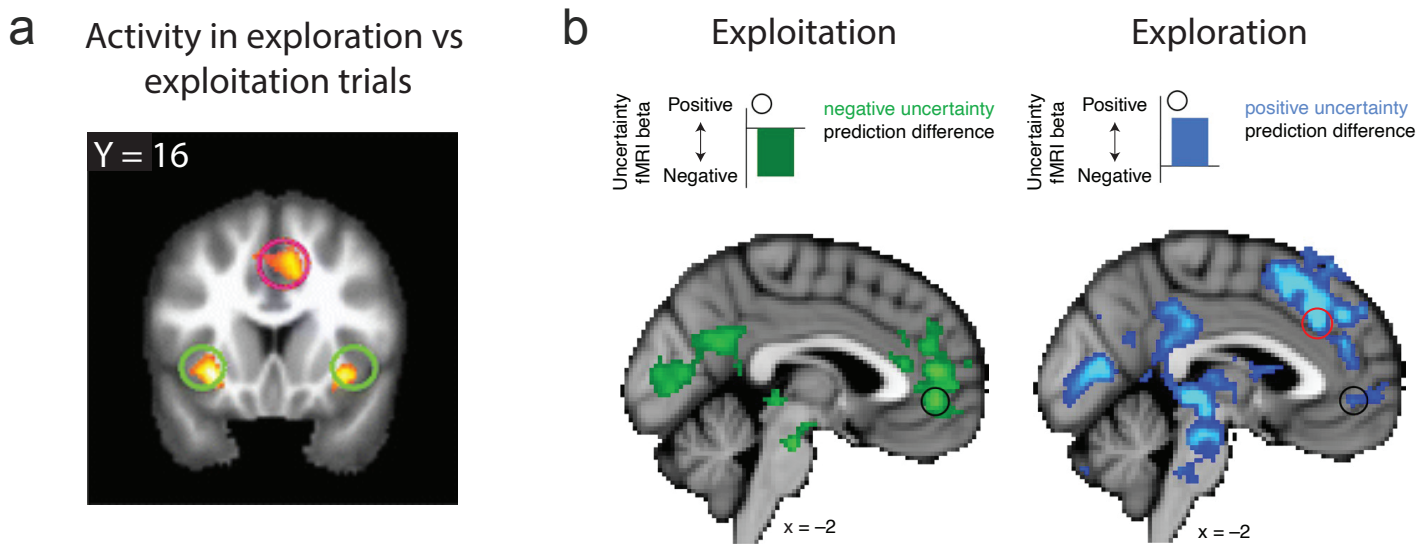


Figure 1.2. Activity in anterior cingulate cortex (ACC) is related to decisions to sample more information in exploitation-exploration tasks. (A) In a task where participants could choose between 'observing' and 'betting' on each trial, trials in which participants chose to 'observe', larger activity in dACC (red circle) and insula (green circles) was observed in an fMRI study (Blanchard & Gershman, 2018). (B) In another fMRI study, dACC only encoded uncertainty when participants were being exploratory (red circle, right panel), while vmPFC encoded uncertainty throughout the task (black circle), but uncertainty-related activity changed polarity depending on whether the participant was in an exploitative (left panel) or exploratory phase (right panel) (Trudel et al., 2021).

1.3.3 The value of information versus choosing to sample

1.3.3.1 Dissociable neural activity in information sampling tasks

Other recent studies have separated information sampling from choice by giving subjects the explicit opportunity to actively sample information before making a choice (Hunt et al., 2018; Kaanders et al., 2021; Kobayashi & Hsu, 2019; Wang & Hayden, 2020). In these tasks, prefrontal activity has again been found to be related to information sampling (Figure 1.3; Hunt et al., 2018; Kaanders et al., 2021; Kobayashi & Hsu, 2019; Wang & Hayden, 2020; White et al., 2019). Looking more carefully at the specific regions related to information sampling, there again appears to be a dissociation between activity in two regions: a more posterior and dorsal region, including pre-supplementary motor area (pre-SMA) and dACC, and ventromedial

prefrontal cortex (vmPFC; Figure 1.3A; Charpentier et al., 2018; Filimon et al., 2020; Iigaya et al., 2020; Kaanders et al., 2021; Kobayashi & Hsu, 2019; Neubert et al., 2015; van Lieshout et al., 2018).

1.3.3.2 Functional coupling of prefrontal sampling regions

At rest, these two regions show strong functional coupling with each other, but pre-SMA/dACC are also strongly coupled with sensorimotor regions, frontal pole, dorsal prefrontal cortex, and insula, among others (Neubert et al., 2015). On the other hand, vmPFC shows strong coupling with temporal cortex, posterior cingulate, precuneus and crucially, ventral striatum, which receives the strongest dopaminergic input from reward-sensitive parts of the dopamine midbrain (Neubert et al., 2015). In contrast, dACC is more strongly connected to dorsal striatum, which mostly receives inputs from motoric parts of the dopamine midbrain (Haber & Behrens, 2014).

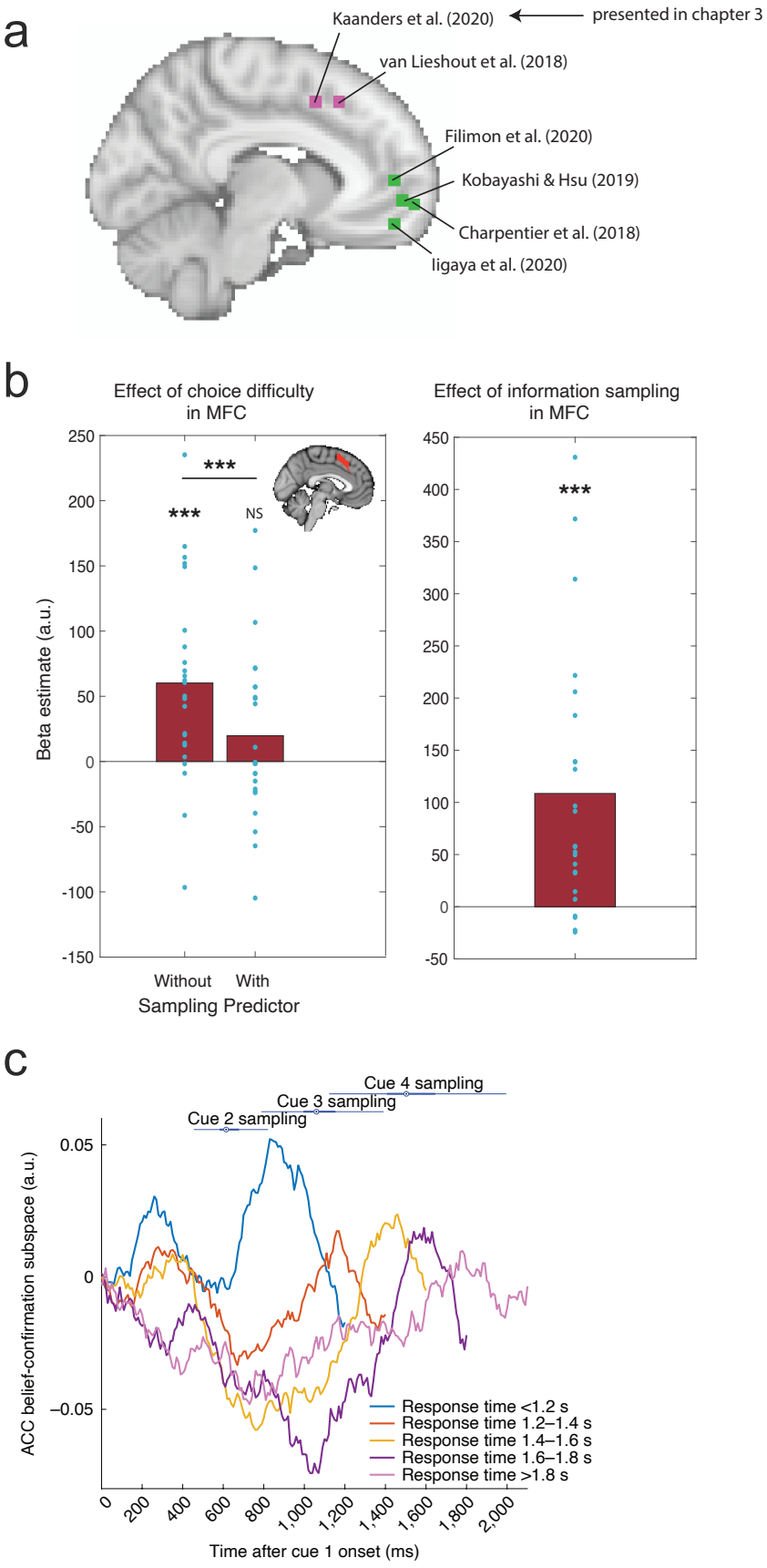

a
Kaanders et al. (2020)
presented in chapter 3
van Lieshout et al. (2018)
Filimon et al. (2020)
Kobayashi & Hsu (2019)
Charpentier et al. (2018)
Iigaya et al. (2020)
b
Effect of choice difficulty in MFC
Effect of information sampling in MFC
Beta estimate (a.u.)
***
***
NS
***
Without
With
Sampling Predictor
c
Cue 4 sampling
Cue 3 sampling
Cue 2 sampling
ACC belief-confirmation subspace (a.u.)
Response time <1.2 s
Response time 1.2–1.4 s
Response time 1.4–1.6 s
Response time 1.6–1.8 s
Response time >1.8 s
Time after cue 1 onset (ms)

Figure 1.3. Prefrontal activity is associated with information sampling in tasks where sampling is explicitly separated from reward-maximizing choices. (A) Locations of peak activity associated with information sampling in pre-SMA/dACC (purple) and vmPFC (green) in a number of recent human fMRI studies, collapsed across the x-axis (Charpentier et al., 2018; Filimon et al., 2020; Iigaya et al., 2020; Kaanders et al., 2021; Kobayashi & Hsu, 2019; van Lieshout et al., 2018). One of these studies (Kaanders et al., 2021) is presented in **chapter 3**. (B) An effect of inverse value difference on BOLD signal in medial frontal cortex (MFC) is significantly reduced when including information sampling as a co-regressor in an economic choice task. This study is presented in detail in **chapter 3** (Kaanders et al., 2021). (C) Macaque single-cell activity in dACC is related to how much newly revealed information confirms a previously formed belief about which of two options to choose to maximize reward in an economic choice task (Hunt et al., 2018).

This dissociation in functional coupling can be linked to the different features of the specific behavioral tasks used to find neural activity in these regions related to information sampling. In **chapter 3** I will show that pre-SMA/dACC activity is predictive of subsequent actions by the agent to gather more information to guide an upcoming choice (Kaanders et al., 2021). In contrast, vmPFC encodes the value of upcoming information (Charpentier et al., 2018; Filimon et al., 2020; Iigaya et al., 2020; Kobayashi & Hsu, 2019). For example, Iigaya et al. (2020) found an anticipatory value signal in vmPFC, which is the modelled value associated with the information that a reward will be received later on (Figure 1.4A-B). Similarly, Charpentier et al. (2018) found an adjacent region is activated in response to the opportunity to receive information about future outcomes (Figure 1.4C).

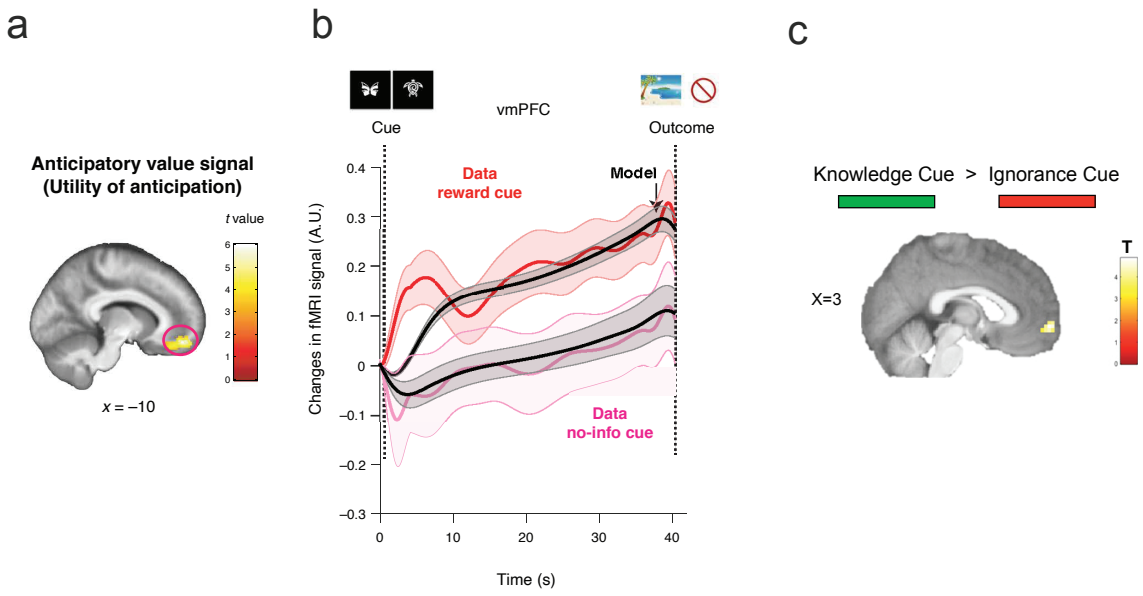


Figure 1.4. Activity in vmPFC is associated with the value of information. (A) BOLD activity in vmPFC is correlated with the anticipatory value of knowing about upcoming reward ahead of time (Iigaya et al., 2020). (B) The temporal evolution of this BOLD signal matches the anticipatory utility signal predicted by a reinforcement learning model that includes a preference for advance reward information (Iigaya et al., 2016, 2020). (C) vmPFC is more active during the delivery of informative cues than uninformative cues (Charpentier et al., 2018).

1.3.3.3 Information value encoding in dopaminergic system

More evidence for such a functional dissociation between these two regions in information sampling comes from the long literature on novelty encoding in the dopaminergic midbrain (Kakade & Dayan, 2002). Recent findings elaborate on this by showing that ventral striatum, which is strongly coupled with vmPFC, itself also encodes the value of information (Charpentier et al., 2018; Filimon et al., 2020; Kobayashi & Hsu, 2019; Lau et al., 2020; van Lieshout et al., 2018). In other words, while more posterior regions, such as pre-SMA and dACC have been shown to drive sequential choices to sample information to guide future choice, vmPFC and the dopaminergic system mostly seem to encode the value of that information but do not necessarily themselves drive behavior. As these regions are strongly coupled with

each other as well, it is likely that pre-SMA/dACC uses value representations in vmPFC to drive sampling. This is in line with broader accounts of the roles of these regions in decision-making in general: vmPFC is often considered to encode the value of stimuli (Padoa-Schioppa & Assad, 2006), while pre-SMA/dACC have frequently been described as driving behavioral shifts (Hare et al., 2011; Kolling et al., 2016; O'Reilly et al., 2013).

1.3.3.4 Parietal involvement in information sampling decisions

Activity in the intraparietal sulcus (IPS) has also been found in some paradigms to be related to sampling, to predict the moment of information gain, or outcome uncertainty (Duan et al., 2020; Horan et al., 2019; Iigaya et al., 2020; Kaanders et al., 2021; van Lieshout et al., 2018). Both IPS and dACC are part of a network often identified in decision-making tasks, responding to outcomes different from the agent's predictions that require belief updating and behavioral shifts (Akam et al., 2021; Isoda & Hikosaka, 2007; M. P. Karlsson et al., 2012; Kolling et al., 2016; Molenberghs et al., 2007; O'Reilly et al., 2013; Schuck et al., 2015). More specifically, IPS seems particularly sensitive to surprising changes in the environment, while dACC activity is more related to the actual updating of beliefs that leads to behavioral change (O'Reilly et al., 2013). This suggests this network might use estimates of uncertainty to drive information sampling for decision-making. In **chapters 3** and **6**, I discuss the results from fMRI studies using decision-making tasks with active information sampling in which we observe IPS activity predicting upcoming sampling decisions.

1.3.4 Individual differences in information search

Individual differences in novelty seeking have been observed and investigated for decades (Cloninger, 1987; Hughes, 1965; Keltikangas-Järvinen et al., 2004). Specifically, it has been shown that individuals with increased novelty seeking also show increased sensitivity to drugs of abuse (Bardo et al., 1996; Bradberry et al., 1991; Deroche et al., 1993; Hooks et al., 1991; Hooks et al., 1991; Piazza et al., 1989). The rewarding sensation caused by these drugs is induced by a dopaminergic response (Le Moal & Simon, 1991; Nemeroff & Kalivas, 1988), similar to that produced in response to novelty

As the dopaminergic system is thought to play a role in many psychiatric conditions (Belujon & Grace, 2017; Bogerts et al., 1983; Dong et al., 2020), it is likely that individual differences in information search might also relate to mental health. Indeed, Hauser et al. (2017) showed that individuals with a high level of obsessive-compulsive traits sampled more information in a sequential information gathering task. By modelling this behavior they showed this was due to a raised decision threshold, thus requiring compulsive individuals to sample more evidence before they could commit to a choice. This account differs somewhat from traditional accounts of compulsivity, which describe compulsivity as the tendency to exhibit more habitual actions, which are not guided by an internal model of the environment (Gillan et al., 2011). This habitual account of compulsivity would predict repetitive sampling behavior that does not flexibly adapt in response to changing demands in an information sampling task, whereas Hauser et al.'s (2017) results also show increased sampling behavior, but this behavior does arise from a decision-making process similar to that in non-compulsive individuals. In **chapter 5** I discuss the findings from a large-scale

online study in which we used our own novel information sampling task to test these predictions.

1.4 New approaches and challenges to studying information sampling for decision-making

1.4.1 Need for more complex tasks

Behavioral tasks used to investigate information sampling in decision-making are typically highly simplified, often with only two available choice options, and the amount of information that can be sampled to inform choice therefore also tends to be small (Gottlieb & Oudeyer, 2018). This approach has been crucial in characterizing decision-making mechanisms from behavioral and neural data (Glimcher & Fehr, 2014; Gold & Shadlen, 2007), but it is less useful to begin to disentangle the many potential drivers of information seeking, their neural representations and how they affect choice (Gottlieb & Oudeyer, 2018). While we have argued that there is a functional dissociation in frontal cortex between signals that encode information value and those that directly drive sampling, our understanding of the different factors and their neural representations that lead an information source to be considered more or less valuable is poor. To improve it, the tasks we use need to better reflect the large amount of goals, choice options and information available in the world.

1.4.2 Use of priors in sampling and decision-making

This is easier said than done. Firstly, human agents do not just sample information from their environment, they also use their prior knowledge of the world to make decisions or learn new tasks (Spelke & Kinzler, 2007). Without a full account of

this prior knowledge, it is difficult to identify what exactly drives information seeking when faced with a new task. This is also an important distinction between humans and RL agents: the latter typically do not enter a new task with such prior knowledge. Dubey et al. (2018) found that gradually removing this prior knowledge from human agents (by modifying the visual information in a video game environment), drastically reduced performance, suggesting prior knowledge is a very important factor in effectively driving sampling to solve new problems. Another way in which human participants obtain such prior knowledge is through the instructions given by the experimenter before starting the task. Human participants perform much better at a novel task when given prior information in the form of instructions or by watching another agent play the game, which could be one of the main reasons why human agents learn new tasks much faster than RL agents (Tsividis et al., 2017).

The ability to construct and maintain a complex model of the world is likely crucial to the effective driving of information seeking for learning and solving new complex problems, and could be the main reason human agents struggle less with 'hard exploration' problems than their RL counterparts. Ideas derived from RL algorithms have previously been successfully applied to study human behavior and neural data (Niv, 2009). The knowledge that the performance of RL agents on 'hard exploration' problems is dramatically improved by including intrinsic motivation to seek out new information, could be used in a similar fashion to better understand human information sampling. This is unlikely to be a very good model, though, as human agents probably also use their model of the world to guide sampling of the information likely to be most valuable in the task, which is hard to replicate in an RL algorithm. As such, the presence of such a model in human agents might severely limit how much

we can learn about human information sampling from studying RL agents. However, it is not inconceivable that some elements of a model of the world might effectively be introduced into AI algorithms, such as intuitive physics and psychology (Lake et al., 2017), which could lead to behavior that approximates that of humans.

1.4.3 Future directions

Progress along these lines will likely come from two sides: First, human studies need to be suitably complex in the stimulus or action domain (e.g. by using many possible stimulus dimensions, or using video game-like tasks with many possible actions). This will allow researchers to harness recent advances in machine learning algorithms that specifically target good compressions of information into a useful state space to enable efficient learning (Nair et al., 2018). In doing so, researchers can use tasks and stimuli that are novel to participants. Second, researchers will likely benefit from the recent advances in 'meta-learning' (Wang et al., 2018). Here, more complex tasks allow researchers to explicitly derive normative (or at least performative) exploration algorithms directly from data without the need to define them by hand. For instance, Zheng et al. (2020) encoded knowledge about likely task structures in the intrinsic reward signal, showing that this enabled the agent to behave efficiently in novel environments. Such work will allow researchers to interrogate learned curiosity signals and use these insights to design human experiments or to compare human performance against a largely assumption-free, yet performative baseline. Taken together, these tasks would yield predictions about both "what" to sample (through the learned state space) and "whether" to sample (through the learned exploration or

intrinsic reward function) and would recover some of the tractability often lost when introducing more complexity into task design.

In this thesis, I have started moving in this new direction by developing a novel decision-making task in which participants are allowed to actively sample information from choice alternatives before committing to a choice. In addition, several task features are manipulated to enable us to investigate their effect on information sampling, such as sampling cost, volatility of option value and time horizon. The development of this task is discussed in **chapter 4**. **Chapter 5** discusses the findings from a large-scale online study using a version of this task to assess the effect of compulsivity on information sampling. **Chapter 6** discusses an fMRI study conducted using a version of this task that will enable us to study the neural signatures of different drivers of information sampling as well as the hypothesized dissociation between neural mechanisms driving sampling discussed above.

1.5 Conclusion

In this chapter I have outlined some previous literature on information seeking in relation to decision-making. Firstly, I have discussed the varying terminology used to refer to different types of information seeking behavior. I then used a widely used decision-making framework, accumulator frameworks, to hypothesize how information sampling may be driven by the brain.

I have highlighted some recent findings from a range of behavioral tasks studying different types of information sampling. In these studies, prefrontal activity is often related to information sampling, but there appears to be a functional dissociation

between more posterior regions, such as pre-SMA and dACC, and the more anterior vmPFC. We argue that pre-SMA/dACC drives the agent to sample more information before committing to a decision, while vmPFC activity encodes the value of upcoming information, but does not directly affect decisions to sample. The functional connectivity profiles of these regions, with pre-SMA/dACC being strongly connected to sensorimotor regions and vmPFC to the reward-sensitive dopaminergic system, support this hypothesis. **Chapter 3** describes the findings of an fMRI study demonstrating the involvement of medial frontal cortex in information sampling in economic choice.

What remains unclear is what the different drivers are of information seeking and how they are represented in the brain. In a number of the findings described here, the information sought by agents was non-instrumental to the decision at hand, suggesting information seeking is often driven by goals other than those set by the experimenter. In **chapter 4**, I describe the development of a novel task that directly addresses these issues. **Chapters 5 and 6** discuss findings from a large-scale online study and an fMRI study illustrating how information sampling has different drivers, how these are related to psychiatric symptoms and how these are represented in the brain.

Finally, I propose that to better understand how these information representations arise, we must develop behavioral tasks that better reflect the real decision-making problems humans face. RL may help here, but its use may be limited as RL agents do not possess the complex representation of the world we live in that we use to make decisions every day.

Chapter 2: Methods for analyzing fMRI data

Abstract

Two experiments presented in this thesis use functional magnetic resonance imaging (fMRI) to investigate the neural signatures of information sampling. In this chapter I will discuss two analysis methods I have used to analyze this data: mass univariate analysis and representational similarity analysis (RSA). In a mass univariate analysis, we design a general linear model containing explanatory variables from our task to predict neural activity in response to events and conditions in the task, which produces statistics at each voxel of the brain or within a region-of-interest (ROI). By contrast, RSA employs (dis)similarity measures to assess the similarity in neural responses between task conditions to build up a representation dissimilarity matrix (RDM) reflecting which task features activity in a particular brain region represents. I will also discuss some of the recent debate on how useful the later approach is to analyze fMRI data of prefrontal cortex (PFC).

Functional magnetic resonance imaging (fMRI) is used to study the neural signatures of information sampling in two of the chapters in this thesis. fMRI detects the level of oxygenation in blood in the brain, which is a measurement of blood flow as oxygenated blood flows to regions in the brain that are expending more energy (Logothetis et al., 2001). This is also referred to as the hemodynamic response and the MRI contrast used to obtain these measurements is the blood-oxygen-level dependent (BOLD) contrast. In this chapter, I will discuss the analysis methods used

in this thesis to relate these measurements of blood flow to behavior on decision-making tasks.

Firstly, I will discuss model-based fMRI analysis (**section 2.1**). In this section, I will explain how general linear models (GLMs) can be used to link parameters from the task and behavior to the imaging data, and how this can be used to test hypotheses. Secondly, I will discuss representational similarity analysis (RSA; **section 2.2**), a method to link task parameters to patterns of brain activity, as well as across species or methodologies, as well as some recent work using this method.

2.1 Model-based fMRI analysis

2.1.1 General linear model

In collecting an fMRI dataset from a participant completing a decision task inside the scanner, we obtain a measure of blood oxygenation at each voxel (a 2x2x2mm box) in the brain at each measured timepoint (defined by the repetition time of the MRI sequence used) for the duration of the task. To link these measurements to our task parameters, we need to model the hypothesized hemodynamic response to these parameters and fit this model to the brain data. To do this, we use the general linear model (GLM). In this model the BOLD signal in each voxel is the dependent variable, which is predicted by a linear combination of predictors, or explanatory variables (EVs), which can be task parameters such as stimuli, or behavioral responses such as button presses. As we typically use the same model to predict the signal at every voxel in the brain (or within a region of interest), this is also referred to as a 'mass univariate' analysis. Using this model, many different statistical tests can be run on the data by specifying a particular design matrix.

Mathematically, the GLM can be described as follows:

$$Y = X\beta + \varepsilon$$

where Y is the single voxel activity, X the design matrix, β the coefficients describing the weight of each EV in the model and ε the residual error (Jenkinson & Chappell, 2018).

A simple example is a model with a single EV, say, stimulus onset. Y is a vector containing the BOLD signal of a single voxel over time. X is the design matrix, in this case a vector specifying at which timepoints a stimulus was presented, with 1's at the timepoints a stimulus was presented and 0's at all other timepoints. We expect a hemodynamic response to our stimuli, rather than just an immediate increase in BOLD activity, so we need to model the shape of this response as well. This is done by multiplying our EV, the vector containing each stimulus onset, with the hemodynamic response function (HRF). As you can see in Figure 2.1A, this produces a timeseries which can be fit to the measured voxel time series. The parameter that is estimated in fitting the model is β, the unknown parameter, such that $X\beta$ best fits the data (Figure 2.1B), or in other words, that minimizes the sum of the squared residuals, ε (Winkler et al., 2014). ε then describes the difference between the predicted response and the actual response at each time point.

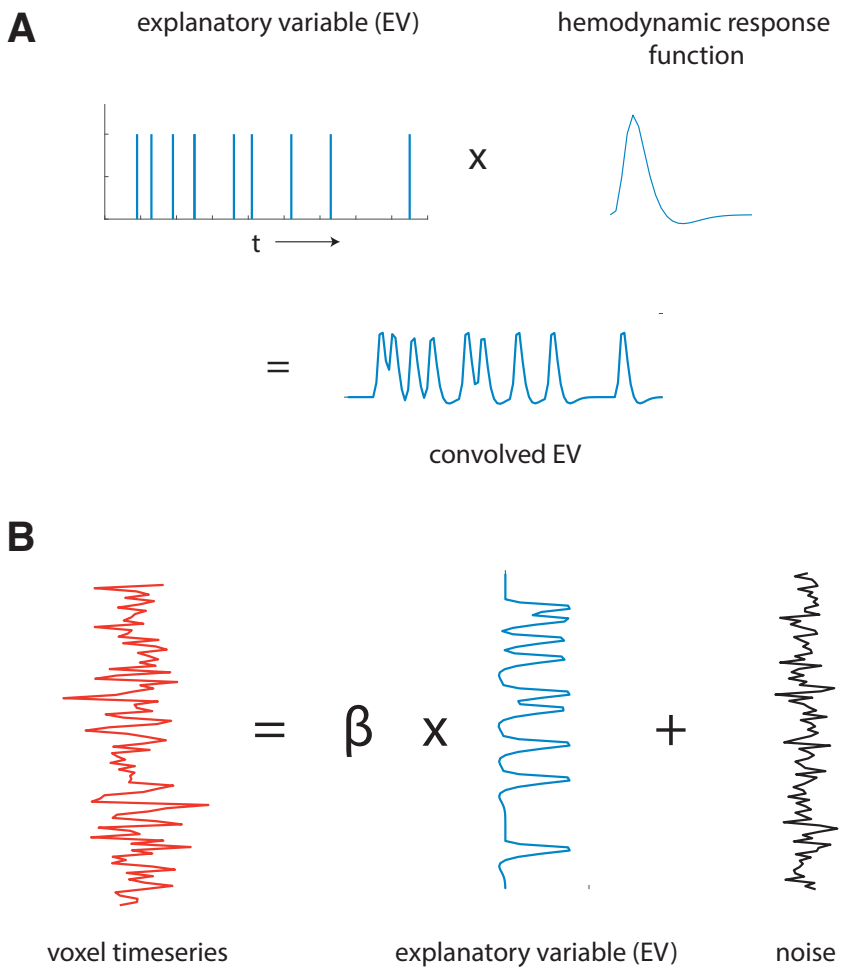


Figure 2.1 We can make regressors by convolving event timings with the HRF to produce timeseries predictors that can be regressed onto the BOLD signal timeseries for each voxel in the brain. (A) Convolving event timings described as stick functions at the time of event onset with the HRF produces a regressor with realistic predictions of activation for a voxel that activates in response to the described event. (B) In fitting a GLM we are estimating the free parameter β, such that noise is minimized.

Of course we usually want to predict the brain response to our task in terms of multiple features of the task and the participant's behavior. We can do this by expanding the design matrix to include multiple EVs. In the simple example, a dummy variable, stimulus onset, was used as a predictor of the BOLD signal (also referred to as a 'constant term'). However, we can also include continuous EVs (also referred to as 'parametric modulators'), such as value. Say each stimulus corresponds to a particular value on the task between 0-10. Then we can include a predicted response

to stimulus value in our design matrix, described by a column containing the value of the stimulus at the time it is presented (Figure 2.X). We would typically 'demean' this regressor (make it run from -5 to +5, rather than 0 to +10). This does not affect the coefficient for the regressor itself, but instead ensures that the constant term reflects the mean response to all trials; it is equivalent to orthogonalizing the parametric modulator with respect to the constant term.

In the analyses presented in chapters 3 and 6, I have used GLMs with many EVs to model brain activity in response to the task. Besides including EVs describing events, such as stimulus onset, at the time when they occur, we can also include EVs describing hypothesized brain activity leading to a response. As I am interested in investigating the neural signatures of information sampling, I can include an EV describing the predicted hemodynamic response leading up to a decision to sample information. In practice this is described by another column in the design matrix with 1's at the timepoints marking the start time of information sampling phases on the task.

The β that is estimated for an EV at a single voxel describes how much the predicted response (the waveform obtained from multiplying a predictor with the HRF) fits the measured BOLD signal at that voxel (Figure 2.2).

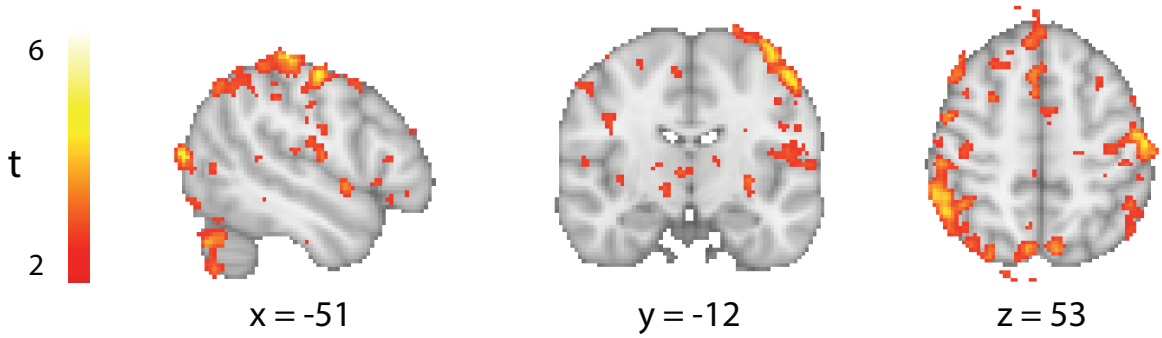


Figure 2.2 T-statistics at each voxel for an EV describing a button press with the right hand in a single participant (shown in a standard brain). The t-statistics at each voxel are computed by dividing the parameter estimates by their standard errors. Only t-statistics > 2.0 are displayed here. MNI coordinates of the displayed slices are shown alongside the maps.

To test specific hypotheses, we can define contrasts, which are linear combinations of parameter estimates. For example, to test whether a single EV has an effect on the BOLD signal compared to baseline, our contrast is just equal to the parameter estimate β for this EV. But we can also define contrasts that compare different conditions in our task. For example, say we want to compare brain activity in response to pictures of faces (EV_1) to that in response to pictures of houses (EV_2). Then the contrast of parameter estimates (COPE) becomes:

$$COPE = 1 \times \beta_1 - 1 \times \beta_2$$

2.1.2 Higher-level analyses

The model described so far can be used to fit our predicted responses to the task and the participant's behavior to the BOLD activity timeseries of a single voxel. This model is fit to activity from each voxel in the brain to obtain βs for each predictor

for each voxel. However, we are not so interested in the brain response to the task in a single participant; we want to know how our entire sample responds to the task on average. For this reason, we combine first-level analyses, which analyze each participant's data as described above, into higher-level analyses (Woolrich et al., 2004).

Two approaches are typically used to combine first-level analyses: a fixed-effects analysis or a mixed-effects analysis. A fixed-effects analysis simply takes the average of the lower-level estimates of β to compute the group effect. However, the only variance that is taken into consideration in computing the group average in this way is the within-subject variance. To be able to infer what the effect β is in a population, rather than just in our sample, we also have to take into account between-subject variance. This can be done with a mixed-effect analysis in which a second GLM is run which estimates the population mean β by taking into account both between-subject variance and within-subject variance. To do this group analysis in chapter 3 and 6 we have used FLAME (FMRIB's Local Analysis of Mixed Effects), which carries up first-level within-subject variances up to the higher-level analysis (Beckmann et al., 2003; Woolrich et al., 2004). T- and z-statistics for β_g can be obtained dividing β_g by its standard error (Figures 2.2-2.3). We can use multiple higher-level analyses to first combine across different fMRI sessions within each participant and then to combine all participants (Figure 2.3), as shown in chapters 3 and 6.

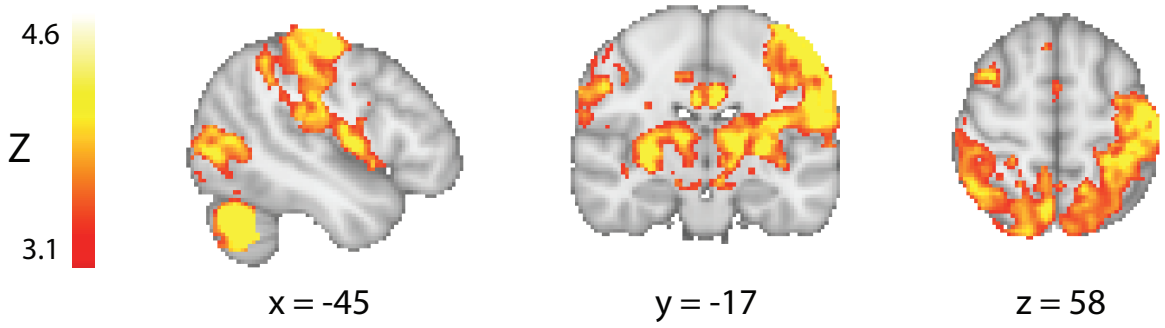


Figure 2.3 Z-statistics at each voxel for an EV describing a button press with the right hand. The z-statistics at each voxel are obtained by standard statistical transformation from t-statistics. Z-statistics are thresholded at > 3.1. MNI coordinates of the displayed slices are shown alongside the maps.

2.1.3 Significance testing

We can use the GLM to test hypotheses about brain activity in response to the task. In chapter 3, my hypothesis was that MFC activity predicts information sampling decisions. We can model this using a GLM by including an EV describing the predicted response as a function of the timepoints at which sampling phases in the task start. By running this GLM at each voxel, and then performing a higher-level analysis across participants, we obtain z-statistic maps for the information sampling EV. There are many regions in the brain with z-statistics larger than zero. However, we want to know in which of these regions the relationship between the BOLD signal and future information sampling is statistically significant. We can do this by thresholding at a particular p-value. Because we ran many analyses (a GLM in each voxel), we cannot simply use $p<0.05$, as it is likely that of the many voxels we are analyzing for some voxels the result will be significant by chance (in 5% of voxels if we use a threshold of $p<0.05$). Therefore, we must correct the threshold used. There are several approaches by which multiple-comparisons correction can be implemented.

One approach is to perform cluster-based inference. We first make sure that the original z-statistic image is family-wise error (FWE) corrected, meaning that the whole brain family wise error rate is controlled such that the probability of one or more false positives (voxels showing a statistically significant effect that is not truly there) is <0.05. Then we can use a cluster-forming threshold to mask the z-statistic image (Woolrich et al., 2004). A typical value to use for this is, for example, z=3.1. We are then left with some clusters (groups of adjacent voxels where z>3.1) and have eliminated single voxels with high z-statistics likely to be noise. The validity of this approach has recently been questioned in an investigation that showed cluster-based inference can lead to high false-positive rates due to differences between actual and assumed spatial autocorrelation between voxels (Eklund et al., 2016).

A different method to define clusters in our data that does not involve setting an arbitrary z-threshold is threshold-free cluster enhancement (TFCE; Smith & Nichols, 2009). This approach uses an algorithm that enhances the intensity of voxels that are 'cluster-like', meaning that the values of its surrounding voxels are similar. This approach allows us to look for clusters in the data, as this is how we expect activity in response to our EVs to occur, without having to set arbitrary thresholds that can potentially have a large effect on our results.

In many cases we will have a hypothesis about a specific part of the brain. It then makes sense to only run the above analyses on the brain region we are interested in, the region of interest (ROI; Poldrack, 2007). By doing so, our FWE controlled significance threshold is also less stringent, as we are running fewer models.

This way of analyzing fMRI data predicts one output variable at a time: the BOLD signal at one voxel. We fit the GLM to all recorded voxels in parallel to make

inferences about how different brain regions respond to the task. This is referred to as *mass univariate analysis*, 'mass' referring to the fact that the same model is fit to many data sources (voxels) and 'univariate' signifying that each individual analysis predicts only one variable (activity in one voxel).

Our hypotheses usually make predictions about brain responses in particular areas of the brain, not that of individual voxels. We can also predict *patterns* of activity in a brain region using multivariate techniques, which take into account the activity in all voxels within an ROI rather than just statistically significant activity (Haxby, 2012). In this way, activity patterns can be compared across conditions and through this we can infer what a brain region is encoding. An example of a multivariate method we used in chapter 3 is discussed in the next section.

2.2 Representational Similarity Analysis

Representational similarity analysis (RSA) is an analysis technique in which activity patterns are compared across conditions. This is done by determining the similarity between two sets of data associated with different conditions. We are typically not interested in the absolute activity in a brain region, but rather in how activity levels relate to each other in response to different conditions, for which RSA is especially useful. Furthermore, by projecting data into a common space of representational content, RSA facilitates comparisons between (imaging) modalities, participants, species and brain regions.

2.2.1 Representational geometry

RSA relies on measurements of the representational geometry of data. Thinking of neural activity in terms of representations of information is useful in describing how neural activity can produce behavior (Kriegeskorte & Kievit, 2013). It is thought that the population of neurons in a particular brain region represents information in the form of a neural population code, meaning that the *pattern* of activity in a region represents meaningful information, rather than the activity of individual neurons (Averbeck et al., 2006). The large number of neurons then allows for a wide range of possible activity patterns. Each neuron can then be thought of as a different dimension in the representational geometry of a brain region, and its activity in response to a stimulus as a different point in space (Kriegeskorte & Kievit, 2013). In the context of the experiments presented in this thesis, for example, we are interested in brain representations of information value of stimuli, as well as representations of belief updates in response to new information.

2.2.2 Similarity measures

There are many methods that can be used to determine the (dis)similarity between two datasets (Walther et al., 2016). In chapter 3, we used Pearson correlation (*r*), which was also used in Hunt et al. (2018), of which our study is a human replication. The data from each of two conditions can be defined as row vectors **a** and **b** extending from the origin (Fig 2.4). *r* is then the cosine of the angle between **a** and **b**, after normalization (which makes sure the distance measure is not changed when scaling the vectors). The higher *r*, the higher the correlation between two datasets, and thus

the more similar they are. If we want an indication of *dissimilarity* instead, we can simply use 1-*r*.

$$r_{ab} = \cos\theta = \frac{\boldsymbol{a}\cdot\boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|}$$

The Euclidean distance, on the other hand, is the actual distance between two vectors, while the Mahalanobis distance is the distance between noise-normalized response patterns (Walther et al., 2016). Euclidean/Mahalanobis distance are not affected by changes in response baseline, and Pearson correlation is not affected by response scaling (Figure 2.4).

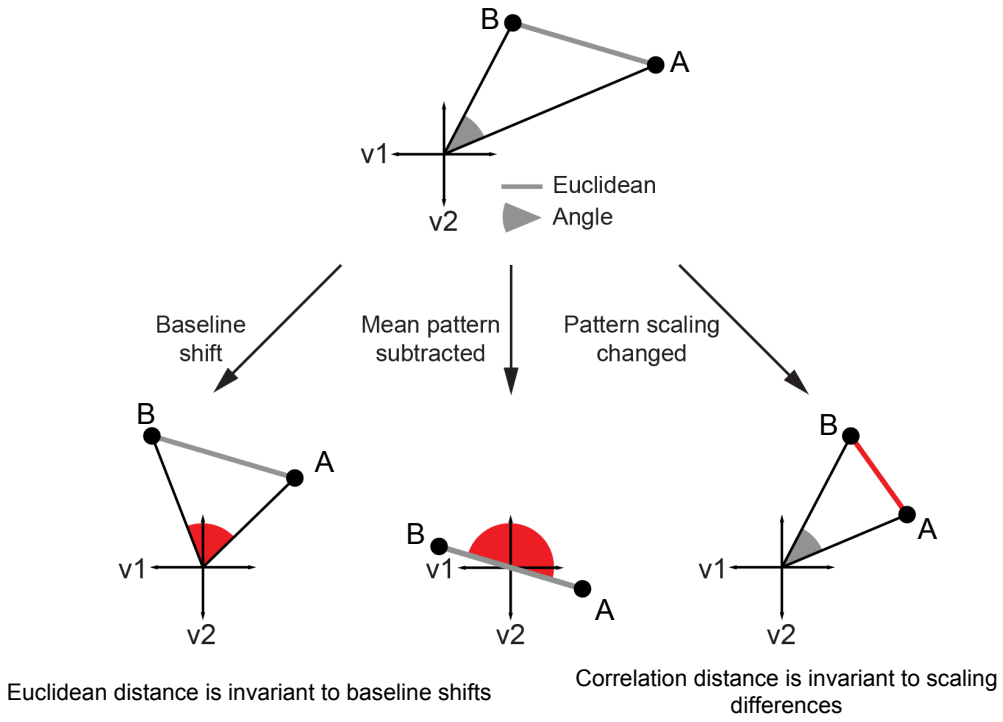


Figure 2.4. Illustration of how similarity between two vectors can be computed in different ways. Pearson correlation is derived from the angle between them, while Euclidean distance is the distance between them in fMRI pattern space (where each spatial dimension reflects activity of a voxel). Reproduced from Walther et al. (2016).

2.2.3 Analysis

Knowing how to compute the dissimilarity between two sets of data, we can compute this for each possible pair of conditions in our task. From this I can build a *representational dissimilarity matrix* (RDM), a summary matrix of how different representations of stimuli are from one another within the full dataset. RDMs are often visualized as two-dimensional matrices of colored pixels, where the color of each pixel represents the dissimilarity between the two conditions on the x- and y-axes. The data used can be brain activity in response to each of the different stimuli, or behavioral data or even output from a model or neural network.

For example, to analyse human fMRI data in this way, we can estimate the activity in response to different conditions using the conditions as EVs in a GLM as described in section 2.1 (Figure 2.5). Rather than deriving an RDM from neural activity from a specified ROI, we can also look for brain regions across the whole brain where neural activity is most similar to a specified RDM template in a *searchlight analysis*, as shown in chapter 3.

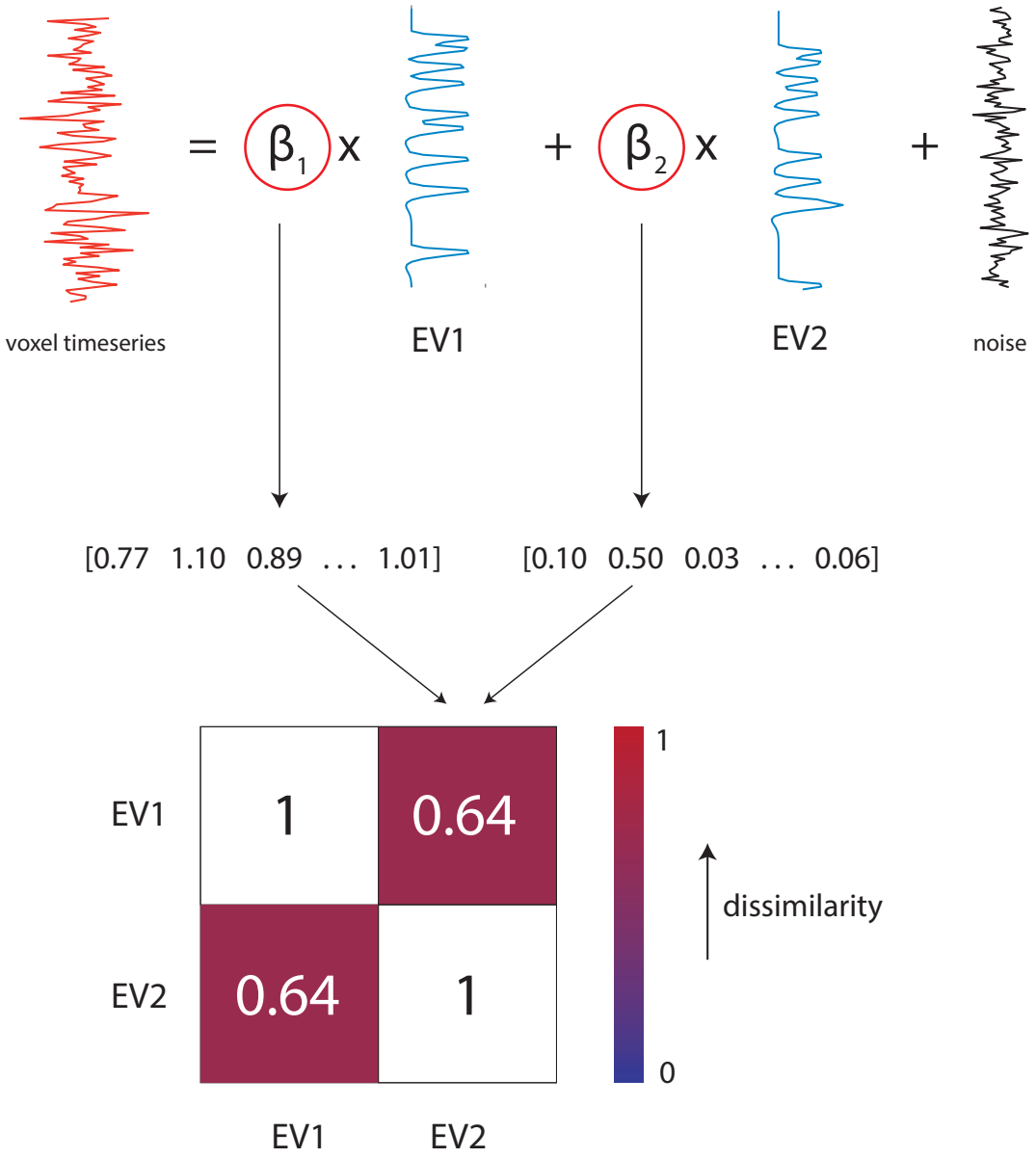


Figure 2.5 Simplified example of how to do RSA in fMRI data. Using a GLM in which EVs capture different task conditions, parameter estimates for the β's in this model can be computed for each voxel. From this we can derive an RDM by computing the dissimilarity between each pair of EVs (using Pearson correlation for example).

2.2.4 What are the advantages of this approach?

Describing neural activity in response to conditions in a task in this way makes comparing it to different datasets very easy. Say we want to compare data from single-cell recordings in macaque monkeys to human fMRI data. To do this in another way,

we would have to find some way to map one dataset onto the other, which is not straightforward as data points usually have completely different meanings when collected using different methodologies (e.g. neurons vs voxels). If we instead describe the data in the form of an RDM, the RDMs can just be compared to each other through correlation, as RDMs are simply matrices consisting of numbers without units. Furthermore, this approach allows us to directly compare models of neural activity by designing template RDMs and correlating these with the RDMs derived from data. Finally, RSA is able to reveal subtle differences between representations, which a simple classifier cannot (Figure 2.6).

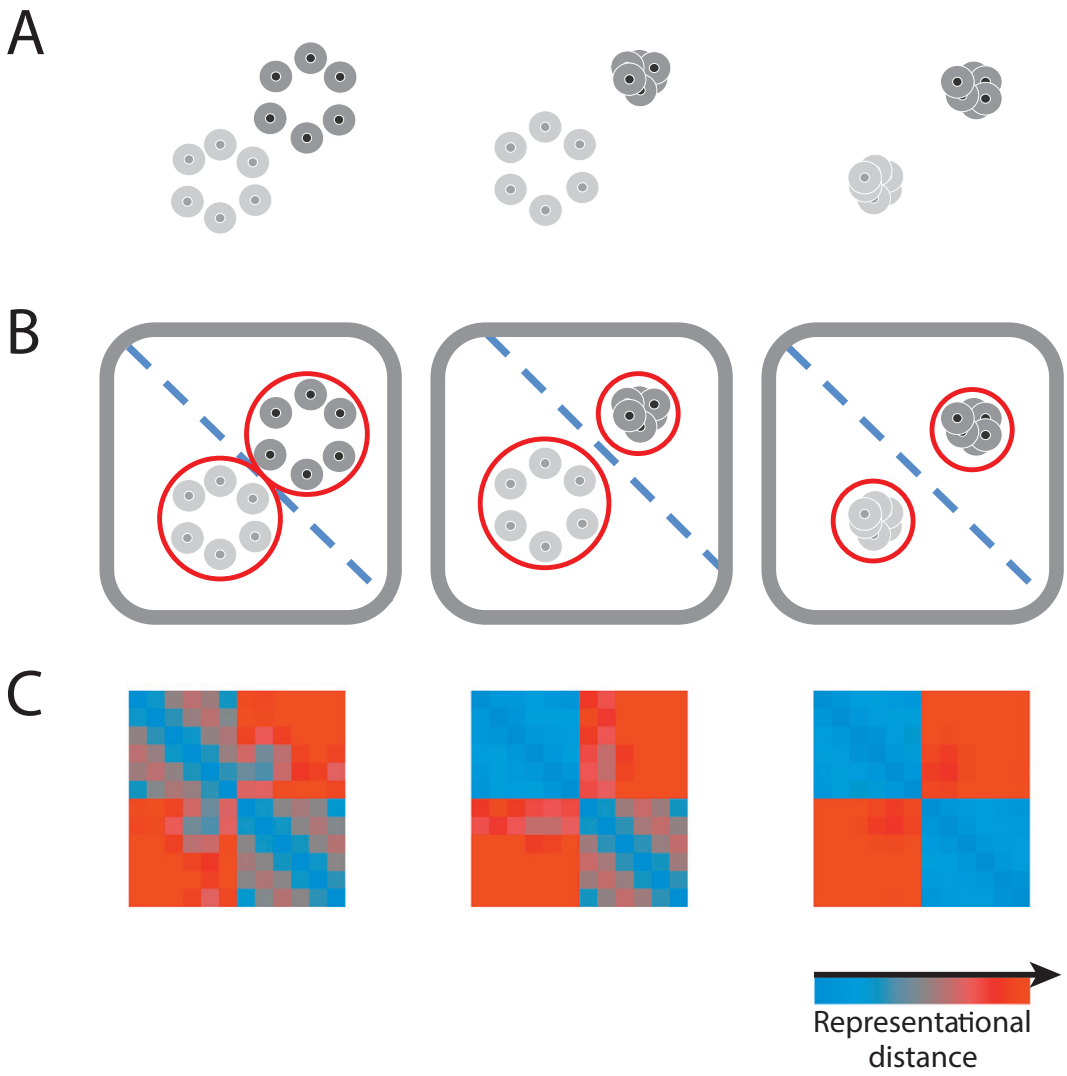


Figure 2.6 RSA can tell apart subtle differences in neural representations of information which a simple classifier cannot. (A) Neural activity in response to 12 different stimuli represented as points in space. (B) A classifier would give a similar result for all three patterns of activity, and would return a binary response to the two classes of stimuli. (C) RDMs can reveal subtle differences in activity patterns while also still showing the clear categorical difference between the two stimulus sets. Reproduced from Kriegeskorte & Kievit (2013).

2.2.5 How is RSA used in neuroscience?

RSA has a long history in psychology in determining the internal representations of stimuli from behavioural responses (Shepard & Chipman, 1970). It has particularly been used to study response generalisation to similar stimuli in

mathematical psychology (Shepard, 1987). More recently, RSA has been most commonly used to dissociate different brain regions in the visual stream. Specifically, Kriegeskorte et al. (2008) showed that activity in inferior temporal (IT) cortex in response to visual stimuli represented stimulus category in both macaque single-cell recordings and human fMRI data (Figure 2.7). Furthermore, they showed this representation was distinct from representations in human early visual cortex, where categorical representation was weak (Figure 2.8). This is a good example of how RSA can be used to show the functional role of a brain region across species and imaging modalities. In a later publication, RSA was used to compare different computational models of object recognition in terms of how similar their representational geometry was to that in IT cortex (Khaligh-Razavi & Kriegeskorte, 2014). They found that better performing models carried more similar representations to those in IT and that a deep convolutional network trained using supervised learning showed the most similar representation.

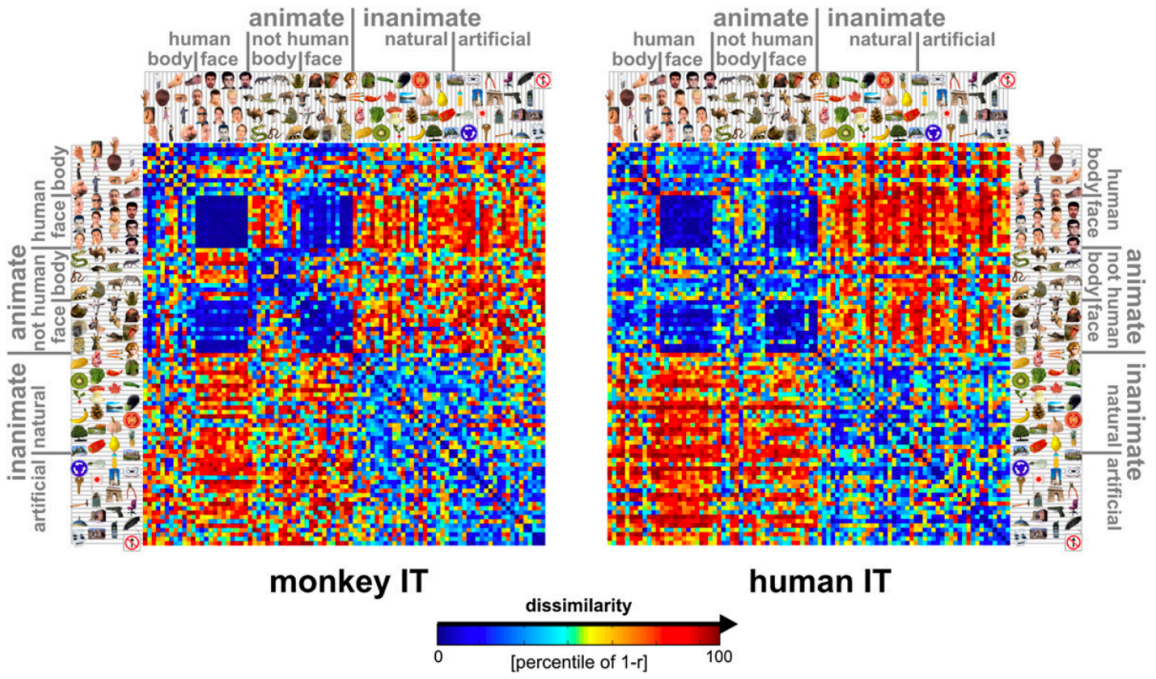


Figure 2.7 RDMs of neural activity in macaque and human IT cortex. Both show a similar categorical representation of visual stimuli. Reproduced from Kriegeskorte et al. (2008).

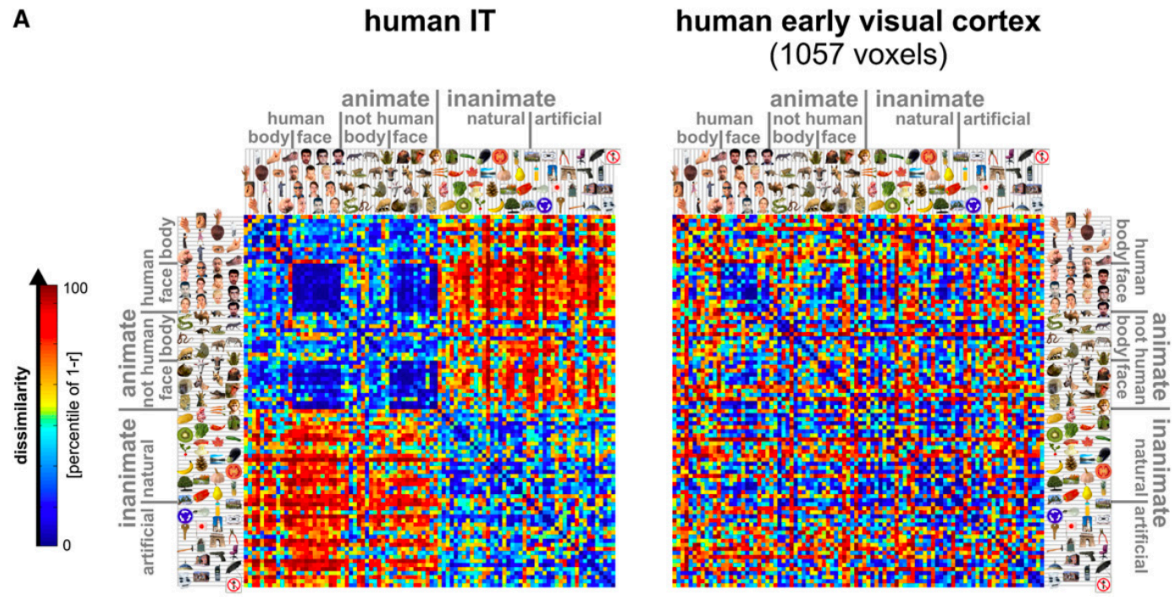


Figure 2.8 RDMs of neural activity in human IT and early visual cortex. Human IT shows a categorical representation of visual stimuli while early visual cortex does not. Reproduced from Kriegeskorte et al. (2008).

The RSA used in chapter 3 of this thesis as well as in most of the example studies presented in this chapter consider the representational geometry of spatial activity (ie. activity at each voxel or neuron is a dimension), but information can also be encoded temporally. To investigate the representational geometry of activity over time, RSA can be used to analyze imaging data of high temporal resolution, such a single-cell recordings or magneto- or electroencephalography (MEG/EEG) data. Luyckx et al. (2019) used RSA on EEG recordings to show that neural activity in response to numbers and probability in different value-based choice tasks shared a neural code of 'magnitude', showing that this may be a conceptual basis for learning new information used to guide new choices or actions. Cichy et al. (2014) used RSA to show that temporally early MEG responses carried similar representations to those in fMRI data from early visual areas, while later MEG responses had representations similar to those in fMRI data from IT cortex, showing how the visual response evolves in space and time.

Compared to sensory regions, fewer studies have so far used RSA to specifically investigate representations in prefrontal cortex (PFC). Holroyd et al. (2018) used an RSA searchlight to identify brain regions with BOLD activity similar to that of the layers of a recurrent neural network (RNN) in an action-sequence task and identified midcingulate cortex (MCC) as the region encoding action sequences most robustly. Performance of RNNs on action-sequence tasks had previously been linked to MCC activity univariately (Shahnazian & Holroyd, 2018), but RSA allowed for a direct comparison of activity levels between the model and the human brain. Kragel et al. (2018) showed representations of cognitive control, negative affect and pain,

generalizable across participants and tasks in fMRI data of anterior midcingulate cortex using RSA.

Hunt et al. (2018) established a triple functional dissociation in an information gathering task in single-cell recordings from macaque prefrontal regions (Figure 2.9). Orbitofrontal cortex (OFC) activity represented stimulus value, dorsolateral prefrontal cortex (dlPFC) activity represented the current locus of spatial attention, ie. the side of the screen on which a stimulus was presented, and anterior cingulate cortex (ACC) represented a belief confirmation signal. We replicated this study in human participants using fMRI and made direct comparisons between macaque and human RDMs (see chapter 3).

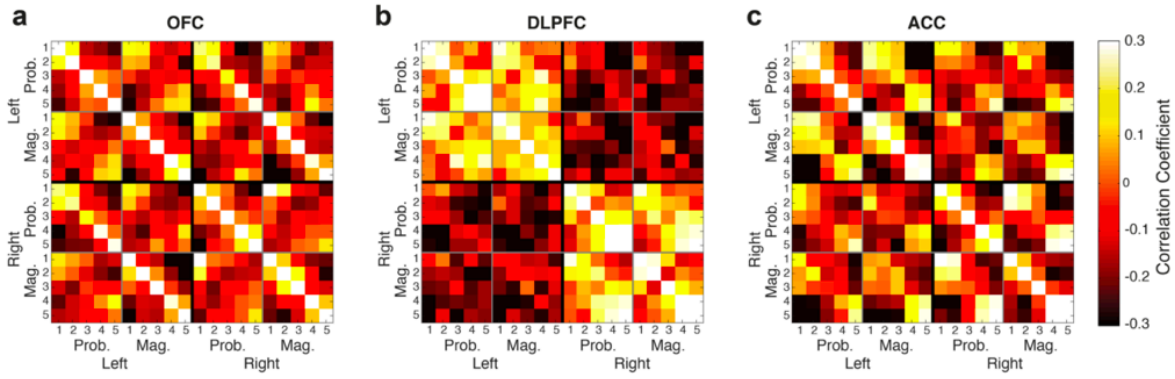


Figure 2.9 Triple functional dissociation in macaque prefrontal cortex in response to an information gathering task. Shown are RDMs from activity in three prefrontal regions in macaque in response to an information gathering task. The conditions being compared in the RDMs are ten stimuli presented either on the left or right side of the screen. Reproduced from Hunt et al. (2018).

2.2.6 How successful is RSA performance in PFC?

As you will see in chapter 3, we did not manage to find strong correlations between the macaque and human prefrontal RDMs, either using direct homologues of the three prefrontal regions studied in Hunt et al. (2018; Neubert et al., 2015; Sallet et

al., 2013) or using a searchlight analysis searching across the whole human brain for representations similar to those see in macaque PFC. However, we did find meaningful task representations in human PFC using RSA, meaning that in our study RSA does not seem to have contributed any additional understanding of representations in PFC over and above using a mass univariate analysis. In this section I will discuss some recent thoughts on to what extent RSA can be useful to study PFC function

There is much evidence that neurons in PFC show *mixed selectivity* (Fusi et al., 2016), meaning they respond to a wide range of tasks (Mansouri et al., 2006; Mante et al., 2013; Rigotti et al., 2010, 2013) and their activity can change flexibly over time (Duncan, 2013). It is thought that this property enables complex cognition, as mixed selectivity allows for a wide range of neural representations that can flexibly change to learn new tasks (Tang et al., 2019). High-dimensional representations allow for a linear readout by downstream areas that can easily separate combinations of different dimensions, while that is not possible for low-dimensional highly specialized neurons (Figure 2.10; Fusi et al., 2016). As RSA focuses on comparing the high-level geometry of neural codes rather than low-level selectivity of neurons, we might expect it to perform quite well in prefrontal regions, though an outstanding question is whether fMRI can actually pick up this mixed selectivity. On the other hand, Flesch et al. (2021) recently used RSA to show that PFC activity during a decision-making task was consistent with low-level representations of the task, raising the possibility that the specific experimental conditions in which neural activity is collected may have a large influence on how the experimental task is represented in the brain.

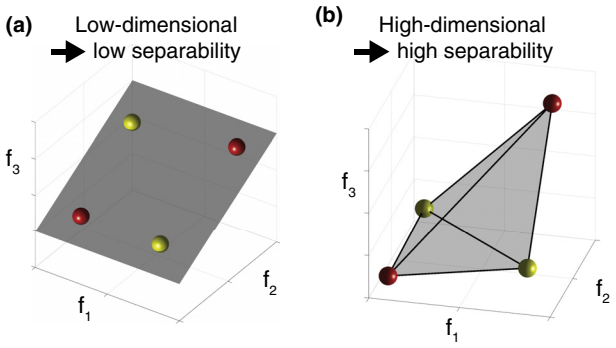


Figure 2.10 Linear readout from a set of highly specialized neurons is more difficult than from neurons with mixed selectivity. (A-B) The four spheres represent four different experiment conditions of two categories of interest (red and yellow) and the three axes the firing rates of three neurons. (A) If the neurons are highly specialized such that their firing rate responds linearly to a stimulus, the possible combination of firing rates is low-dimensional and a linear readout is not possible as no plane can be found that separates the two categories. (B) If just one of the three neurons has mixed selectivity instead (firing rate responds non-linearly to stimulus), a linear readout becomes possible. Reproduced from Fusi et al. (2016).

It is generally unclear whether multivariate analyses of fMRI data are really able to pick up fine-grained spatial information similar to that in single-cell recordings, or whether they mostly show the activity of large-scale neural population codes which a mass univariate analysis would also reflect (Op de Beeck, 2010). Hunt et al. (2018) observed intermixed positive and negative coding in their single-cell recordings of macaque prefrontal regions, which may not be reflected in voxel activity when collecting human fMRI data. Local averaging, which happens when we summarize activity into voxels, has been shown to reduce the detectability of differences between conditions in neural data (Dubois et al., 2015; Tremblay et al., 2015). Decoding accuracies in multivariate pattern analyses (MVPAs) of fMRI data are typically lower for prefrontal regions than other brain regions (Bhandari et al., 2018), while such decoding accuracies were generally high in other human brain regions as well as single-cell recordings of prefrontal regions (Bhandari et al., 2018; Rigotti et al., 2013).

This suggests the mixed selectivity of neurons in PFC is not well measured through fMRI. Furthermore, the activity measured in voxels using fMRI is not simply the average activity of neurons in that voxel, but the "complex spatiotemporal transform of the hemodynamic response" (Kriegeskorte et al., 2008), meaning patterns of neural activity are not identical to patterns in fMRI signals. Besides this, BOLD patterns have been shown to be less reliable in PFC than in other brain regions (Bhandari et al., 2018). Activity in PFC also shows a high degree of temporal autocorrelation, which may be a confound in studies trying to dissociate the representations of different sequential task conditions (Murray et al., 2014).

There is also evidence that neural representations of complex tasks are constantly changing, even long after decision errors have stopped decreasing, both in neural networks (Goldt & Seifert, 2017) and humans (Ruitenberg et al., 2018). As we usually average across trials from a long fMRI session (of an hour or longer), it is possible that RSA is not effectively capturing moment-to-moment representations of neural populations and the average representation may not be an accurate reflection of neural activity on a trial-by-trial basis.

Other brain regions where RSA has been successfully applied to fMRI data, such as the ventral visual stream, do not suffer from the same issues as PFC. Neural selectivity in sensory areas is thought to be less mixed than that in PFC: activity of single neurons in visual areas shows selectivity to fewer and lower-dimensional features (Hubel & Wiesel, 1968; Olshausen & Field, 1996) and this selectivity is strongly spatially clustered (Kamitani & Tong, 2005; Sengupta et al., 2017). This means that fMRI measurements of visual cortex are less likely to suffer from local

averaging. BOLD patterns have also been shown to be more reliable in visual areas than in PFC, making decoding these patterns easier (Bhandari et al., 2018)

An alternative approach to studying neural representations which does not suffer from some of the limitations outlined above is *repetition suppression* (Barron et al., 2016; Bhandari et al., 2019; Grill-Spector & Malach, 2001; Rigotti & Fusi, 2016). This is a technique that leverages the adaptational properties of neurons to test their involvement in a task. Neurons adapt to a repeatedly shown stimulus through decreasing firing rates and this is reflected in the BOLD response (Grill-Spector & Malach, 2001). This reduction in activity is not affected by the local distribution of neural activity, so local averaging into voxels should not reduce the ability to dissociate neural representations using this method. Representations can be studied by comparing repetition suppression responses to different task conditions, in other words, by comparing how much the BOLD response decreases in response to repeated stimulation by different conditions. This has been shown to produce more reliable distance measures using fMRI data than MVPA in dlPFC (Bhandari et al., 2019). A downside of this approach is that in order to compare repetition suppression across stimuli or conditions, stimulus transitions have to be perfectly counter-balanced, making it less well-suited for tasks in which events depend upon the participant's behavior (Barron et al., 2016). Besides this, whether the neuronal adaptation mechanism described above actually drives repetition suppression is debated (Barron et al., 2016). An alternative explanation is that the repetition suppression effect is in fact a reflection of predictive coding (Auksztulewicz & Friston, 2016; Summerfield et al., 2008). In this framework, a reduction in a prediction error signal, which indexes the difference between predictions about future events from downstream brain regions

and observed stimuli, would cause a reduction in BOLD signal within a voxel. This lack of clarity on what repetition suppression is measuring is a disadvantage when using it to infer what task features or behavior is being represented in brain activity.

2.3 Summary

I have shown the step-by-step process of testing hypotheses on fMRI data using GLMs with predictors defining task events. This approach has enabled me to investigate the neural signatures of sampling decisions in chapters 3 and 6. Furthermore, I have outlined how RSA can be used on fMRI data to compare neural representations across brain regions, species and modalities. I have also given some examples of the use of RSA in fMRI data in other studies, which has proved very useful in studying sensory brain regions, but shows mixed results in PFC, as also shown in chapter 3. Finally, I have provided some potential limitations of the use of RSA in fMRI data of the PFC.

Chapter 3: Medial frontal cortex activity predicts information sampling in economic choice

This chapter was published as a paper:

Kaanders, P., Nili, H., O'Reilly, JX, & Hunt, LT (2021). Medial frontal cortex activity predicts information sampling in economic choice*. Journal of Neuroscience, Vol. 41*(40), 8403-8413.

Abstract

Decision-making not only requires agents to decide what to choose, but also how much information to sample before committing to a choice. Previously established frameworks for economic choice argue for a deliberative process of evidence accumulation across time. These tacitly acknowledge a role of information sampling, in that decisions are only made once sufficient evidence is acquired, yet few experiments have explicitly placed information sampling under the participant's control. Here, we use functional MRI to investigate the neural basis of information sampling in economic choice, by allowing participants (n=30, sex not recorded) to actively sample information in a multi-step decision task. We show that medial frontal cortex (MFC) activity is predictive of further information sampling prior to choice. Choice difficulty (inverse value difference, keeping sensory difficulty constant) was also encoded in MFC, but this effect was explained away by the inclusion of information sampling as a co-regressor in the general linear model. A distributed network of regions across prefrontal cortex encoded key features of the sampled information at the time it was presented. We propose that MFC is an important controller of the extent to which

information is gathered before committing to an economic choice. This role may explain why MFC activity has been associated with evidence accumulation in previous studies, in which information sampling was an implicit rather than explicit feature of the decision.

Significance statement

The decisions we make are determined by the information we have sampled prior to committing to a choice. Accumulator frameworks of decision-making tacitly acknowledge the need to sample further information during the evidence accumulation process until a decision boundary is reached. However, relatively few studies explicitly place this decision to sample further information under the participant's control. In this fMRI study, we find that medial frontal cortex (MFC) activity is related to information sampling decisions in a multi-step economic choice task. This suggests an important role of evidence representations within MFC may be to guide adaptive sequential decisions to sample further information, before committing to a final decision.

3.1 Introduction

Decisions are determined by the information that the decision-maker samples before they commit to a choice. Since the 1970s, cognitive psychologists have developed ways to determine how participants decide to sample information as a decision unfolds (Newell & Simon, 1972; Payne, 1976). This has led to a rich understanding of the role of information sampling in economic choice behaviour (Bettman et al., 1998; Hunt et al., 2016; Kobayashi et al., 2019; Navarro et al., 2016; Stewart et al., 2016).

Perhaps the most pervasive behavioural effect is that more difficult decisions – that is, where options are particularly close in value – take longer, potentially allowing more time for the agent to acquire information about the alternatives (Busemeyer & Townsend, 1993; Hunt et al., 2012; Jamieson & Petrusic, 1977; Milosavljevic et al., 2010). Evidence accumulation frameworks such as the drift diffusion model (DDM) tacitly assume that more difficult decisions with longer reaction times are those in which more information is sampled (Ratcliff & McKoon, 2008; Usher & McClelland, 2001; Shadlen & Shohamy, 2016). Recent models have further argued that in economic choice, visual fixation provides a window into which option is currently being sampled, biasing evidence accumulation towards this attended option (Krajbich et al., 2010, Smith & Krajbich, 2019). However, the links between reaction time, fixation and information sampling remain implicit assumptions of such frameworks. This is because in these decision-making experiments, all information is typically presented immediately and simultaneously to the participant, rather than allowing them to decide whether more information is needed.

Indeed, most studies of the neural basis of economic choice have not placed information sampling under the participant's control (Gottlieb & Oudeyer, 2018). Although neural mechanisms supporting attention (Bisley & Goldberg, 2010; Thompson & Bichot, 2005) and information sampling (Blanchard et al., 2015; Horan et al., 2019; White et al., 2019) are increasingly understood, much less is known about how future economic choices are guided by it. On the other hand, several studies have used neural implementations of evidence accumulation models to identify brain networks that show similar trial-by-trial changes in aggregate activity as the model. A region of medial frontal cortex (MFC) that encompasses the dorsal anterior cingulate cortex (dACC) and the pre-supplementary motor area (preSMA) has been identified as correlating with accumulator activity (Gluth et al., 2012; Hare et al., 2011; Pisauro et al., 2017; Rodriguez et al., 2015; Venkatraman et al., 2009; Zhang et al., 2012). Typically, accumulator activity is greatest on trials where choice difficulty (i.e. the value of the unchosen option minus that of the chosen option) is highest, and choice difficulty itself has also previously been associated with activity in MFC BOLD fMRI signal (Grinband et al., 2006; Pochon et al., 2008; Wunderlich et al., 2009, Fitzgerald et al., 2009, Klein-Flugge et al., 2016, Pisauro et al. 2017). It is possible, however, that accumulator fluctuations may also reflect greater information sampling on these trials. Because none of these studies explicitly measured information sampling, it is unclear which of these two accounts – choice difficulty or information sampling – better explains BOLD responses in these areas (although importantly they are not mutually exclusive).

Here, we used BOLD fMRI to investigate the neural basis of information sampling in economic choice, by allowing participants to actively sample information in a multi-step decision task. Our experimental paradigm mirrors equivalent recent paradigms used in non-human primates (Hunt et al., 2018) and a large-scale human behavioural study (Hunt et al., 2016). As expected from our recent work, subjects sampled more information when decisions were more difficult. When participants were first able to make a choice, fMRI signal within MFC correlated with the difficulty of the decision. Crucially, this effect was significantly reduced by including a co-regressor indicating whether the subject sampled further information prior to choice. Activity in MFC (and bilateral intraparietal sulcus) was instead explained by the decision to sample further information. Furthermore, we hypothesized that activity in several frontal regions in response to the task could be dissociated using representational similarity analysis (RSA), as was seen in single-unit recordings in macaque using the same task (Hunt et al., 2018). While we found both sensory- and value-related activity using RSA, we did not find dissociable activity in response to different task features as in our previous study (Hunt et al., 2018).

3.2 Methods

30 healthy human participants (aged 18-50, sex not recorded) attended two study sessions: one behavioral training session (1hr) and one fMRI session (2hr 15min). One participant was excluded from the representational similarity analyses as their fMRI study session was ended prematurely, meaning we did not obtain enough data from this participant to appropriately balance across different conditions. In the behavioral session, participants learned the meanings of ten stimuli: five faces and

five houses. Five of these stimuli represented a reward probability (10%, 30%, 50%, 70%, or 90%) and the other five a reward magnitude (10, 30, 50, 70, or 90 points). We ensured that whether the cues were faces or houses was orthogonal to cue values by pseudo-randomizing the meaning of each cue for each participant. Specifically, there were always 3 face probability cues and 2 house probability cues (and similarly for magnitude), or vice versa, and these were interspersed such that faces or houses could not be associated with generally low or high value (Figure 3.1B). After learning the cues, participants were trained on the main task.

3.2.1 Experimental Design and Statistical Analyses

In the main task, participants were asked to choose between two pairs of cues (Figure 3.1A). Each pair consisted of a probability and a magnitude cue. The reward associated with that option was the number of points represented by the magnitude cue awarded probabilistically in accordance with the probability cue. 'Optimal' behavior in the task (maximizing long-run expected reward) would be to choose the side with the higher expected value (reward probability multiplied by reward magnitude). However, a trial started with the four cues being hidden (presented as grey squares), where the pair of cues on the left was one option and the pair on the right another. Participants were initially shown two different cues sequentially, the selection of which was pre-determined by the experimenter in order to balance cue presentation for representational similarity analyses. These could either be two cues from the same option ('option trials', 50% of trials) or two cues from different options representing the same attribute (probability or magnitude; 'attribute trials', 50% of trials). Note that the location of each attribute was counterbalanced to be the top or bottom cue, but that

this was consistent between the two options within a trial (e.g. if the top cue of the left option was the probability cue, the top cue of the right option was also the probability cue). The location and identity of the first cue were counterbalanced throughout each block. The reason the cues were shown sequentially is because this enabled us to study brain activity in response to each cue separately. At the second cue, we can also study choice difficulty and belief confirmation by looking at how the neural response depends on the value difference between the options and on how well the evidence provided by the second cue conforms to that provided by the first cue. Each cue presentation lasted 2s and was followed by a 0-4s jittered interval during which all the cues were hidden again.

After this, participants were given the opportunity to view the remaining two hidden cues by pressing the corresponding buttons. However, participants had to pay points to do this: 3 points to see a third cue, and another 6 points to see the fourth and final cue. As such, participants had to judge how useful the additional information provided by the last two hidden cues might be in ensuring they choose the option with highest expected value. In a pilot study, it was found that these costs typically led participants to sample further information on a subset of trials (Figure 3.1C/D; although we note that the propensity to sample further information still varied across individuals). Note that participants could not choose to see the first two cues again. When participants finished sampling the additional cues, or decided they did not want to sample further cues at all, they made a choice between the two options. On half of all trials, subjects received feedback on how many points they earned that trial; feedback was only revealed on 50% of trials, as this made sure stimulus-locked value-related signals and feedback-locked reward-related signals were decorrelated in the design

matrix. Inter-trial interval was jittered and between 3-7s. Participants were only invited to take part in the fMRI part of the experiment if they chose the highest valued option on 70% of trials in the last practice block of the main task.

In the fMRI session, participants first received further training on the task by doing two blocks of 40 trials each outside the scanner. Inside the scanner, participants did 10 practice trials and five blocks of the main task (a total of 200 trials). Participants received £1 for every 300 points earned in the task, in addition to a £20 show-up fee for coming to both study sessions. A counter in the right bottom corner of the screen kept track of the money earned up to that point throughout the task. As such, participants earned an average of £41 overall. This experiment was realised using Cogent 2000 developed by the Cogent 2000 team at the FIL and the ICN and Cogent Graphics developed by John Romaya at the Laboratory of Neurobiology at the Wellcome Department of Imaging Neuroscience.

3.2.2 fMRI Data Collection

Whole-brain fMRI measurements were made using a Siemens Prisma 3T scanner with a 2 x 2 x 2mm voxel size, repetition time (TR) = 1.235s, echo time (TE) = 20ms, flip angle = 65° with an axial orientation angled to AC-PC using a 64-channel head coil. The sequence used was MB3 PAT2. Participants performed five blocks of the main task inside the scanner with short breaks in between. For each new block a new run was initiated. Given that the task was not speeded, the runs were of variable length and so were the number of volumes collected per run. The average number of volumes collected per participant was 2587. T1-weighted structural images were

obtained using an MPRAGE sequence with 1 x 1 x 1mm voxel size, on a 174x192x192 grid, TE = 3.97ms, TR = 1.9s. A field map with dual echo-time images was also acquired (TE=7.38ms, whole-brain coverage, 2.4 x 2.4 x 2.4 mm voxel size).

3.2.3 fMRI Data Analysis

fMRI analysis was carried out using the FMRIB Software Library (FSL), and custom-written RSA scripts in MATLAB that built upon the Representational Similarity Analysis toolbox (Nili et al., 2014).

The pre-processing performed on the data was the following: motion correction using MCFLIRT (Jenkinson et al., 2002); non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 5mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=50.0s). Registration to high resolution structural and/or standard space images was carried out using FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001).

3.2.4 Behavioral data analysis

We studied the effects of cue value on choice and sampling decisions using single-subject logistic regressions followed by one-sample t-tests on the single-subject regression coefficients (Figure 3.1C/D). Specifically, we performed two logistic regressions on attribute trials where the independent variable was the value rank difference between the left and right presented cues and the dependent variable was whether or not the left option was chosen or whether a third cue was sampled, respectively. Similarly, on option trials two logistic regressions were executed where

the independent variable was the summed cue ranks of the two presented cues minus the mean value rank of an option and the dependent variable was whether or not the fully revealed option was chosen or whether a third cue was sampled, respectively.

3.2.5 Representational Similarity Analysis (RSA)

We first estimated a whole-brain general linear model (GLM) with regressors encoding onsets of the first two cues, cue category (face or house) and what side of the screen the cue appeared on. This GLM was executed on the spatially smoothed data. Then, we performed searchlight RSA, meaning RSA is performed for a 100 voxel sphere of voxels, centered around each voxel in the brain (Nili et al., 2014). These were then regressed against model representational similarity matrices (Figure 3.2) to find areas of the brain the activity of which showed similar representations. We then performed a statistical test at each voxel and significant clusters were identified using threshold-free cluster enhancement (TFCE; Smith & Nichols, 2009). The resulting p-values were family-wise error rate controlled at a threshold of $p<0.05$. This same method was used with representational similarity matrices obtained from Hunt et al. (2018) to investigate similarities between representations in monkey and human prefrontal cortex, although here we present unthresholded statistical maps as no clusters survived multiple comparisons correction in PFC.

3.2.6 Mass Univariate Analysis

Two GLMs were fit to the individual subject data, one including and one excluding regressors describing information sampling behavior. We also included regressors describing onset of the first two cues, cue values (separated into cues

where the associated option was eventually chosen or unchosen), side of cue presentation, trial type (option or attribute trial), whether feedback was presented, and if so, how much reward the participant received on the preceding trial. Lastly, our model included a regressor that reflected belief confirmation, encoded as an index defining how much the second cue should change the agent's belief about what the best option is (Figure 3.5). We used contrasts of parameter estimates to find regions that encoded the inverse value difference between the revealed cues (i.e. unchosen minus chosen value, or choice difficulty). All regressors were modelled as 2s boxcar functions, to match the duration for which the cue was presented. We report mass univariate results at the group-level using whole-brain family wise error (FWE) corrected statistical significance and a cluster significance threshold of $Z>2.3$ and $p<0.05$. Note that similar results were obtained using TFCE (Smith & Nichols, 2009).

To measure how much the main effect of choice difficulty was reduced by adding the information sampling regressor (Figure 3.4), without being biased by the initial GLM, we used a leave-one-out procedure. BOLD activity was extracted for each participant from the choice difficulty (in both GLMs) and information sampling contrasts using an ROI defined by a group model excluding that participant. The ROI was defined as any positively significant clusters ($Z>3.1$, $p<0.05$, corrected) for the contrast of inverse value difference found within a large mediofrontal mask. We then performed a one-sample t-test on the extracted parameter estimates for the inverse value difference contrast, both including or excluding information sampling behaviour as a co-regressor.

Unthresholded Z-statistic maps for all analyses are available at http://identifiers.org/neurovault.collection:9314.

3.2.7 Code Accessibility

Code for the task, behavioral data, and all behavioral analyses can be found at https://git.fmrib.ox.ac.uk/comp-cog-neuro/behavioral-data-and-analyses-from-kaanders-2021.git. The RSA results were produced using the RSA toolbox from the Cognition and Brain Sciences Unit, University of Cambridge, which can be found here: https://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes/.

3.3 Results

30 healthy human participants (aged 18-50) performed an information gathering and economic choice task inside the MRI scanner (Figure 3.1A). In the task, they were asked to choose between two pairs of cues, where the pair of cues on the left were one choice option and the pair on the right another. Each pair consisted of a probability and a magnitude cue. The reward associated with an option was the number of points represented by the magnitude cue awarded probabilistically in accordance with the probability cue. The cues were pictures of faces and houses and the magnitude or probability associated with them was learned by the participants in a behavioural session that took place before the main experiment (Figure 3.1B). 'Optimal' behaviour in the task (maximizing long-run expected reward) would be to choose the option with the higher expected value (=reward probability * magnitude).

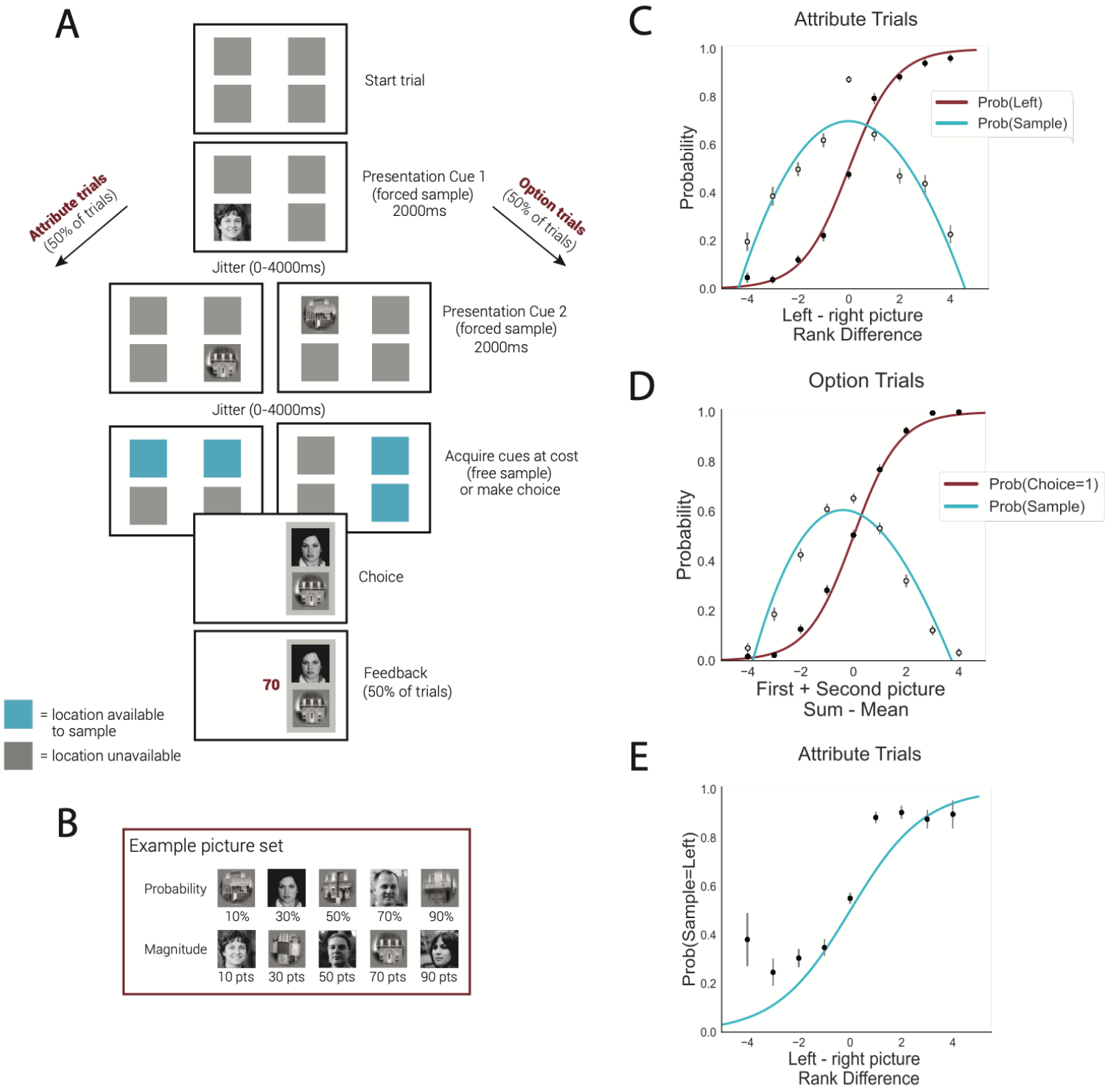


Figure 3.1. Experimental paradigm and participant behaviour. (A) Task structure. The task starts with all four cues being hidden. Two cues are shown sequentially. After this, participants can choose to sample more cues or make a choice between the two options. (B) Example picture set and their associated magnitude and probability amounts. (A,B) Note that the pictures of faces used in these panels do not belong to individuals and were generated. (C, D) Participants effectively use the cue values to make correct choices and to guide whether more cues need to be sampled. (C) In attribute trials, participants are more likely to choose the left option, the larger the value difference is between

the left and right observed cues. Participants are more likely to sample additional cues the closer this value difference is to 0. These psychometric curves are plotting the probability of choosing the left option and sampling a third cue as a function of the difference in value rank between the two revealed cues. (D) In option trials, participants are more likely to choose the option of which the cues have been revealed, the higher the sum of these presented cue values is compared to the mean value of an option. Participants are more likely to sample additional cues the closer this value sum is to the mean. (C-D) These psychometric curves are plotting the probability of choosing the option the cues of which were revealed and sampling a third cue as a function of the sum of the ranked values of the revealed cues subtracting the mean cue rank of an option. (E) Participants were more likely to sample the remaining hidden cue from the option with the highest revealed cue value in attribute trials. This psychometric curve plots the probability of sampling cues from the left option as a function of the difference in value rank between the two revealed cues and is only plotting trials on which participants sampled additional cues.

Crucially, each trial started with the four cues being hidden. Participants were initially shown two of the hidden cues sequentially in a pseudorandom order determined by the experimenter. These could either be two cues from the same option ('option trials', 50% of trials) or two cues from different options representing the same attribute (probability or magnitude; 'attribute trials', 50% of trials). After these first two cues had been viewed, participants could choose whether to sample the remaining hidden cues by pressing the corresponding buttons and paying a small number of points (-3 points to sample the third cue, -6 points to sample the fourth cue). Alternatively, participants could forgo the opportunity to sample further information and instead make a choice between the two options based on the information sampled thus far. Subjects then received feedback on how many points they earned that trial, although in order to decouple this haemodynamic response from that of the decision onset, feedback was only delivered on half of all trials (Guitart-Masip et al., 2011).

Both participants' eventual choices, and their propensity to sample more information prior to choice, depended upon the values associated with the first two presented cues. In attribute trials, their choices were a function of the difference in value between the left and right cues ($t(29)=10.79$, $p<0.001$; Figure 3.1C), whereas in

option trials their choices depended upon the value of the option relative to the mean expected value ($t(29)=15.88$, $p<0.001$; Figure 3.1D). Importantly, participants' decision to sample a third piece of information also depended upon these cue values, such that the more difficult the decision, the more information was sampled ($t(29)=-2.11$, $p=0.04$ on attribute trials; $t(29)=-7.20$, $p<0.001$ on option trials; Figure 3.1C-D). In attribute trials, the direction of this sample tended to be biased towards the option that currently had the highest value (Figure 3.1E), a bias also observed in our previous monkey (Hunt et al., 2018) and large-scale human behavioural studies (Hunt et al., 2016).

We first performed a whole-brain searchlight representational similarity analysis (RSA), using templates representing several features of the task at *the time of presentation of the first cue* (Figure 3.2). At each 100-voxel searchlight sphere we regressed the templates onto the RSA matrix of that location to find the regions where brain activity showed a similar representation. Several of these templates were the same as those used in our previous study using single unit recording in macaques (Hunt et al., 2018). First, as a positive control for the success of our searchlight RSA pipeline, we included a face/house template which represented which category a cue belonged to (Figure 3.2A). Replicating numerous RSA results in the prior literature (Dima et al., 2018; Kriegeskorte et al., 2008), this produced a strong correlation with visual and inferior temporal cortex (peak MNI coordinates at [31, -61, -9] and peak $Z=12.37$, $p<0.001$, Figure 3.2C). Consistent with previous studies, these results were robust and replicable even at the level of individual participants. The next template identified regions the multivariate activity of which distinguished between whether the currently presented cue was on the left or right hand side of the screen (Figure 3.2C).

As expected, this was strongly represented in primary visual cortex (peak MNI coordinates at [-6, -85, 4] and peak Z=11.28, $p<0.001$, Figure 3.2D). Surprisingly, however, we did not find activation in frontal eye fields or dorsolateral prefrontal cortex, despite this pattern clearly being represented in dlPFC in our previous single-unit data (Hunt et al., 2018).

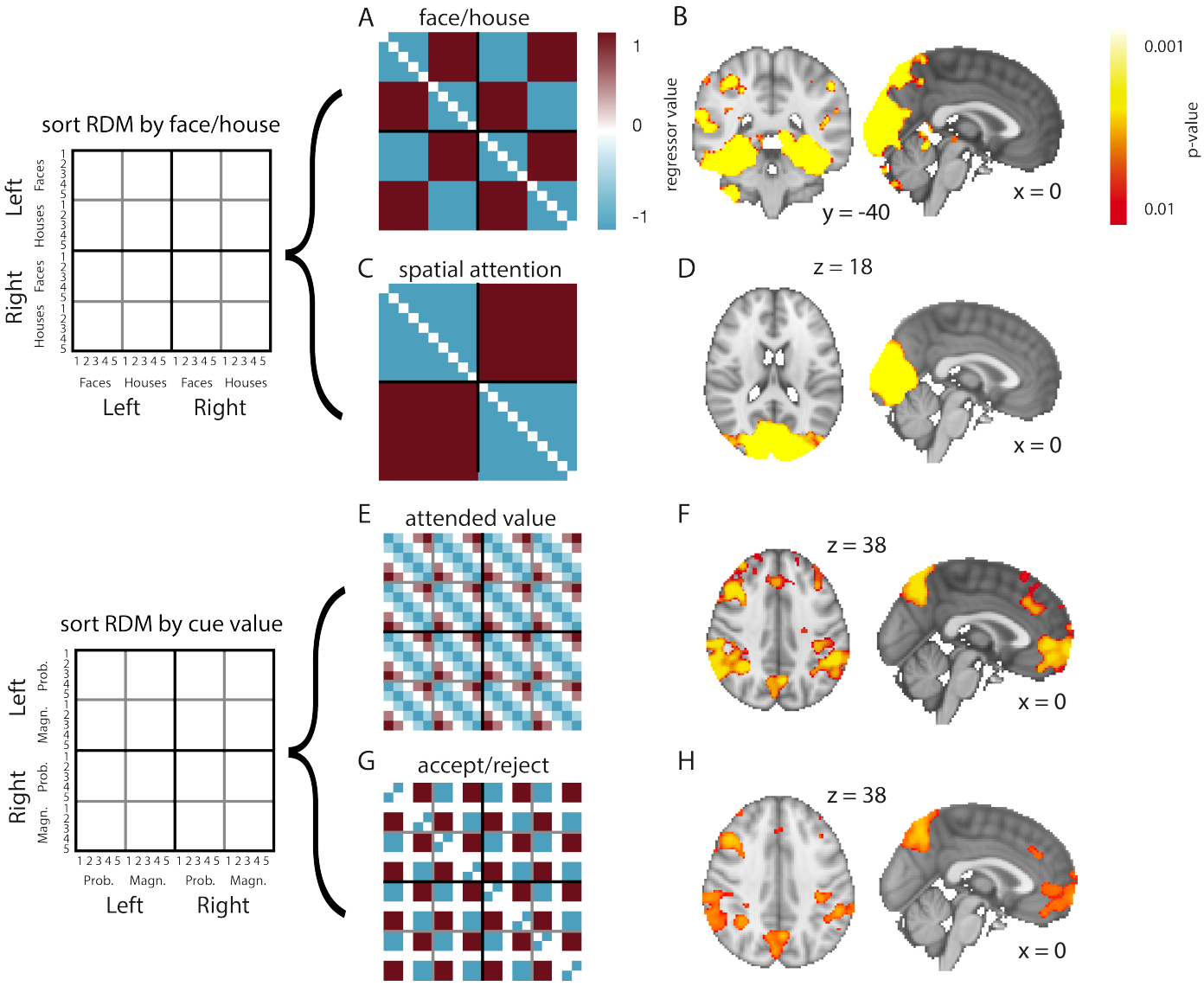


Figure 3.2 Whole-brain searchlight analysis reveals representations of several features of the task. (A,C,E,G) RSA templates for spatial attention, face/house category, attended value, and belief confirmation. (B) Cue category (face or house) is represented in visual and inferior temporal cortex. (D) Spatial attention is represented in visual cortex. (E-H) We find activations for attended value and belief confirmation in an elaborate network including vmPFC, ACC, and posterior parietal cortex. All parametric maps are depicting p-values derived from using threshold-free cluster enhancement (TFCE) on GLM z-score maps.

To investigate the neural correlates of the value-related features of the task, we included 'attended value' and 'accept/reject' templates, mirroring those in our previous study (Hunt et al., 2018). The 'attended value' template simply represents the value of the first cue, ranked from 1-5 and then normalized, such that low-value cues are predicted to be similar to other low-value cues, and high-value cues are predicted be similar to other high-value cues (Figure 3.2E). The template assumes that magnitude and probability cues were treated the same in terms of how valuable they were, because to calculate the expected value of an option participants had to collapse across the two cues, and because magnitude and probability cues were found to be similarly encoded in macaque OFC single-cell recordings (Hunt et al., 2018). The 'accept/reject' template also represents the value of the first cue but binarises this into 'above average' or 'below average' in value (or 'neutral', for the mid-value cue; Figure 2G). In our single-unit data, this correlated with activity in dACC (Hunt et al., 2018). In our fMRI data, we found that both the 'attended value' and 'accept/reject' templates correlated with the activity in an extensive network within frontal and parietal cortex, including ventromedial PFC (vmPFC; peak at MNI coordinates [2, 50, -1] and peak $Z=4.36$, $p<0.001$ for 'attended value' template; peak at MNI coordinates [2, 51, 1] and peak $Z=3.90$, $p=0.006$ for 'accept/reject' template), dACC (peak at MNI coordinates [0, 37, 29] and peak $Z=4.80$, $p=0.003$ for 'attended value' template; peak at MNI coordinates [0, 35, 31] and peak $Z=5.91$, $p=0.006$ for 'accept/reject' template) and posterior parietal cortex (peak at MNI coordinates [52,-49,51] and peak $Z=5.91$, $p<0.001$ for 'attended value' template; peak at MNI coordinates [52, -49, 51] and peak $Z=5.14$, $p=0.002$ for 'accept/reject' template; Figure 3.2E-H). As expected, this

network corresponds well with areas previously identified in meta-analyses of subjective value encoding using fMRI (Bartra et al., 2013; Clithero & Rangel, 2014).

We then investigated the neural correlates of the decision to sample further information, which first occurred at the time of presentation of the second cue, and whether this might explain activations that would otherwise be described as encoding choice difficulty. To test this, we used a mass univariate analysis, as the task comprised too many conditions to perform RSA at the time of presentation of the second cue. In an initial analysis, we included only the values of the different cues as regressors, and found a subregion of MFC that was negatively related to the difference between the values of the chosen and unchosen options, which we will refer to as 'inverse value difference' (Figure 3.3A, peak at MNI coordinates [-4, 15, 50] and peak Z=6.27). In other words, activity in this region was greater for more difficult decisions, replicating a large number of previous findings in the economic choice literature (Grinband et al., 2006; Pochon et al., 2008). Low-level sensory features were counterbalanced across value levels (Figure 3.1B) and across participants, so this effect was only due to the value difference between the choice options. In option trials, the choice difficulty was defined as the value difference between the revealed option and the mean expected value of a choice option. In attribute trials, choice difficulty was defined as the value difference between the revealed cues. While a main effect of choice difficulty was seen looking across all trials, when splitting by trial type the effect was present both on option trials (peak at MNI coordinates [-2, 14, 52] and peak Z=5.37) and attribute trials (peak at MNI coordinates [-4, 15, 50] and peak Z=5.72).

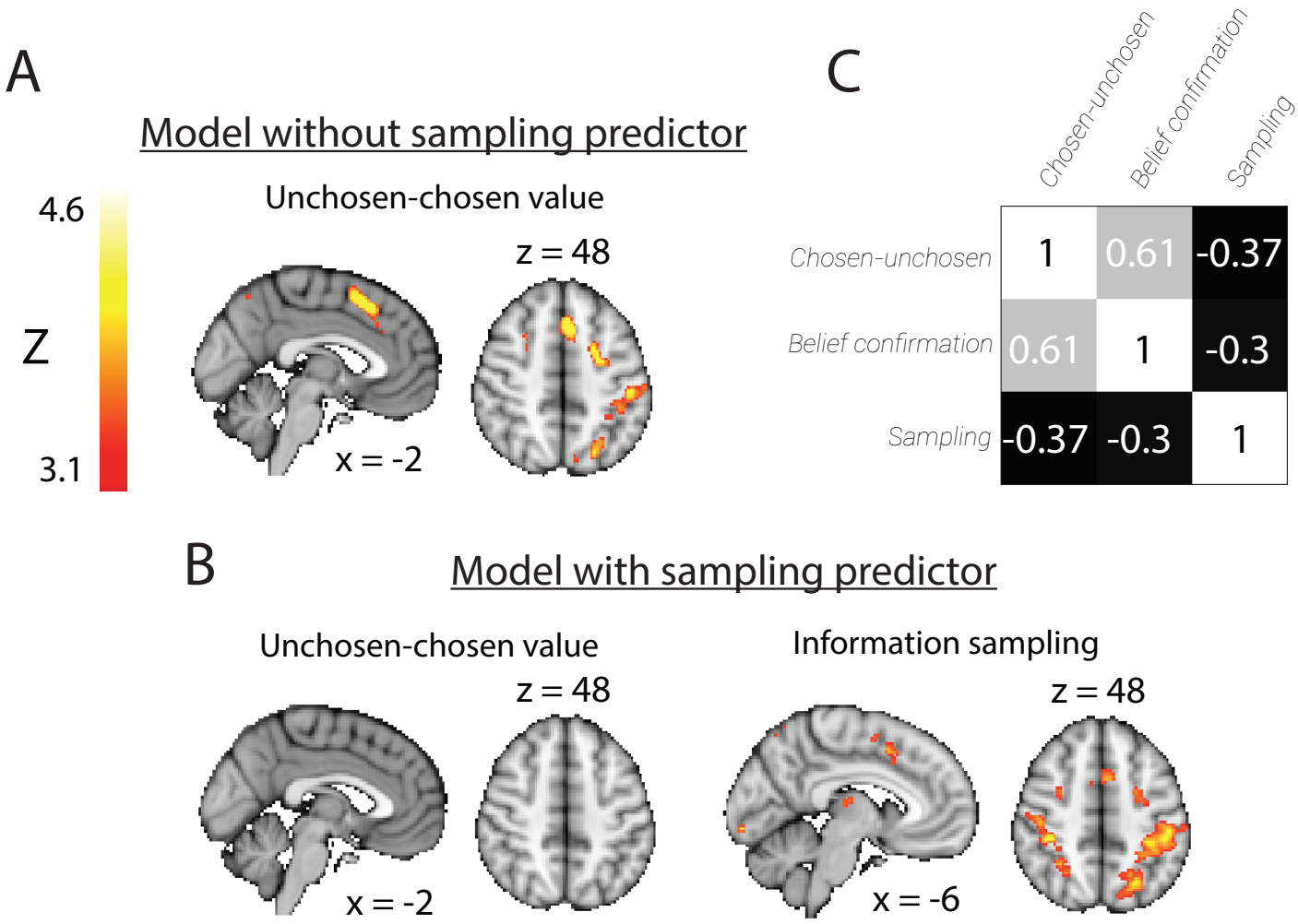


Figure 3.3. There is a main effect of inverse value difference (choice difficulty) at the time of presentation of cue 2 in MFC, which disappears when subsequent information sampling is included as a regressor. (A) Linear positive effect of inverse value difference in MFC in 30 subjects. (B) There is no effect of inverse value difference in MFC when information sampling is included as a regressor, while there is a linear positive effect of information sampling in MFC in 24 subjects. All parametric maps are cluster-corrected and thresholded at z>3.1. (C) Correlation matrix of value difference, belief confirmation and sampling predictors in mass univariate analysis in 24 subjects.

Crucially, this model did not yet account for whether participants chose to sample further information before committing to one of the options. In a further regression model, we included a binary information sampling regressor describing whether or not participants chose to sample more information on a given trial. As expected from the relationship between difficulty and information sampling (Figure 3.1C/D), this regressor had some correlation with the regressor encoding value difference (average correlation of r=-0.3; Figure 3.3C; we note that 6 participants were excluded from this analysis, as they sampled additional information before choosing less than once per block). We found that the main effect of choice difficulty in MFC disappeared when the information sampling predictor was included, and this was not a consequence of the exclusion of the 6 participants who rarely sampled further information. Instead, a main effect of information sampling was seen in a small section of MFC, which was more active on trials where subjects decided to sample further information (Figure 3.3B, Figure 3.4, peak activation at MNI coordinates [-6, 10, 46] and peak Z=4.04). We also tested whether any other regions encoded choice difficulty after inclusion of the regressor that reflects information sampling; no regions were found to survive correction for multiple comparisons, although some subthreshold activations remained (unthresholded statistical maps, both including and excluding information sampling co-regressors, are available at http://identifiers.org/neurovault.collection:9314). It is important to note that this effect cannot be attributed to variations in response times, and thus, trial durations, as participants were only allowed to respond after viewing the second cue for 2s and a 0-4s jitter. All regressors in this analysis were modelled as 2s boxcar functions to reflect the duration of cue presentations. Consistent with recent reports that highlight

the role of parietal cortex in reducing uncertainty about future rewards (Horan et al., 2019), we also found bilateral activation of intraparietal sulcus for this regressor (peak activation at MNI coordinates [-36, -32, 42] and peak Z=5.78, Figure 3.3B).

To further explore whether choice difficulty encoding was significantly reduced by the inclusion of the information sampling regressor, we performed a region-of-interest analysis. We focused our results on the MFC region that encoded inverse value difference (i.e. choice difficulty) in the whole brain analysis (Figure 3.3A), but in order to avoid circularity (Kriegeskorte et al., 2009), we used a leave-one-out procedure in which each participant's ROI was defined by taking significant MFC voxels from a group model including all participants except them (Boorman et al., 2013). This revealed a significant reduction in the encoding of inverse value difference as a consequence of including the information sampling regressor (Figure 3.4B; $t(23)=3.28$, $p=0.003$), such that choice difficulty was no longer significantly encoded in this region ($t(23)=1.17$, $p=0.26$). Instead, this region of interest significantly encoded the decision to sample further information (Figure 3.4C; $t(23)=4.17$, $p<0.001$).

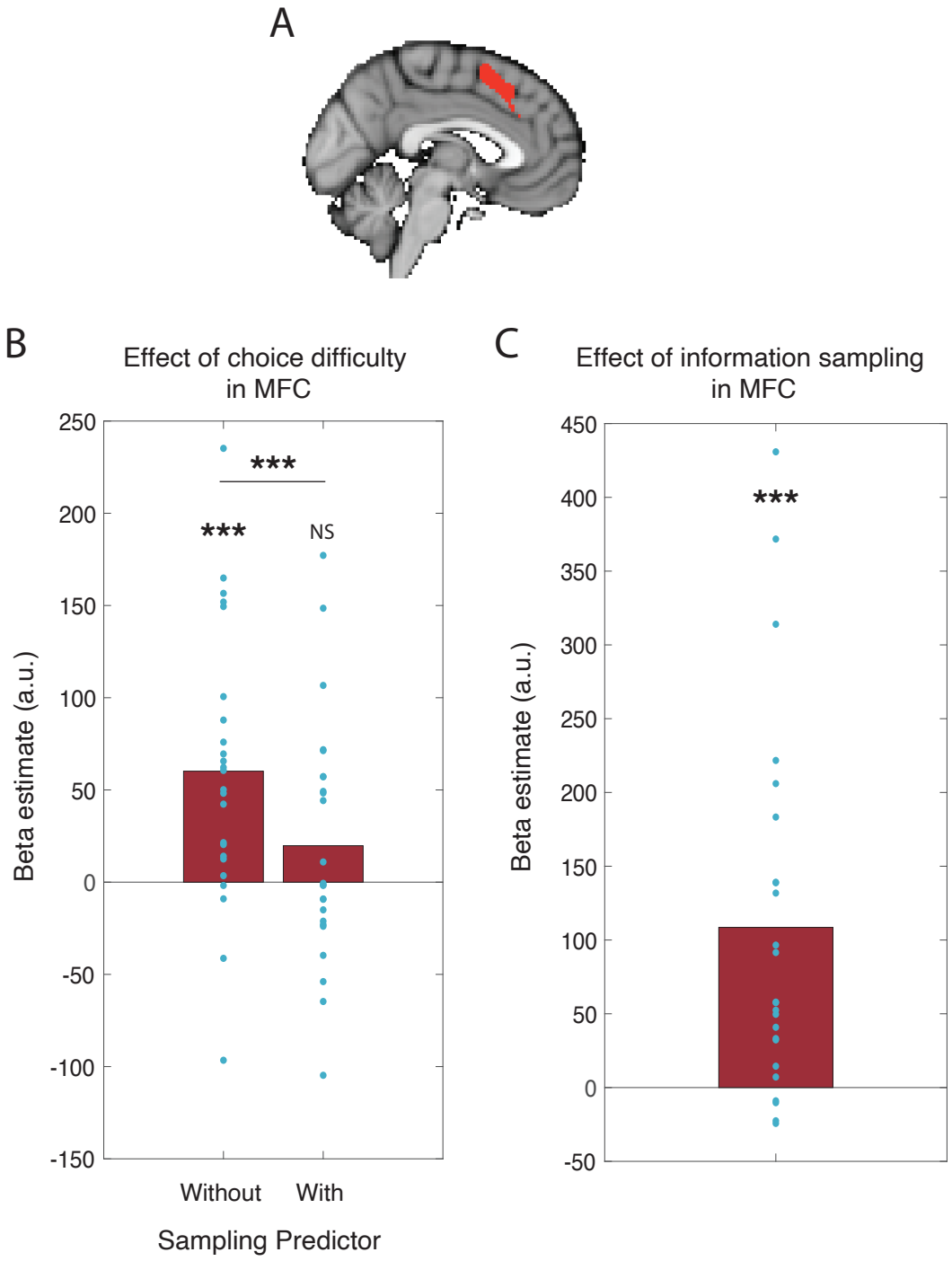


Figure 3.4. When information sampling is included as a co-regressor here is a significant decrease in the main effect of inverse value difference (in an MFC region of interest equivalent to the region defined in Figure 3A, defined by a leave-one-out procedure). (A) MFC region of interest derived from mass univariate analysis including all 30 subjects, focusing on the region that maximally encodes choice difficulty. (B) Effect of inverse value difference in MFC in GLMs with and without the information sampling co-regressor in 24 subjects. (C) Effect of information sampling in the same brain region. A stronger BOLD signal was found in this region on trials where participants would subsequently decide

to sample more information before committing to a choice compared to trials where they made a choice straight away.

We also investigated encoding of belief confirmation, which was defined as the extent to which the value of the second cue confirmed the belief formed after presentation of the first cue. In our single unit study, we found that belief confirmation was reliably encoded in dACC (Hunt et al., 2018), and similarly in the present study we found fMRI activity in a more anterior portion of MFC encoded belief confirmation on 'option' trials (Figure 3.5A; peak Z=4.81 at peak at MNI coordinates [8, 34, 8]). We did not find significant activation related to belief confirmation on 'attribute' trials. The effect of belief confirmation in 'option' trials did not survive whole-brain correction in the reduced sample of 24 participants who regularly sampled more information before choosing. We again used a leave-one-out procedure to determine how this effect changed as a result of reducing the sample size and including the information sampling regressor (Figure 3.5C). While there is only a significant main effect of belief confirmation in MFC in the large sample of 30 participants ($t(28)=2.54$, one-tailed $p=0.008$), this effect is not significantly reduced by excluding participants and including the sampling predictor from the model ($t(51)=0.06$, $p=0.952$; $t(23)=-0.002$, $p=0.998$).

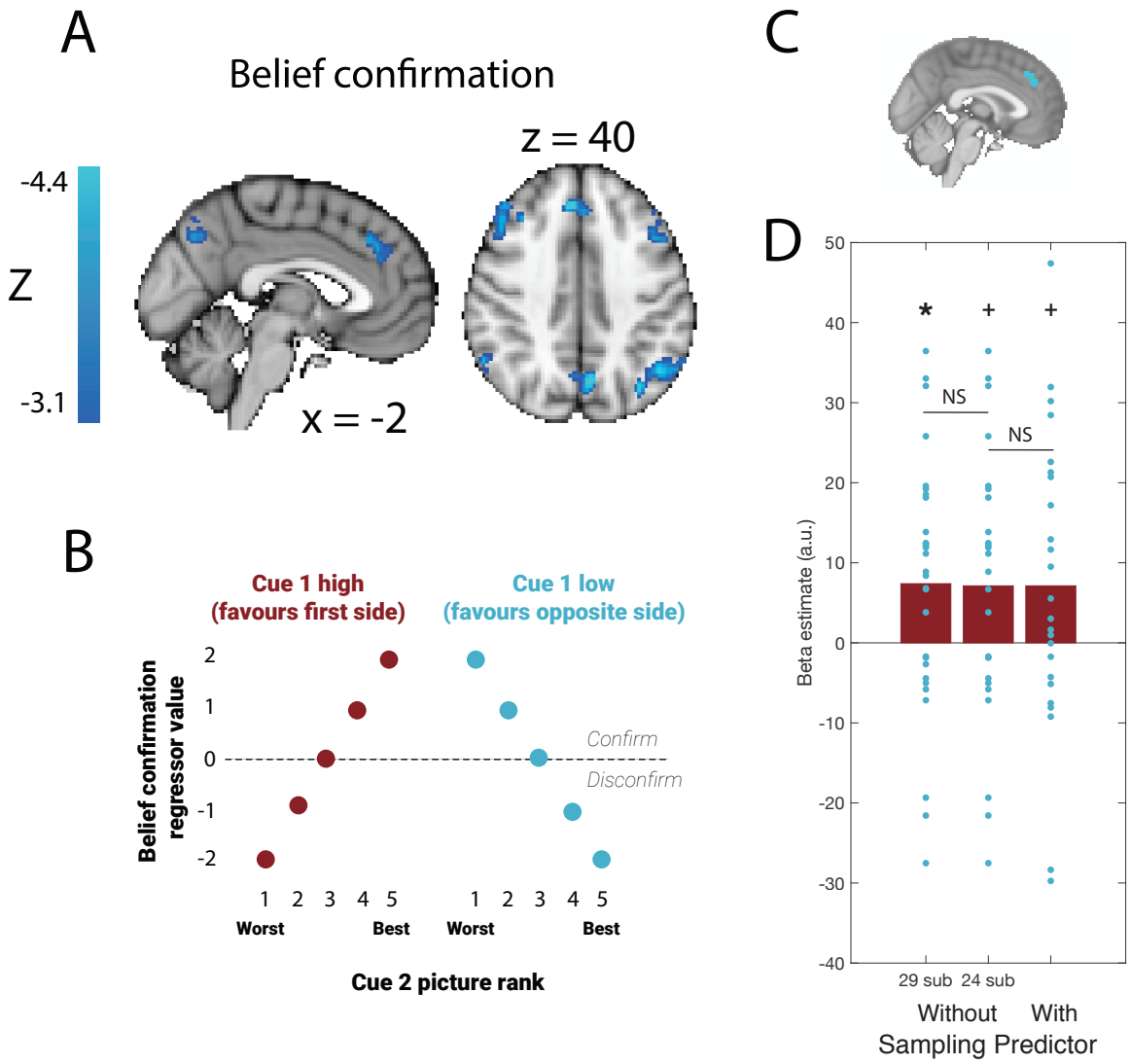


Figure 3.5. A main effect of belief confirmation was also found in MFC on option trials, but this effect was no longer statistically significant when using the reduced sample of 24 subjects. (A) Linear negative effect of belief confirmation in MFC on option trials in 30 subjects. Parametric maps are cluster-corrected and thresholded at z<-3.1. (B) Encoding of belief confirmation as a function of the ranks of the two presented cues. (C) MFC region of interest derived from mass univariate analysis including all 30 subjects. (D) Effect of belief confirmation in option trials in MFC in GLMs with 29 subjects or 24 subjects, and with and without the sampling predictor.

Finally, we hypothesized RSA could also be a way of identifying functional homologues of macaque prefrontal subregions in the human brain. To study task representations across species, we used the RSA matrices found in single-cell recordings in three prefrontal regions in the macaque (Hunt et al., 2018) for another whole-brain searchlight RSA (Figure 3.6). Prefrontal BOLD activity from the human brain was not found to be significantly correlated with any of these macaque RSA matrices. However, some subthreshold activations of interest were found: the macaque ACC matrix was correlated with human anterior insula activity (peak at MNI coordinates [-28, 24, -3] and peak Z=3.97) and the macaque OFC matrix was correlated with human vmPFC activity (peak at MNI coordinates [0, 46, 9] and peak Z=3.34). A strong correlation was found between the macaque dlPFC matrix and human visual cortex, which was expected as macaque dlPFC strongly represented spatial attention on this task (peak at MNI coordinates [-5, -88, 3] and peak Z=21.16).

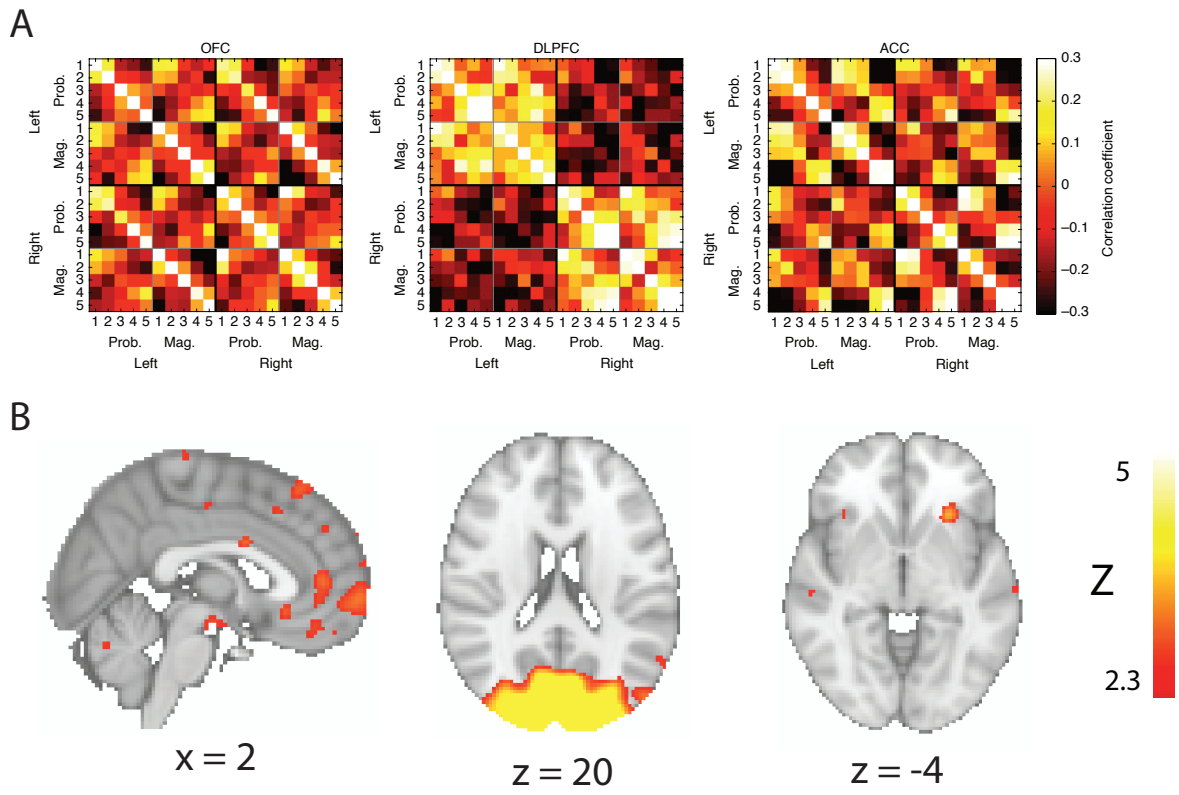


Figure 3.6. To further investigate RSA as a method to link functional data across species (Kriegeskorte, 2008), we performed a whole-brain searchlight on the human fMRI data using (A) RSA matrices derived directly from the single unit recordings measured in three relevant prefrontal regions in macaques, instead of using model RSA matrices as in Fig. 2 (Hunt et al., 2018). (B) No significant whole-brain cluster-corrected correlations were found within human prefrontal cortex, though subthreshold activations were found in human insula correlated with the macaque ACC matrix, and human ventromedial prefrontal cortex (vmPFC) correlated with the macaque orbitofrontal cortex (OFC) matrix. As expected from the result in Figure 2C/D, human visual cortex was found to be related to the macaque dorsolateral prefrontal cortex (dlPFC) matrix, which strongly reflected the participant's current locus of spatial attention. All parametric maps are uncorrected and thresholded at z>2.3.

3.4 Discussion

Here we have shown that in an information sampling and economic choice task, MFC activity encodes whether a subject will sample further information before committing to a decision. We moreover found that this effect could explain away the encoding of choice difficulty (inverse value difference) in MFC, a signal that has been commonly observed in many previous studies of economic choice (Grinband et al., 2006; Pochon et al., 2008). Our findings should not be interpreted as suggesting that

difficulty is not encoded in MFC, but instead that difficulty is a fundamental determinant of whether further information is sampled prior to choice. Here, we have defined choice difficulty purely in terms of the complexity of the value comparison between two choice options. As the low-level sensory features of the cues are similar across the different cues, value comparison is the only source of choice difficulty in this task. However, difficulty is not necessarily identical to choice uncertainty. In fact, trial-to-trial changes in decision confidence can in certain circumstances be dissociated from choice difficulty (De Martino et al., 2013); it is possible that such metacognitive fluctuations may be linked to or even predictive of information sampling.

Our findings form a logical extension of previous work on the role of MFC as an evidence accumulator (Gluth et al., 2012; Hare et al., 2011; Pisauro et al., 2017; Rodriguez et al., 2015; Venkatraman et al., 2009; Zhang et al., 2012), as evidence accumulation involves an implicit decision to sample further information when a decision bound has not yet been reached (Shadlen & Shohamy, 2016). An advantage of the current information gathering paradigm is that it allowed us to study the relationship between MFC signals and subsequent behavior; most paradigms in the field are much simpler and do not allow for this type of analysis.

Our findings are consistent with previous studies that propose a role for MFC in guiding exploratory behavior (Badre et al., 2012; Boorman et al., 2009; Daw et al., 2006; Domenech et al., 2020; Hayden et al., 2009; Kolling et al., 2012). For example, one recent study by Blanchard and Gershman (2018) contrasted trials in which participants decided to place a bet, versus trials where they decided to obtain further information about the value of that bet. Activity in MFC and bilateral insula was significantly greater on observe trials than on bet trials (we also found subthreshold

activity in insula predictive of information sampling, see maps on http://identifiers.org/neurovault.collection:9314). Similarly, a single unit recording study by White et al. (2019) identified dACC as the leading node in a network of areas that encoded gaze shifts to resolve uncertainty about upcoming rewards, and further studies have shown that activity in this area predicts 'checking' behaviours when monkeys are close to receiving a reward (Stoll et al., 2016). Collectively, this work is consistent with a role for this region in driving information sampling to resolve uncertainty (Monosov, 2020), which in turn gives rise to shifts in visual fixation and spatial attention (Krauzlis et al., 2014; Gottlieb et al., 2020).

More broadly, different subregions of MFC have been associated with diverse parameters in different paradigms (Kolling et al., 2016), whether it is sampling value (Stoll et al., 2016), conflict (Botvinick et al., 1999), choice difficulty (Grinband et al., 2006; Pochon et al., 2008), or belief (dis)confirmation (Boorman et al., 2013; Hunt et al., 2018; Quilodran et al., 2008; Wessel et al., 2012). It is notable that all of these task parameters may affect the degree to which additional information sampling is required before committing to a choice. On the other hand, it has also been suggested that dACC is particularly active when participants switch away from a default behavior (Boorman et al., 2013; Kolling et al., 2016). It is possible that we find mediofrontal activity predictive of information sampling because in our study not sampling additional information could be considered a default behavior, given that participants commit to a choice straight away more often than they decide to sample more cues (Figure 3.1C/D).

Our behavioural results also indicate that humans are more likely to sample information from an option that they believe is good (Figure 3.1E), consistent with our previous behavioural findings (Hunt et al., 2016). We hypothesize that MFC may therefore drive information sampling through active hypothesis testing (Markant & Gureckis, 2012). This is consistent with the presence of a 'belief confirmation' signal in ACC, which we find here (Figure 3.5) and also in our previous single unit study (Hunt et al., 2018). This effect was not reduced by including an information sampling predictor. It is also consistent with recent behavioural findings in perceptual decision-making (Talluri et al., 2018), and the recent suggestion that gaze affects the integration of goal-relevant evidence rather than value in economic choice (Sepulveda et al., 2020). Such an account of economic choice has recently been proposed as particularly ecologically valid in decisions containing multiple options (Hayden, 2018). More detailed testing of this hypothesis will require paradigms in which multiple options are available for choice, and where information sampling and choice are explicitly dissociated from one another.

We were able to identify brain networks encoding both sensory- and value-related aspects of the task using RSA, some of which encompassed areas in PFC (Figure 3.2) whose value responses have previously been shown using multivariate pattern analysis (Kahnt et al., 2010). A similar region was found to encode continuous and categorial ('below' or 'above' average) attended value. We tested for the presence of these two regressors despite their correlation because in the macaque single-cell data from Hunt et al. (2018), continuous value was found to be mostly encoded in OFC, while categorical value was encoded in ACC. In the same analysis, we also observed value-related activity in bilateral parietal cortex, which is consistent with a similar RSA

regressor used to describe parietal cortex EEG responses to numerical magnitude in an economic choice task (Luyckx et al., 2019).

The clear correlation between templates describing the visual properties of cues in the task and visual/inferotemporal cortical activity replicates earlier work showing that RSA distinguishes objects of different categories in these areas in human fMRI data (Dima et al., 2018; Kriegeskorte et al., 2008; O'Toole et al., 2005). It has previously been suggested that such an approach can be used to study representations across species, found using different types of data (Kriegeskorte et al., 2008). We hypothesized that our task might allow us to perform a similar cross-species analysis in order to identify functional homologues of the PFC subregions studied with single unit recordings in Hunt et al. (2018). We did not clearly identify the same triple dissociation that was seen in our previous study (Figure 3.6).

It is possible that we were not sensitive to detect a dissociation in PFC here similar to that in Hunt et al. (2018) due to the lower signal-to-noise in human fMRI data compared to that in single-cell data. It still remains unclear whether RSA really probes fine-grained spatial information that is similar to that identified by single neuron data, or whether it is primarily driven by the same macroscopic signals that can also be isolated in mass univariate analyses (Op de Beeck, 2010). It is possible that the intermixed positive and negative coding that supported the successful RSA in Hunt et al. (2018) is not observable at the voxel level in human fMRI. This may also explain why decoding accuracies in multivariate pattern analyses are invariably far lower in prefrontal cortex than in occipital and temporal cortices (Bhandari et al., 2018). It is also clear that there are differences in the nature of the spiking and BOLD signal which mean that certain features which are readily decodable in single-unit data cannot be

found using equivalent analysis of fMRI data (Dubois et al., 2015). In addition to this, despite the similarity between the present task and our previous monkey study, there are also some obvious differences. For example, monkeys were trained on the task for several months beforehand (Hunt et al., 2018) whereas humans were only trained shortly before entering the scanner. Although both performed the task well with similar patterns of behaviour, they may have approached the task using different strategies. It is also possible that the differences could result from anatomical differences between human and monkey PFC, yet all of the regions previously studied in the macaque have direct homologues in the human (Sallet et al., 2013; Neubert et al., 2015); we therefore believe this to be an unlikely explanation for the difference.

In summary, we have shown that MFC activity in economic choice predicts subsequent information sampling. We replicated previous studies showing that MFC activity in different subregions is related to choice difficulty, but critically this effect was explained away by the inclusion of information sampling as a co-regressor. This suggests that the role of MFC may extend beyond that of evidence accumulation, and implies an important role for MFC in guiding adaptive information sampling during economic choice.

Chapter 4: Developing a new information sampling task

Abstract

We developed a new task in which we wanted to implement three key features: the opportunity to actively sample choice options, complete separation of information sampling and reward-maximizing choice, and manipulation of information volatility and goal proximity. To do so, we developed multiple iterations of the task and piloted these on small samples online. Each task version was altered as a result of conclusions drawn from the previous pilot. In particular, we did not find an effect of information volatility and goal proximity on sampling in the first few task versions. These limitations were addressed in the final task version, which we then used in an fMRI study (chapter 6). We also developed a version of the task optimized for investigating heterogeneous sampling behavior, which was used in a large-scale online psychiatric study (chapter 5).

4.1 Ecological information search

Consider how you behave when visiting a new city: your attention is likely to be drawn to many things that are not relevant to any decisions at that moment. It might be drawn to a restaurant across the street, even if you are not currently hungry. So, information search might be driven by a model of the world that knows what features in the world *tend* to be goal-relevant and which might become relevant in the *future*, such as when I become hungry later in the day.

Most economic choice tasks place all decision variables under the control of the experimenter rather than the participant, as this makes decisions considerably easier to study and model. Unfortunately, these frameworks do not capture some of the features of real-life information search (see section 1.4).

4.1.1 Active sampling

There is evidence that even when the information we sample is directly relevant to an imminent choice, such as which restaurant to pick when hungry, information sampled passively affects choice differently from information sampled actively (Chambon et al., 2020; Kaanders et al., 2021). This might be due to attention's hypothesized role as an internally generated factor that is allocated according to optimal sampling policies (Callaway et al., 2021; Jang et al., 2021). In other words, information might be processed differently if it is not presented 'in sync' with the agent's internal sampling process. As such, traditional economic choice tasks may not reflect how information is naturally sampled and used in decision-making. To investigate this properly, we set out to develop a choice task where participants are *in control* of the information presented to them.

4.1.2 Separating information sampling from choice

Reinforcement learning has typically considered exploration compared to exploitation, which is a form of active information search, but one where information gain is traded off against reward maximization (Blanchard & Gershman, 2018). This means that it is unclear whether a choice alternative was selected because the agent wanted to know the value of this option or because the agent expected the reward

associated with this option to be high. By *explicitly separating* decisions to sample information from decision to select a choice alternative, we may be better able to characterize the drivers and neural signatures of information sampling.

4.1.3 Manipulating volatility and goal proximity

If it is the case that most of our information sampling is driven by an internal model of which sources of information tend to be goal-relevant, sampling should be affected by how frequently information sources change and goal proximity. Firstly, agents should prefer sampling information that is likely to be valuable for longer. Using our example of visiting a new city, your attention should be drawn less to a temporary hotdog stand than to a restaurant, as the stand may no longer be there by the time you get hungry. To investigate this, we can *manipulate the speed at which recently sampled information becomes irrelevant*. In this chapter we will refer to this as *volatility.* Volatility of choice options has been shown to affect decision-making within the context of bandit tasks, where the rates of change of drifting values of bandits are manipulated by the experimenter (Behrens et al., 2007). As discussed in section 1.2, information sampling can be thought of as many sequential sampling decisions. As such, we might expect the volatility of choice options to affect sampling decisions in a similar way in an active information sampling task.

By the same token, the closer an agent is to a goal, the more important it should become to the agent to collect information relevant to that goal. Wilson et al. (2014) showed that participants were more exploratory when the information gained could be used in upcoming choices compared to a condition in which only a single choice was

made. In our task, we will manipulate *goal proximity* and thereby vary the usefulness of gathered information for upcoming choices.

4.1.4 Overview of task development

In summary, I set out to develop a novel decision-making task where participants could actively sample information, sampling decisions and reward maximization are decoupled, and the volatility of information sources and goal proximity could be manipulated. As we developed this task from scratch, we went through many different iterations of the task and investigated whether each version could adequately address the research questions set out above by piloting them on online participants on Prolific. In Table 4.1, I have listed the different task versions we developed, their main features and the main conclusions from data collection which led us to adapt the task and create the next version.

Version	**Task features changed from previous version**	**(pilot) Sample size**	**Conclusions**
1	• Two alternating goals with three choice options each. • Choice options increased in value across trials at three different rates. • Each trial consists of one sampling round, in which one choice option could be sampled, and one choice round.	9	Participants **rarely sampled information relevant for the upcoming goal** close to a goal switch.
2	• Single goal with three choice options (eliminating the 'upcoming goal' element of version 1).	28	Participants' sampling strategies varied considerably and this task version was **too complex to distinguish different drivers** of sampling.

3	• A block consisted of 5-10 trials and at the start of each block all choice options were reset to a random new value, which were either all hidden or revealed to the participant.	30	**No effect was observed of the rate of change of the choice options on sampling**.
4	• Mean rate of change is the same for all three choice options, but the range of possible increases was varied.	17	This version was **difficult to learn** for online participants.
5	• On each sampling round one choice option is picked at random which the participant can sample at a cost if they so choose. Sampling rounds are structured into blocks of 3-8 sampling rounds, following by a choice round.	30	**Goal proximity was not a strong driver of sampling**, presumably due to the fact that the choice options' values changed in a predictable way.
6	• Each trial consists of one sampling round, in which as many choice options could be sampled as the participant wanted, and one choice round.	450	See **chapter 5** for results of large-scale online study.
7	• Choice options maintained the same value until they 'popped', at which point their value was reset to a new, randomly sampled value. 'Popping' frequency varied across the three options. • On each sampling round one choice option is picked at random which the participant can sample at a cost if they so choose. Sampling rounds are structured into blocks of 3-8 sampling rounds, following by a choice round.	15	See **chapter 6** for results of fMRI study using this task version.

Table 1. Overview of all task versions, main changes made due to conclusions drawn from previous pilots, how many pilot participants were used to test it, and conclusions drawn from the pilot data.

4.2 Task development

4.2.1 Task description

In all versions of the task, the main goal of was to find the balloon containing the most points among a set of multiple balloons by sampling them before having to make a decision. This explicitly separates out the 'sampling' of the balloon information from the use of that information for decision-making.

Participants were presented with multiple balloons of values between 1 and 100 points represented by value bars next to each of the balloons. To vary uncertainty throughout the task, each of the balloons increased in value on each trial until their value reached more than 100 points, at which point the balloon would 'pop' and start increasing in value again from 1 point. Crucially, all the balloons' values were hidden from view unless participants chose to sample them during the sampling trials. Uncertainty was then highest when a balloon was predicted to be of high value, as there was a high probability that the balloon would pop on the next trial. Therefore, we expected to see that sampling would be more likely when participants estimated a balloon's value to be high.

To study the effect of volatility of the balloons on sampling, the balloons varied in how quickly they increased in value. Fast balloons increased at a rate of 21-35 points/trial, medium balloons at 11-25 points/trial and slow balloons at 1-15 points/trial (see 'rate of change' in Figure 4.1). The increase on each trial was sampled from a uniform distribution with the corresponding range and values increased after each choice trial. Therefore, the value of the 'fast' balloon would be uncertain on more trials compared to the other balloons: due to the fast increase of its value, it would more frequently be in an uncertain range, as it could either be high or already have popped.

By contrast, the value of the 'slow' balloon would not be uncertain as often, due to its slow increase. We predicted that as a result, participants should want to sample the 'fast' balloon more frequently than the 'slow' balloon, as the information gathered from any individual sample would be relevant for less long.

Furthermore, each of the balloons had a probability of 5% of being reset to a randomly sampled value between 1-100 points on any given trial. This ensured that participants would need to sample each balloon at least occasionally.

To independently investigate the drivers of sampling and choice, sampling decisions to gather information were completely separated from choice decisions to maximize reward by alternating sampling trials and choice trials (see 'current round' in Figure 4.1). On sampling trials, participants were allowed to sample one of the balloons. This meant that that balloon's current value was revealed in the value bar next to the balloon (see 'value' in Figure 4.1). This value bar would stay on the screen until the balloon was sampled again in a future trial, at which point the value bar would update to the balloon's new current value. Dots shown next to each of the balloons represented the number of trials that had passed since that balloon was last sampled, to prevent working memory from being a confound on this task (see 'dots indicating trial since last sample' in Figure 4.1).

On choice trials, participants had to choose one of the balloons. Participating received the number of points corresponding to their chosen balloon on that trial, but did not receive feedback on the number of points they had received after choosing. In other words, participants' goal was to identify the balloon with the highest value using the sampling rounds.

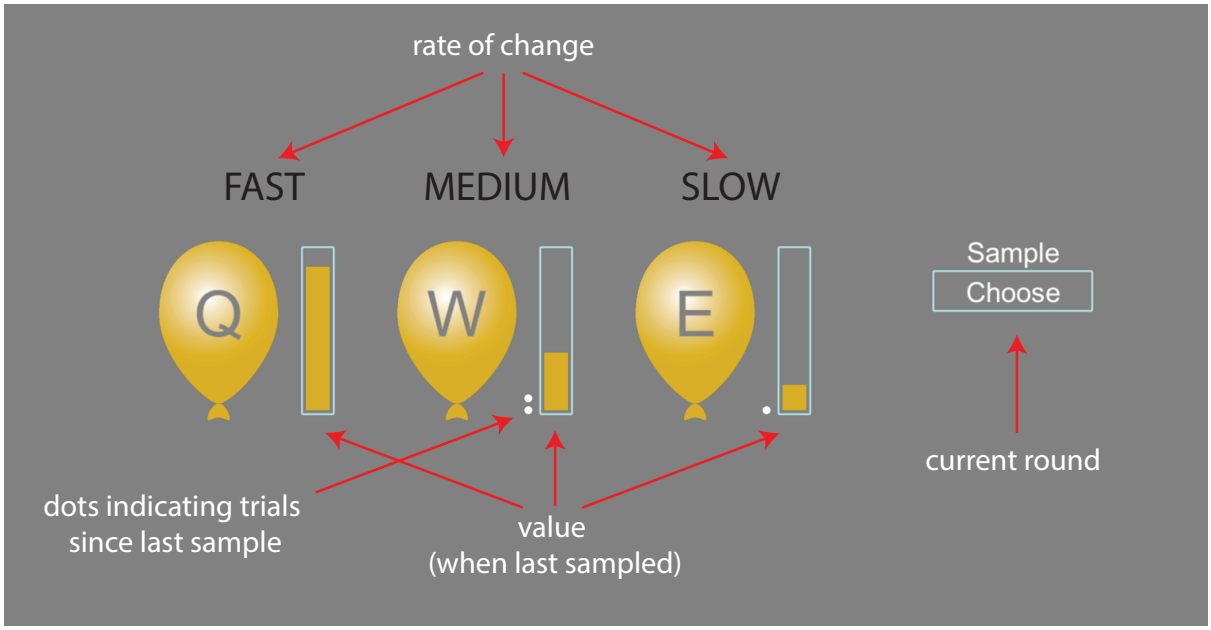


Figure 4.1 Screenshot of task version 1 pointing out task features.

4.2.2 Task version 1: Dual goal manipulation

In the initial version of the task, participants actually had two goals: finding the best gold balloon and finding the best silver balloon, one of which was always more proximate than the other. This manipulation was intended to investigate whether participants would sample information to inform future goals rather than the current one (as we predicted people do in the real world, see section section 4.1). The balloons were of two types, 'gold' and 'silver', and each choice trial was either a 'gold' or 'silver' trial, meaning that participants could only choose from the gold balloons or the silver balloons, respectively. Participants could see of which category the next 6 choice trials would be. The choice trial balloon category would stay the same for 4-8 trials at a time after which there would be a 'goal switch'. On sampling trials, participants could sample any of the 6 balloons, regardless of what category the upcoming choice trial was. By giving participants this option, we expected to see sampling of the *not* currently relevant balloon type increase on trials leading up to a goal switch.

9 online participants were recruited to perform 150 blocks of this task (each block consisting of 1 sampling trial and 1 choice trial). One participant was excluded, because it became clear from their answers to the debriefing questions they had not understood the task properly. It became apparent that the dual goal manipulation in this task was not working, as participants rarely sampled a balloon of the 'other' category than of the choice trial category coming up directly after. We had expected to see a linear effect of goal proximity such that participants would be more likely to sample the 'other' category the closer to a goal switch they were, but as the overall rate of sampling of the 'other' category was so low, this could not be measured (Figure 4.2).

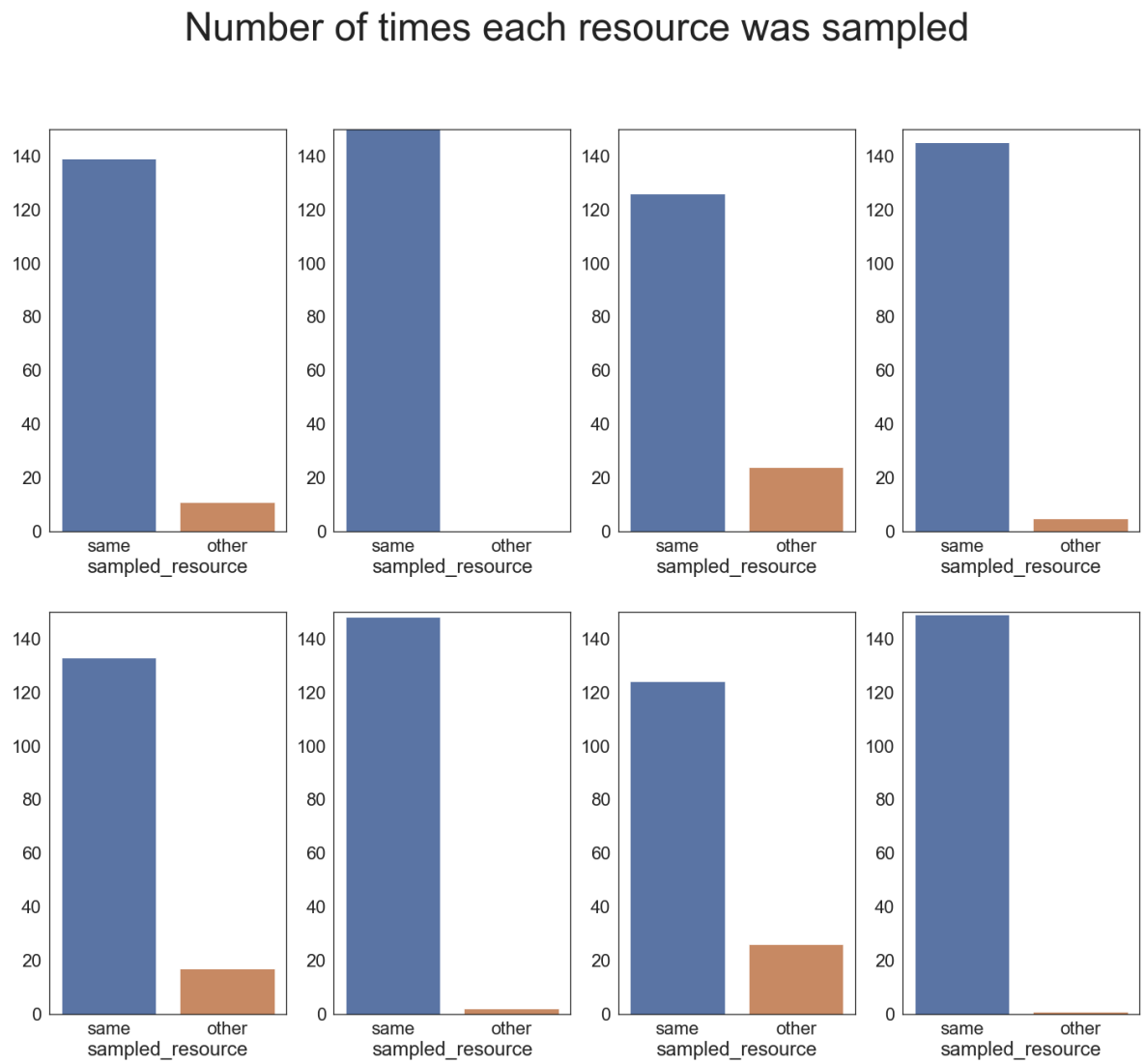


Figure 4.2. Pilot participants rarely sampled one of the balloons of the category opposite to that of the upcoming choice trial. Plotted are the number of trials a balloon of the upcoming goal category was sampled next to the number of trials a balloon of the opposite goal category was sampled for each participant.

In this version we attempted to study prospective sampling behavior, by including two goals and giving participants the option to sample information for future goals. However, participants rarely did this, perhaps because they had enough sampling options that they did not feel prospective sampling was necessary. We therefore decided to drop this dual goal manipulation and focus on studying the effects

of some of the other key task features in a simplified task with a single goal. Manipulating the number of goals in a decision-making task might remain an interesting avenue for future research, but this would require some further task development to figure out in which scenarios prospective sampling can and cannot be observed.

4.2.3 Task version 2: Information obtained from sampling effectively used to guide choices

Due to the results of the previous pilot, we decided to drop the dual goal manipulation and moved to a version of the task where there were only 3 balloons of the same type. Other than that the task stayed the same. At this stage we also developed an optimal model of behavior on this task, which is explained in detail in the supplementary materials (section 4.6.1).

We recruited 28 online participants to perform 250 trials of this version of the task. 3 participants were excluded because they misunderstood the task as evidenced by their answers on a debriefing questionnaire at the end of the experiment. Firstly, participants performed above chance level at this task (t_{24} = 12.85, $p<0.001$; Figure 4.3). They were more likely to choose a balloon when its value was high the last time they sampled it (t_{24} = 9.18, $p<0.001$, Figure 4.4), meaning participants used the information they gathered in sampling rounds to make better choices.

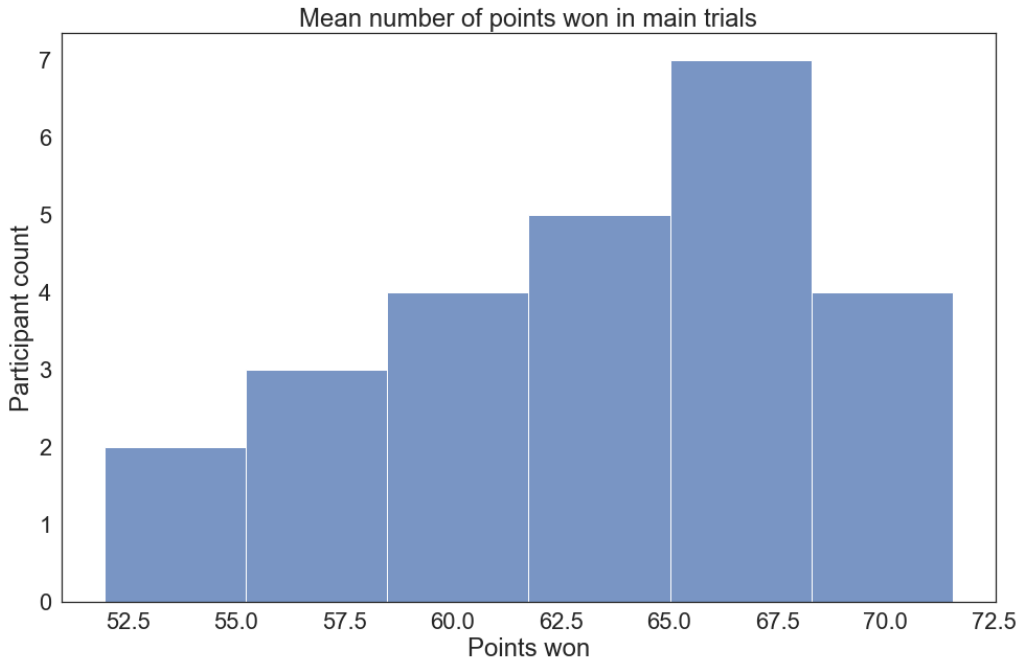


Figure 4.3 Participants performed fairly well on task version 2. Plotted are the average number of points won by each participant. The mean number of points an agent would win if choosing balloons at random would be 50 points.

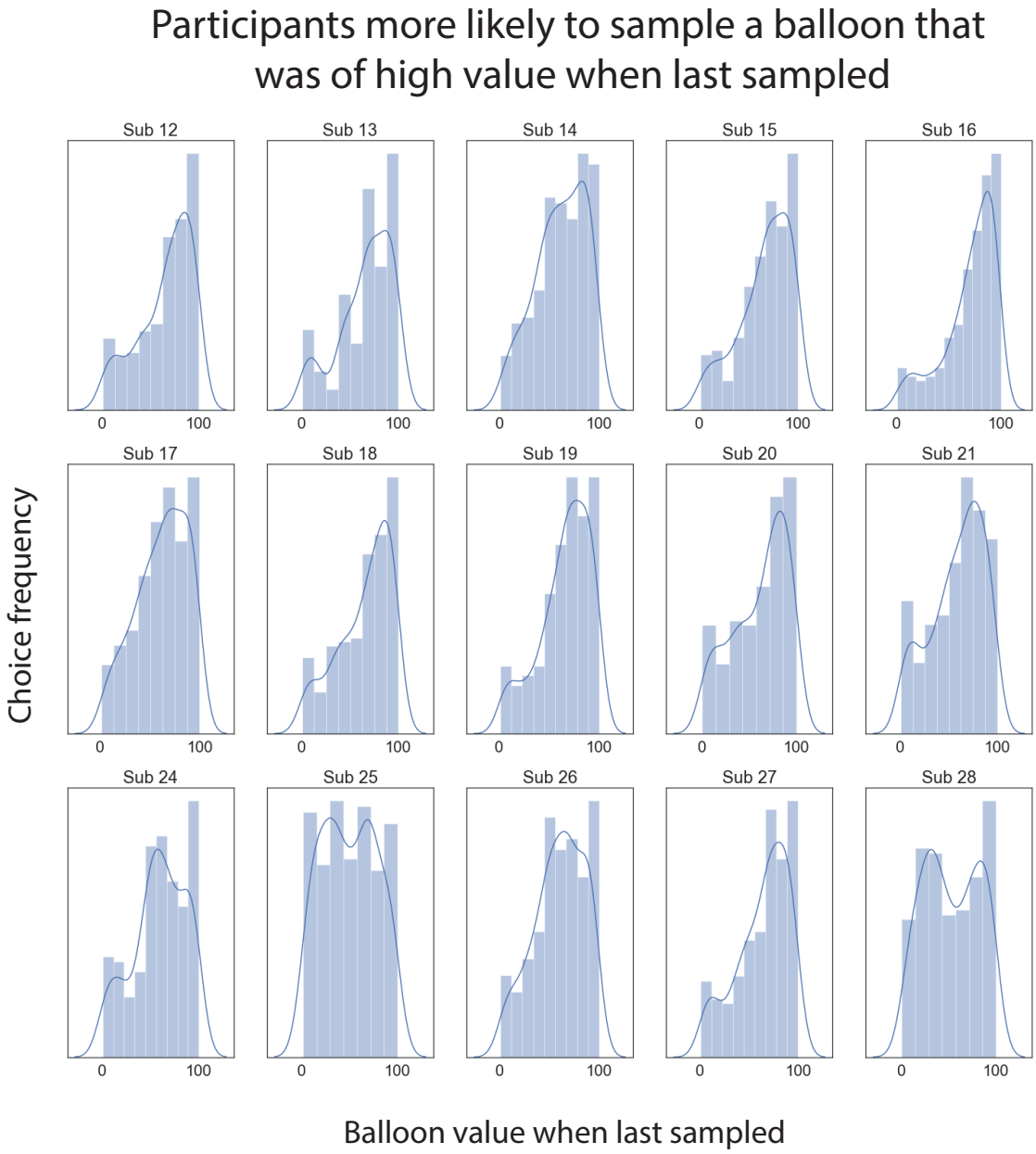


Figure 4.4 Most participants were more likely to choose a balloon on a choice trial if it had a high value the last time they sampled it in task version 2. Plotted is the frequency of sampling by value of the balloon at last time of sampling for 15 example participants.

As participants were only allowed to sample one balloon on each sampling round in this version of the task, we predicted that participants would be more likely to sample a balloon, the longer ago they last sampled it. Furthermore, we expected this to be modulated by the rate of change of the balloon, such that participants would want

to sample a faster changing balloon more frequently, as the information from this balloon would remain relevant for less long (supplemental materials section 4.6.1 describes how this prediction can also be derived from an optimal agent). Participants on average seemed to want to sample the three balloons approximately equally, as evidenced by the fact that balloons were most frequently sampled when they were last sampled two trials ago, suggesting that each balloon was sampled on average every 3 trials (Figure 4.5). However, this was not true for all participants. Some had a strong preference for sampling the balloon that was just sampled on the previous trial, suggesting they were attempting to track a single balloon. We referred to such samples as 'greedy' samples and show that participants showed considerable variation in how often they sampled balloons twice (or more frequently) in a row (Figure 4.5). In other words, participants used different sampling strategies on this task and their behavior is unlikely to be explained by a single (optimal) model. This large variation in sampling strategies seen on this task made it difficult to determine what features of the task were driving sampling decisions in general across our sample. The latter is essential for this task to be suitable for use with fMRI, as large between-subject variations are notoriously difficult to study in fMRI data (Marek et al., 2022).

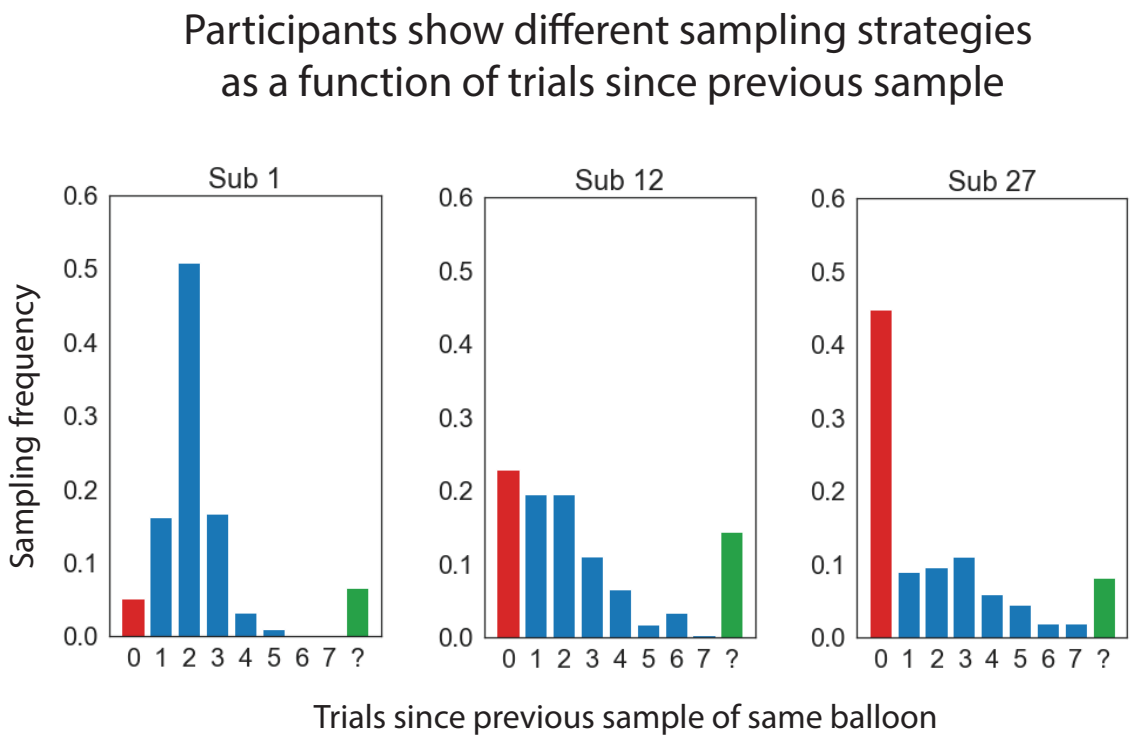


Figure 4.5 Sampling frequency of balloons by how many trials ago they were last sampled by each participant in task version 2. The red bar refers to trials where the same balloon was sampled twice in a row, which we refer to as 'greedy' samples. The green bar refers to balloons that were sampled when unknown (the first sample after a balloon's value was reset), which we refer to as 'exploratory' samples. The different patterns of sampling as a function of how many trials passed since the previous sample suggests the use of different sampling strategies. For example, participant 27 engaged in mostly 'greedy' sampling, while participant 1 mostly sampled the balloons equally.

To see whether participants used the different rates of change of the balloons in their sampling decisions, we looked at sampling frequency of each the three balloons as a function of how long ago it was last sampled, and what value it had then. Sampling of a balloon was most common when combining information gained from prior samples with the knowledge about the rates of change of the different balloons would suggest a balloon was currently potentially of high value, but not when it was likely the balloon had already popped (Figure 4.6). This means participants showed good understanding of the different rates of change of the three balloons and used them to decide whether to sample a balloon.

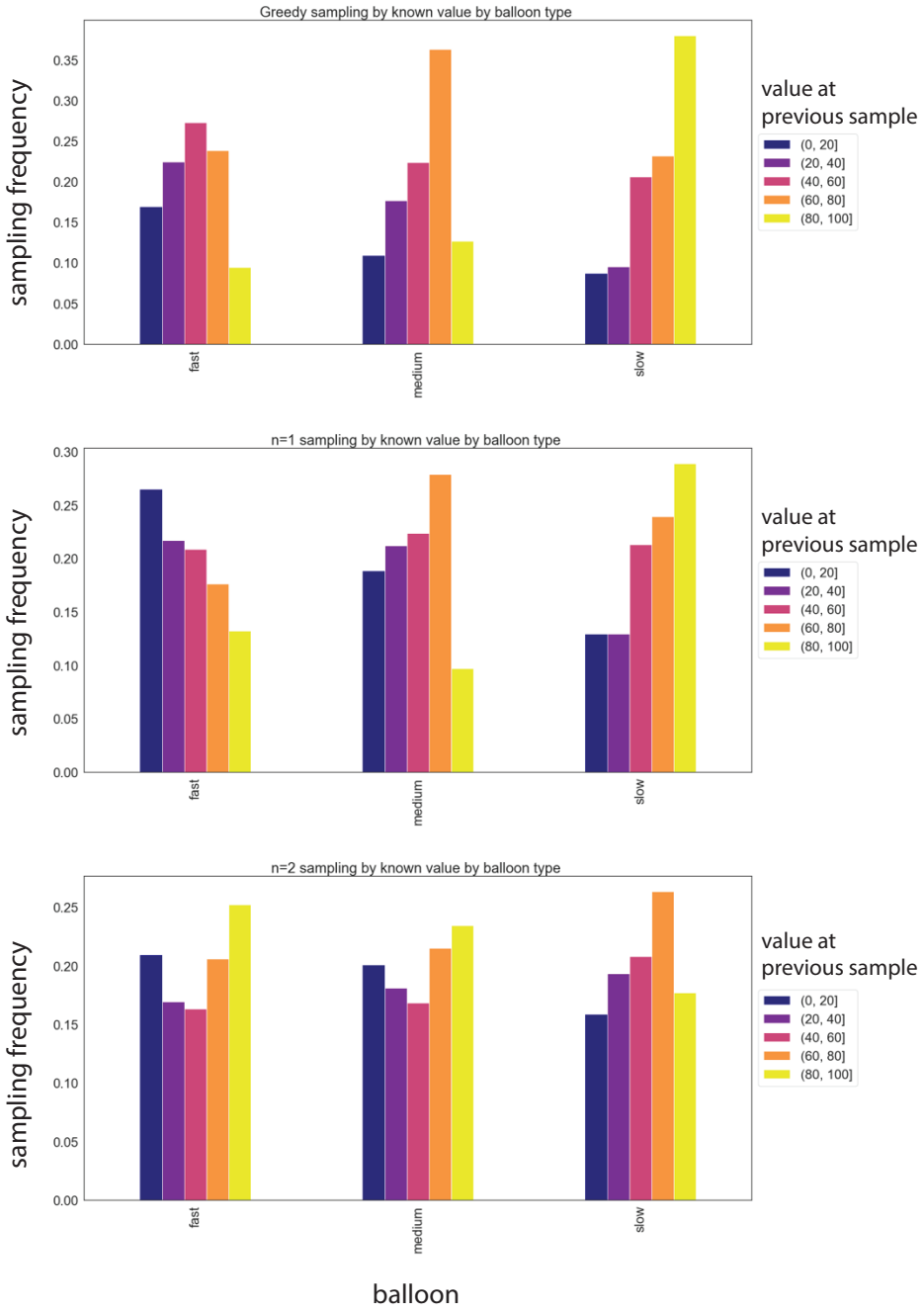


Figure 4.6 Sampling reflects how long ago a balloon was last sampled, depending on how fast that balloon increases in task version 2. Sampling frequency for each balloon is plotted split up by the balloon's value the last time it was sampled, on the previous trial (A), two trials ago (B) or three trials ago (C). It can be seen that while repeated sampling of the slow balloon when it was shown to have a high value on the previous trial is common, this is not the case for the other two balloons, presumably because participants know the other balloons are likely to have popped, while being uncertain about the slow balloon. Instead, repeated sampling of the medium balloon was most common when it was of reasonably good value on the previous trial (60-80 points) and repeated sampling of the fast balloon was most common when it was of mediocre value (40-60 points). The preferred value to sample each of the balloons decreased at predictable step sizes as trials passed since the last sample increased (B, C), suggesting participants used the balloons' different rates of change to inform their sampling decisions.

Furthermore, as in this task sampling is completely separated from choice, we were able to look at participants' exploration of new choice options, unconfounded by the need to maximize reward. In most other paradigms studying exploration, participants also receive the number of points associated with the option they explore, meaning their sampling decisions are not a pure measure of exploration, and they have to trade-off their need to reduce uncertainty with their need to maximize reward (Cohen et al., 2007; Kolling et al., 2012; Navarro et al., 2016; Trudel et al., 2021). In this version of the task, when a balloon's value was reset after having been sampled by a participant, this was indicated by the value displayed in the value bar being reset to a question mark. In other words, participants were aware when a balloon had been reset since they last sampled it. When a participant samples a completely unknown balloon this does not affect their reward, meaning this can only be driven by a need to reduce uncertainty on which balloon to choose on upcoming choice rounds. We found that such 'exploratory' samples of unknown balloons were more common in trials where the last sampled values of other balloons was in the medium range (40-60 points), suggesting that this exploration was indeed likely as a result of increased uncertainty about which balloon was currently best (Figure 4.7). In this task version it was difficult to directly compare exploration of the different balloons as a function of rate of change, though, as the balloon values reset independently from one another. In other words, all three balloons were rarely all unknown at the same time, making it impossible to compare exploratory sampling of the different balloons.

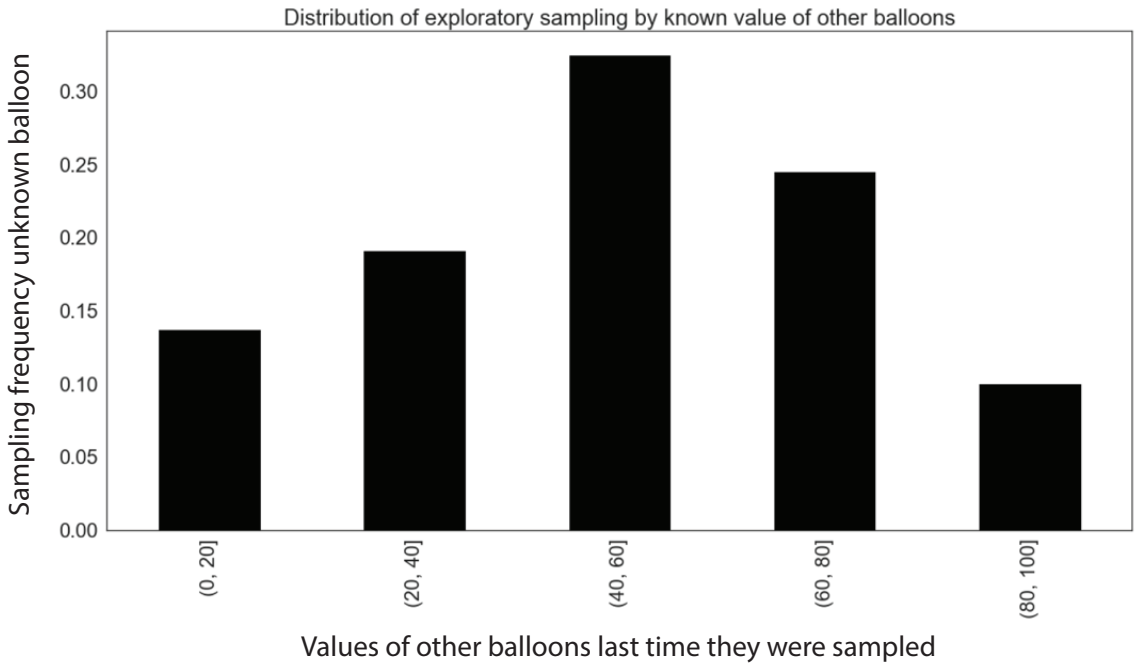


Figure 4.7 'Exploratory' samples were more likely when the last sampled values of the other two balloons was mediocre (40-60 points) and current uncertainty about which balloon was best was thus high. Plotted are the frequency of 'exploratory' samples (samples of a completely unknown balloon) as a function of the value of the other balloons when they were last sampled in task version 2.

In summary, participants were able to effectively use the sampling trials on this task to inform their choices, and this sampling behavior varied considerably between participants. These results suggested the existence of multiple sampling strategies, but we felt the task did not yet allow us to accurately distinguish different drivers of information sampling on this task. Specifically, the fact that balloons' probabilities of resetting were independent from one another meant that exploration of the three balloon types could not be directly compared.

4.2.4 Task version 3: No strong effect of rate of value change on sampling

In order to better investigate the effect of the rate of change of the balloons on sampling we changed the 'reset' trials and introduced 'reveal' trials in task version 3. Instead of having participants continuously sampling the balloons and each balloon randomly resetting occasionally, in this version trials were grouped in blocks of between 5-10 trials. At the start of a block all the balloons would either be reset, i.e. all their values changed and were hidden, or revealed, i.e. all their values changed and were revealed. Participants did not know at any point how many trials were left until a 'reset' or 'reveal' would take place. Participants were not asked to sample on 'reveal' trials as all the balloons' values were already known to the participant on those trials.

We recruited 30 online participants to perform this version of the task. 13 of these had to be excluded, because they misunderstood the task as evidenced by their answers on a debriefing questionnaire at the end of the experiment.

In this version, we changed 'reset' trials such that all balloons were reset at the same time to better compare balloon types on these trials. However, there is no general preference for one of the three balloons across all participants (Figure 4.8). As such, we were not able to draw any clear conclusion on the effect of the rate of change of balloon value on information sampling in this task. We attempted to model the different sampling strategies participants used, which is described in detail in the supplementary materials (section 4.6.2).

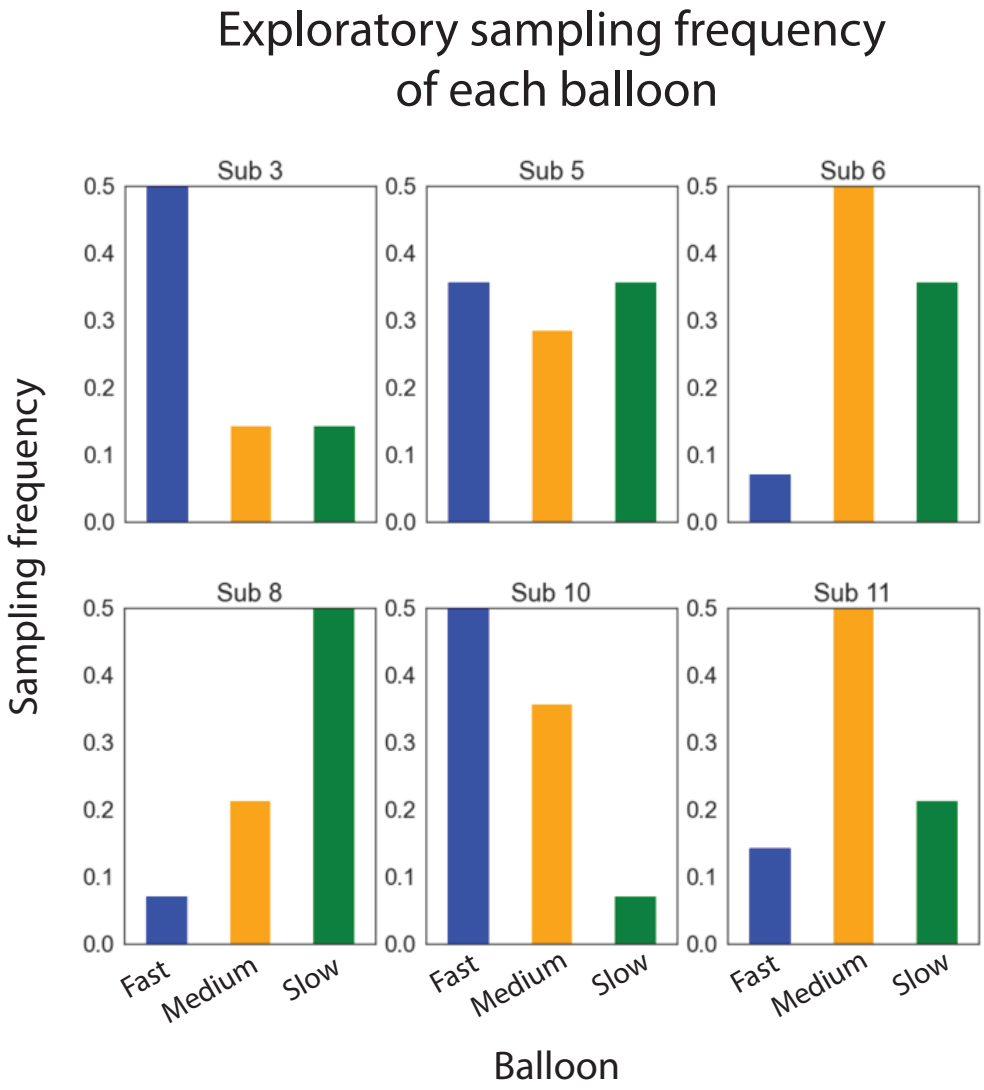


Figure 4.8 Sampling frequency of the three balloons by individual participants including only reset trials (in which all the balloons' values were reset and hidden) for some example participants in task version 3. It can be seen that participants vary substantially in their balloon preference on reset trials.

We suspected that part of the problem here was that on this task, uncertainty actually increased at a similar rate in all 3 balloons, as the width of the range of value increases for each of the balloon was the same (15 points) but only the mean of the range differed (28, 18 and 8 points, respectively). As such, the only way the balloons

differed in terms of uncertainty was that the value of more rapidly changing balloons would more often be harder to predict as they more frequently reached a point at which it was unclear whether they had popped yet. At these points, uncertainty is very high as there is some probability the balloon is of very high value (close to 100 points) as well as some probability it is of very low value (close to 1 point). To investigate the relationship between volatility (and its resulting uncertainty) and information sampling on this task more clearly, we designed a version of this experiment in which the uncertainty of the three balloons increased at different rates.

4.2.5 Task version 4: Manipulating variance of rate of value change affects sampling behavior, but is difficult for online participants to learn

We saw in the previous pilot that there was no clear effect of the rate of change of the balloons on sampling decisions across participants. In order to better investigate the effect of information resource volatility on sampling, we changed the rates of change of the three balloons. Specifically, in this version of the task each of the three balloons had a main value increase of 18 points per trial, but the range of possible increases varied: 3-33 points for the balloon with 'less' certainty, 11-27 points for the balloon with 'medium' certainty, 15-21 points for the balloon with 'more' certainty. The value by which each balloon increased was randomly sampled on each trial from their respective distributions. The task remained the same in all other ways.

17 online participants performed this task. Participants found this task much more difficult than previous versions, as evidenced by their lower mean reward as compared to previous pilots (compared with second pilot: $t_{40} = -2.76$, $p < 0.01$), and by participants' self-report in the debriefing questionnaire.

While participants' performance was less good on this task version than on previous versions, participants did seem to understand the balloons' different uncertainty levels. This can be seen by looking at how often a balloon was chosen as a function of how many trials ago it was last sampled (Figure 4.9). There was a main effect of how many trials ago a balloon was last sampled such that balloons were more likely to be chosen when they were sampled more recently. This effect was modulated by balloon type, where more certain balloons were more often chosen when sampled last longer ago, meaning participants were more willing to choose a balloon without having sampled it recently if this balloon had low uncertainty.

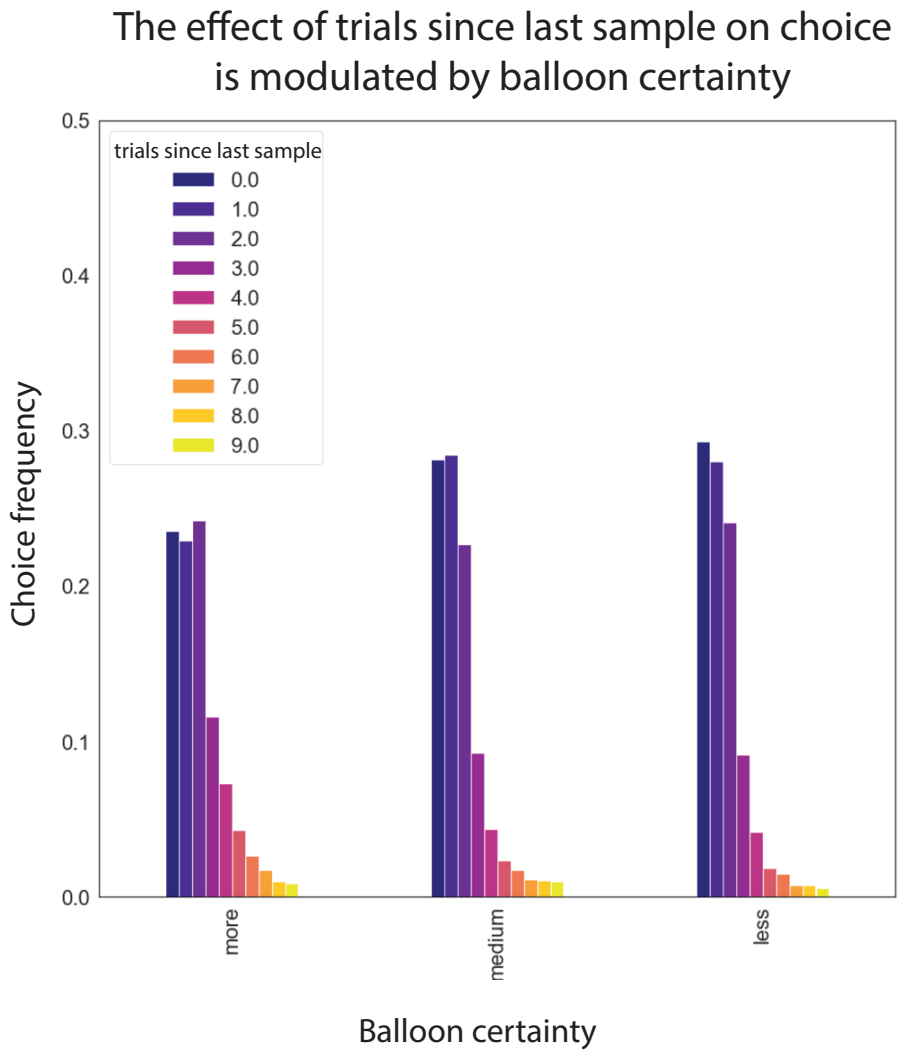


Figure 4.9 Participants were more willing to choose a balloon the shorter ago it was last sampled and this effect was stronger the less certain they were about the balloon's value. Plotted are the proportion of trials in which each of the three balloons were chosen where it was last sampled a particular number of trials ago.

This last result more convincingly shows an effect of volatility on information sampling than the previous task versions and could therefore be a promising task version for future investigation. However, due to online participants' poor performance on this task, we doubted its usefulness as a task for online data collection and thus chose not to pursue data collection with task version 4 further. This version of the task would probably be suitable and interesting for investigation in the lab in-person so the experimenter is better able to ensure the participants properly understand the task before data collection.

4.2.6 Task version 5: Goal proximity not a strong predictor of sampling

We adapted task version 3 to be more suitable for use with fMRI. We expected that a sampling decision between three balloons would not be the best way to detect a neural signature for information sampling, so we decided to instead change the sampling decision into an accept/reject-type decision. We did this by picking one of the three balloons at random on each sampling trial and highlighted it. Participants could either decide to sample or reject to sample. However, in order to compare sampling to *not* sampling in the brain imaging data, we made sampling costly, so that participants would not simply choose to sample on each trial, but only when they believed the benefit of sampling exceeded the cost. A cost bar would appear next to the highlighted balloon indicating a cost for sampling between 1-20 points.

Furthermore, as we previously decided to eliminate the second resource (silver balloons) in this task, and task version 3 was structured in a way that each sampling trial was followed by a choice trial, goal proximity was not a big driver of information

sampling in the results so far. Therefore, we structured trials on this task version into blocks of 3-8 sampling trials, followed by a single choice trial. The number of trials a participant had left until a choice trial was visible on the screen at all times. We expected that sampling decisions would be more frequent on sampling rounds closer to the choice round, as information gained on these rounds was more likely to remain relevant as balloons were less likely to pop in the meantime. On a choice trial, participants could pick any of the three balloons, and after choosing all the balloons' values were revealed. The participants' reward outcome on the choice trial at the end of a block corresponded to the points contained in the chosen balloon subtracted by the total amount of points spent on sampling that block. After a choice trial, all the balloons were hidden again and reset to a new value at the start of a new block.

We collected data from 30 online participants on this task. We excluded three of them as they misunderstood the task as evidenced by their responses on a debriefing questionnaire.

As we expected, sampling was strongly driven by cost, where sampling costs above 15 points rarely led to sampling (Figure 4.10). We predicted this relationship might be mediated by the rate of change of the balloon, as it would make sense for participants to be willing to spend more points on information that would remain relevant for longer. In other words, we predicted the effect of sampling cost to be less strong the lower the rate of change of the balloon. However, we found that the relationship between cost and sampling was unaffected by balloon type. This negative relationship between sampling cost and sampling was present in most participants, though the strength of the correlation varied a lot (Figure 4.11).

We performed a hierarchical logistic regression model predicting sampling on each trial with data grouped by participant, including as predictors sampling cost, rate of change, trials left until choice trial, number of samples already made on this block, and their interaction terms. The only significant predictors of sampling were sampling cost ($z=-4.32$, $p<0.001$), how many samples were already made on that block ($z=-5.64$, $p<0.001$), and the interaction between cost and how many trials were left until a choice trial ($z=-3.95$, $p<0.001$).

We hypothesized that goal proximity was not a strong driver of sampling because of how the balloons' values change on this task. In principle, because the balloons changed in value in a predictable way, even samples many trials away from the choice trial could still be informative. As the total number of samples already made was a predictor of sampling probability, it is likely that this is why goal proximity did not significantly affect sampling. This is the same problem we encountered in the pilot data from task version 3, which led us to design task version 4, which was too difficult for online participants to do well. We addressed this by changing the way the balloons change value in task version 8 (see **section 4.4** and **chapter 6**).

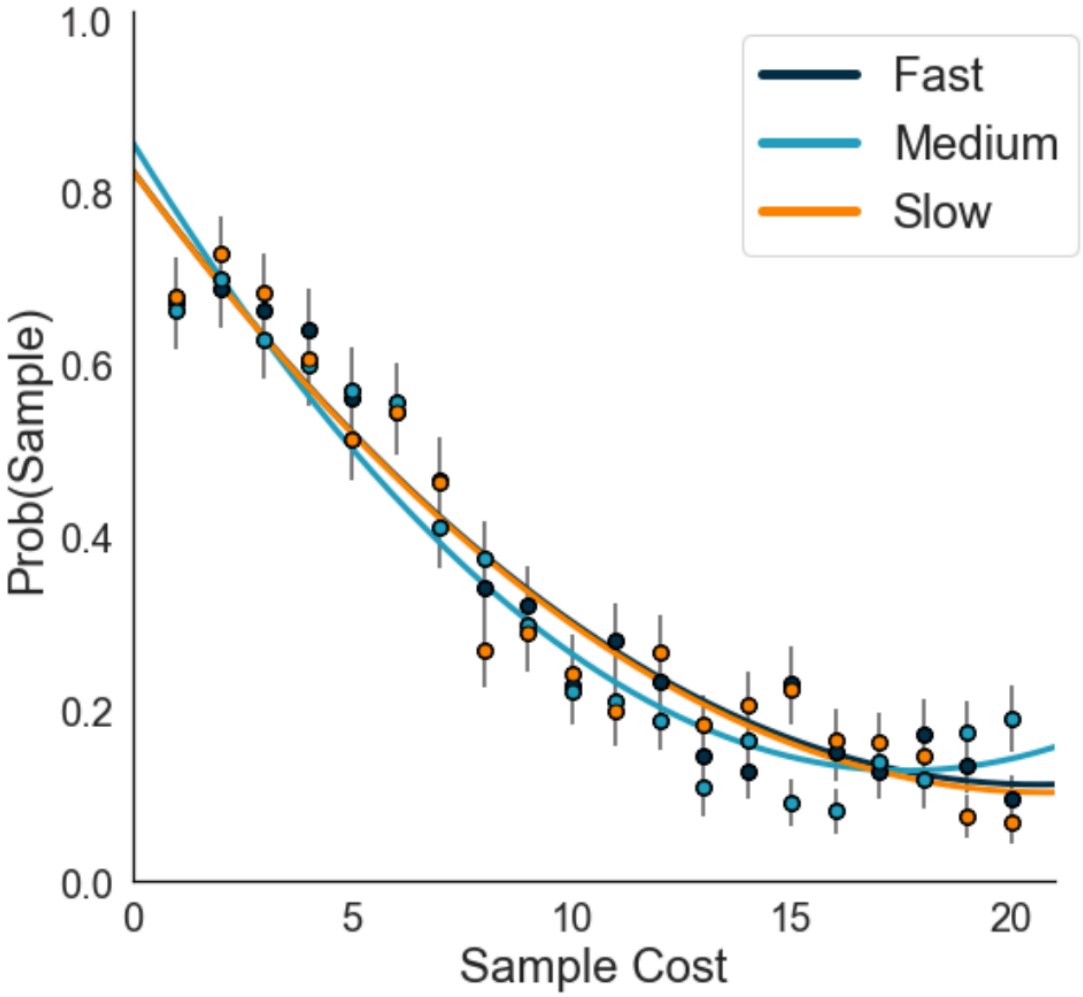


Figure 4.10 Sampling is less likely the more expensive the sample, and this is not affected by balloon type. Plotted is the probability of accepting a sample as a function of sample cost split by balloon type in task version 6.

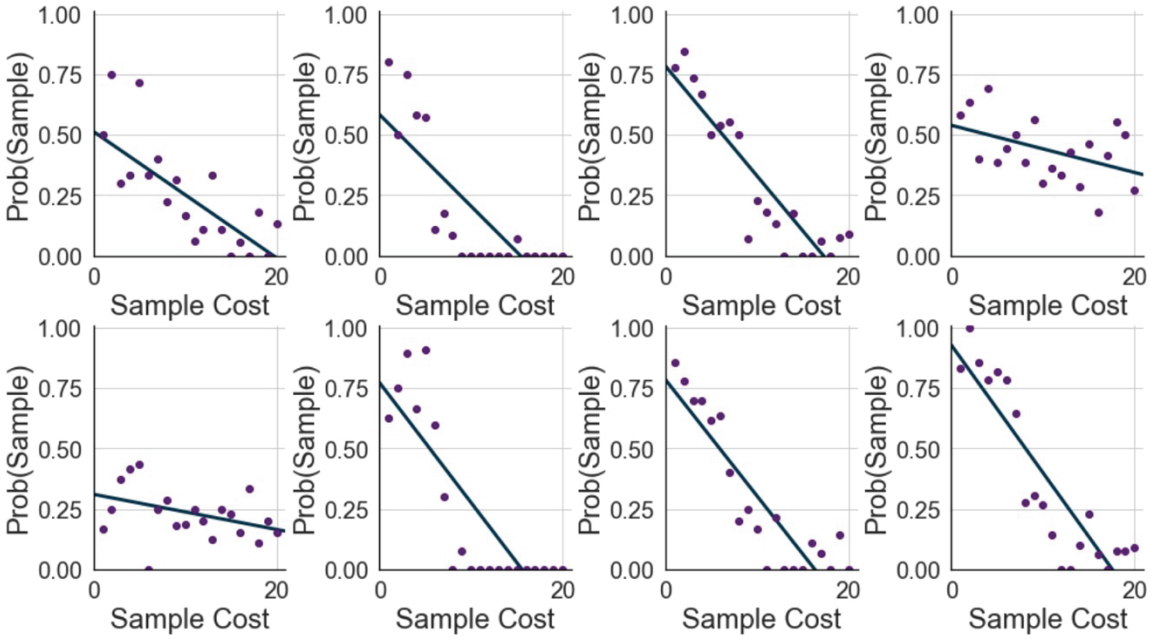


Figure 4.11 The relationship between sampling cost and probability of sampling varied a lot across participants. Plotted are eight example participants from pilot data using task version 6.

4.2.7 Task development conclusions

From piloting several task versions, we discovered a number of flaws and limitations of the task that we have tried to address to develop two task versions for use in a large-scale online study (chapter 5) and an fMRI study (chapter 6). Firstly, we tried to induce prospective sampling by including two goals in the task (section 4.2.2) but found that participants rarely sampled information for future goals. We therefore made a version of the task with a single goal, in which participants effectively used the sampling rounds to inform their choices (sections 4.2.3-4). However, we were unable to draw many general conclusions about the drivers of sampling in these versions, as participants showed a wide range of sampling strategies. Besides this, no effect was found of the rate of change on sampling, so participants did not seem to use the predicted length of time that information would remain relevant to guide their choices

on these tasks. A version in which the rate at which uncertainty increased over time varied more strongly across balloons did produce this effect, but was deemed too difficult for participants to learn (section 4.2.5).

We used a version of this task to study the effect of psychiatric dimensions on overall information sampling as well as on the sampling strategy used, as we hypothesized the heterogeneous behavior seen in this task could be interesting to investigate in relation to previously investigated relationships between psychiatric (or subclinical) symptoms and information sampling. This task version is described below in section 4.3 and its results in chapter 5.

A version of the task more suitable for fMRI was also designed by making sampling decisions binary, but here we also found no effect of rate of change and goal proximity on sampling decisions (section 4.2.6). We hypothesized that this was due to the predictable rates of value increase of the balloons, meaning there was no strong need to sample some balloons more than others or to sample close to the choice round. This task limitation was addressed in the final task version we used in our fMRI study (see section 4.4 below and chapter 6).

4.3 – Unlimited sampling - task version 6

Up to this point, we had mostly focused our task development on designing a version that could show information sampling is driven by features of the information source and the goal, specifically goal proximity and rate of change. In other words, we want to investigate *which* of multiple information resources is sampled to guide choice. The final version of the task that achieves this is presented in section 4.4 and is used in an fMRI study presented in chapter 6. However, we realized that this task could also

be used to investigate *whether* people choose to sample information at all and how much.

Specifically, compulsivity has been related to increased information sampling in previous literature, but it is unclear whether this is due to inflexible, repetitive behavior or altered decision-making parameters (see chapter 5). As we found such a wide range of sampling behavior in our pilot data (section 4.2.3), we decided it could be interesting to combine our task with a set of psychiatric questionnaires to assess whether these heterogeneous sampling patterns could be related to transdiagnostic psychiatric dimensions.

To be able to relate the frequency of information sampling to psychiatric dimensions, we designed a version of this task in which participants could sample however many balloons they wanted on each sampling trial. In this version, participants could continue to sample balloons on a sampling trial until they sampled all of them, or until they elected to move on to the choice trial. However, each sample came at a waiting time cost of 3s and the total experiment time was restricted to 30min. In other words, more samples would lead participants to be able to make fewer reward-maximizing choices within the allocated time. The task remained the same as task version 3 in all other details. We collected data from a total of 450 online participants using this task in combination with a set of psychiatric questions (see **chapter 5** for the results). This version of the task can be tried out at https://run.pavlovia.org/paulak/balloontask_v6_blocked .

4.4 – Values change unpredictably – task version 7

4.4.1 Task description

As we did not find an effect of goal proximity and rate of change of the balloons in previous task versions (sections 4.2.2-3, 4.2.5), we hypothesized this might be because uncertainty did not change linearly over time. Therefore, we made a version where uncertainty increased linearly. in this version the balloons no longer increased in value until 'popping' and then increasing again from 1 point. Instead, whenever a balloon 'popped' it would be reset to a random value between 1-100 points. The three balloons differed in 'popping' frequency, such that balloons had a 12.5%, 25% or 50% chance of popping on any given trial, respectively. This 'popping frequency' was displayed for each balloon on the screen at all times, so that there was no memory load for participants as they performed the task. We predicted that this would make it both easier to model the effects of uncertainty on sampling, as well as easier for participants to understand the effect of the popping frequency of each balloon on how long sampled information remained useful, compared to the previous task versions. We therefore predicted that in this version, participants were more likely to use popping frequency to make sampling decisions.

Furthermore, as in task version 6 it was seen that participants rarely sampled when sampling costs were over 10 points, in this version sampling costs were randomly sampled between 1-10 points. This version of the task can be tried out at https://run.pavlovia.org/paulak/balloons_instr/html/ .

4.4.2 Pilot results

We recruited 15 online participants to perform this version of the task. We excluded 1 participant because they appeared to have misunderstood the task based on their answers on the debriefing questionnaire.

Participants performed above chance on this task (Figure 4.12A) and sampling frequency was higher on this task version than on task version 6 (Figure 4.12B): samples were accepted on about 40% of trials.

As we hypothesized, popping frequency was now a strong predictor of sampling in this task version, where balloons that pop less frequently were sampled more often by most participants (Figure 4.13).

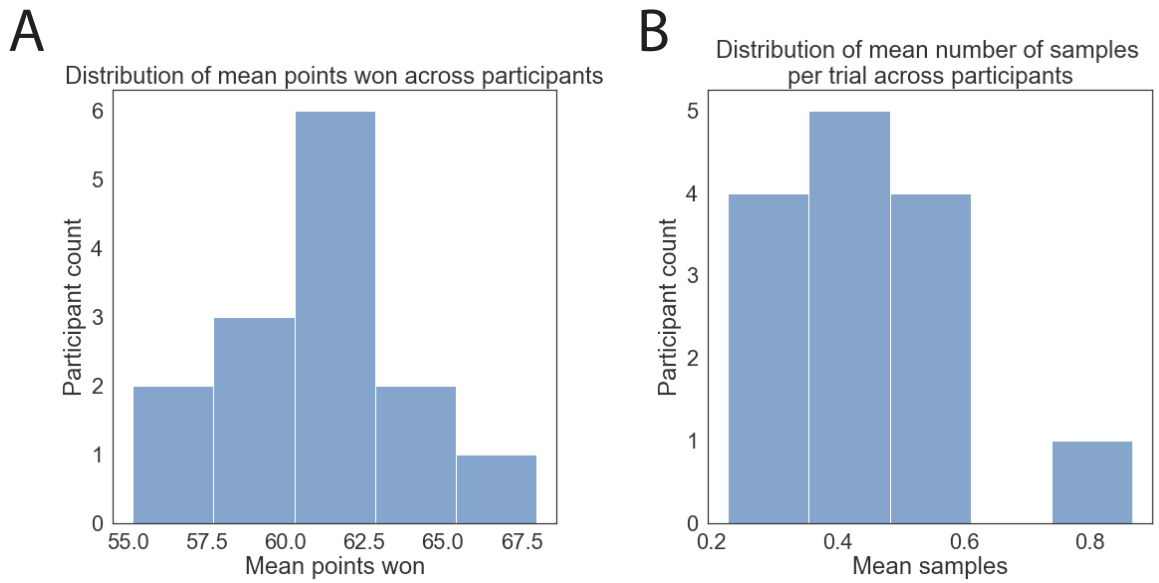


Figure 4.12 Participants performed above chance on task version 7. (A) Plotted are the average number of points won on a choice trial. Note that sampling costs have not been subtracted from these values. (B) Proportion of trials where samples were accepted for each participant in task version 7. Most participants accepted samples on around 40% of trials.

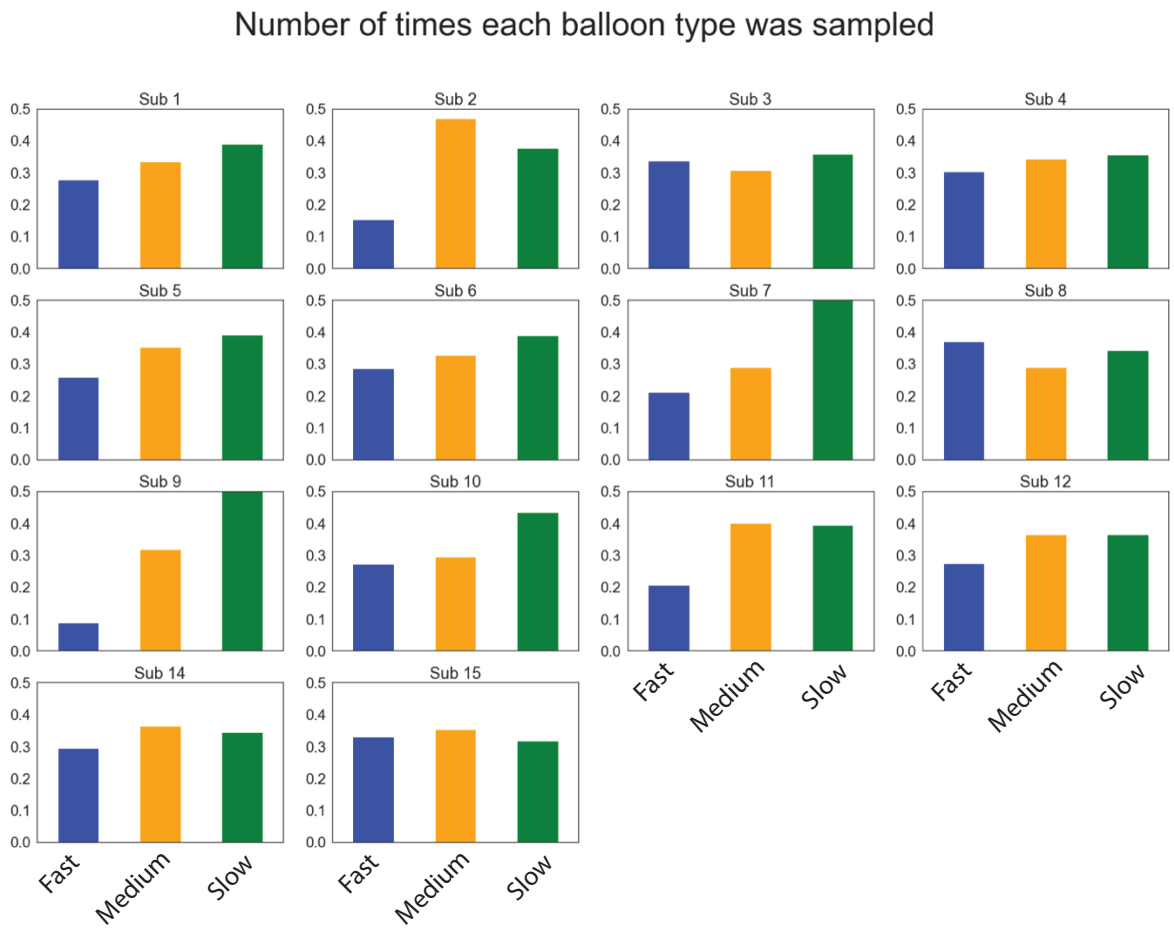


Figure 4.13 Frequency of sampling of each balloon type for each participant in task version 7.

Furthermore, the number of trials left before a choice trial, i.e. goal proximity, was also a strong driver of sampling in this version of the task, as predicted. Participants were less willing to sample balloons when there were still many sampling rounds left and more willing to do so close to a choice trial (Figure 4.14). This effect was mediated by balloon type, with the effect of trials left being stronger the higher a balloon's popping frequency.

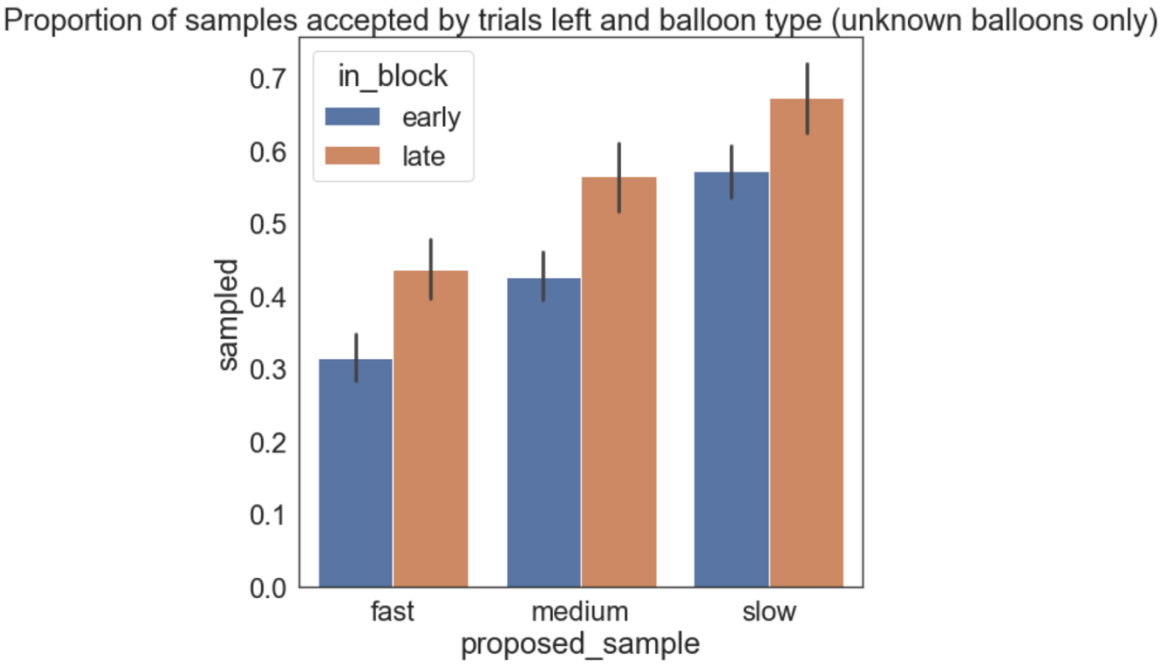


Figure 4.14 Sampling was more likely the closer participants were to a choice trial and this was mediated by balloon type in task version 7. Plotted is the proportion of trials where a sample was accepted by balloon type and whether a trial was early (5-8 trials before choice) or late (1-4 trials before choice) in a block.

Sampling was affected by cost, as before, where more costly samples are less likely to be accepted (Figure 4.15). On this version, however, this effect was also mediated by popping frequency of the balloon, where less costly samples increase the negative effect of popping frequency on sampling.

As we predicted, making value uncertainty of the balloons change linearly over time and then varying this rate of change across the balloons made participants take this rate of change into account in their sampling decisions, in contrast to previous task versions. Furthermore, goal proximity was now also a driver of sampling. In summary, in this task version, sampling is now driven by rate of change, goal proximity and sampling cost. As this task version seemed to have addressed the issues with previous task versions, we used this version in an fMRI experiment (see **chapter 6**).

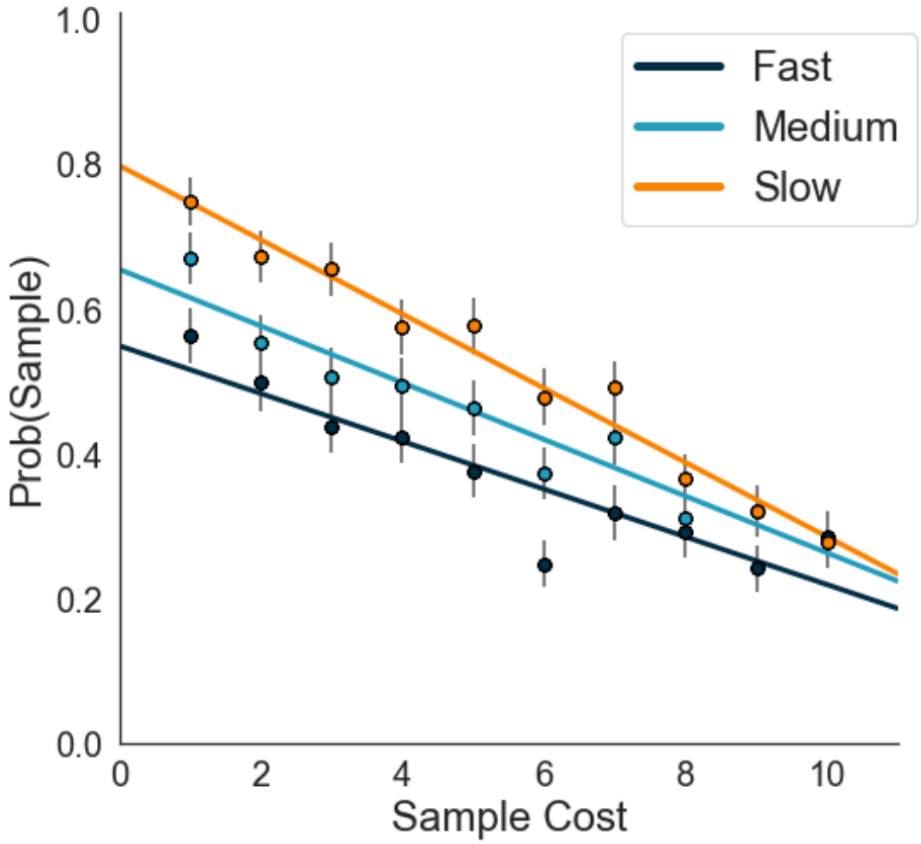


Figure 4.15 Sampling is less likely the more expensive the sample, and this is mediated by balloon type, such that popping frequency is a stronger driver of sampling when sampling is cheap. Plotted is the probability of accepting a sample as a function of sample cost split by balloon type in task version 7.

4.5 Summary

We had set out to develop a novel decision-making task where participants could actively actively sample information, sampling decisions and reward maximization are decoupled, and the volatility of information sources and goal proximity could be manipulated. All the task versions described in this chapter allowed participants to actively sample and successfully decoupled information sampling from reward maximization. However, it took a number of pilots using different task versions to find a version where volatility and goal proximity reliable predicted sampling. To

study the effects of psychiatric dimensions on sampling frequency, task version 5 was used in a large online experiment described in Chapter 5. To investigate the neural signatures of information sampling, task version 7 was used in an fMRI experiment described in Chapter 6.

4.6 Supplementary materials

4.6.1 An optimal model of task version 2

We modelled an optimal agent's behavior (optimal in the sense of maximizing expected average reward) on this task in order to compare participants' behavior to model predictions. This agent uses its knowledge of how the balloons change to compute an expected value for each balloon:

$$EV_{balloon} = \sum_{n=1}^{100} (P(V_{balloon} = n) \times n)$$

where $V_{balloon}$ is the balloon's value and $P(V_{balloon} = n)$ is the probability that the balloon's value is n (which can be between 1-100 points). The probability that a balloon is best at a particular value can be computed by the product of the cumulative sum of probabilities of the other two balloons up to that value (Figure S4.1):

$$P(V_{balloon1} = \max(V) \mid V_{balloon} = x) = \sum_{n=1}^{x} P(V_{balloon2} = n) \times \sum_{n=1}^{x} P(V_{balloon3} = n)$$

In other words, for each possible value a balloon can have (1-100 points), the probability that a balloon is the best of all three balloons at that value is the probability that the two other balloons have a lower value. To get the overall probability that a balloon has the highest value, the agent takes the product of the probability a balloon has a specific value and the probability the balloon is the best balloon at that value and sums across all possible values (1-100 points; Figure S4.1):

$$P(V_{balloon} = \max(V)) = \sum_{n=1}^{100} (P(V_{balloon} = n) \times P(V_{balloon} = \max(V) \,|V_{balloon} = n))$$

The agent computes the value of checking each balloon based on how likely checking a balloon would be to change its policy on the upcoming choice trial and the potential gain in value of that policy change (Mattar & Daw, 2018). The value of checking the currently estimated best balloon can be computed by multiplying the probability that the value of this balloon is in fact lower than the currently estimated second best balloon and the number of points the agent could gain by choosing this second best balloon over the currently estimated best balloon. Specifically, when a balloon is currently estimated to have the highest value of the three balloons, the expected value of the second best balloon is used to compute the point gain of switching to the second best balloon at every possible value lower than the second best balloon. The value of checking the estimated best balloon is then obtained by multiplying this potential gain if the balloon has a specific value with the probability that it has this value and summing across for all values lower than the expected value of the estimated second best balloon:

$$V_{check,balloon} = \sum_{n=1}^{EV_{secondbest}} (P(V_{balloon} = n) \times (EV_{secondbest} - n))$$

The logic here is that *if* the value of a balloon that is currently estimated to be highest is in fact worse than the expected value of the second best balloon, the agent would

only switch to this other balloon if it had sampled the estimated best balloon, giving the action of checking a value in itself.

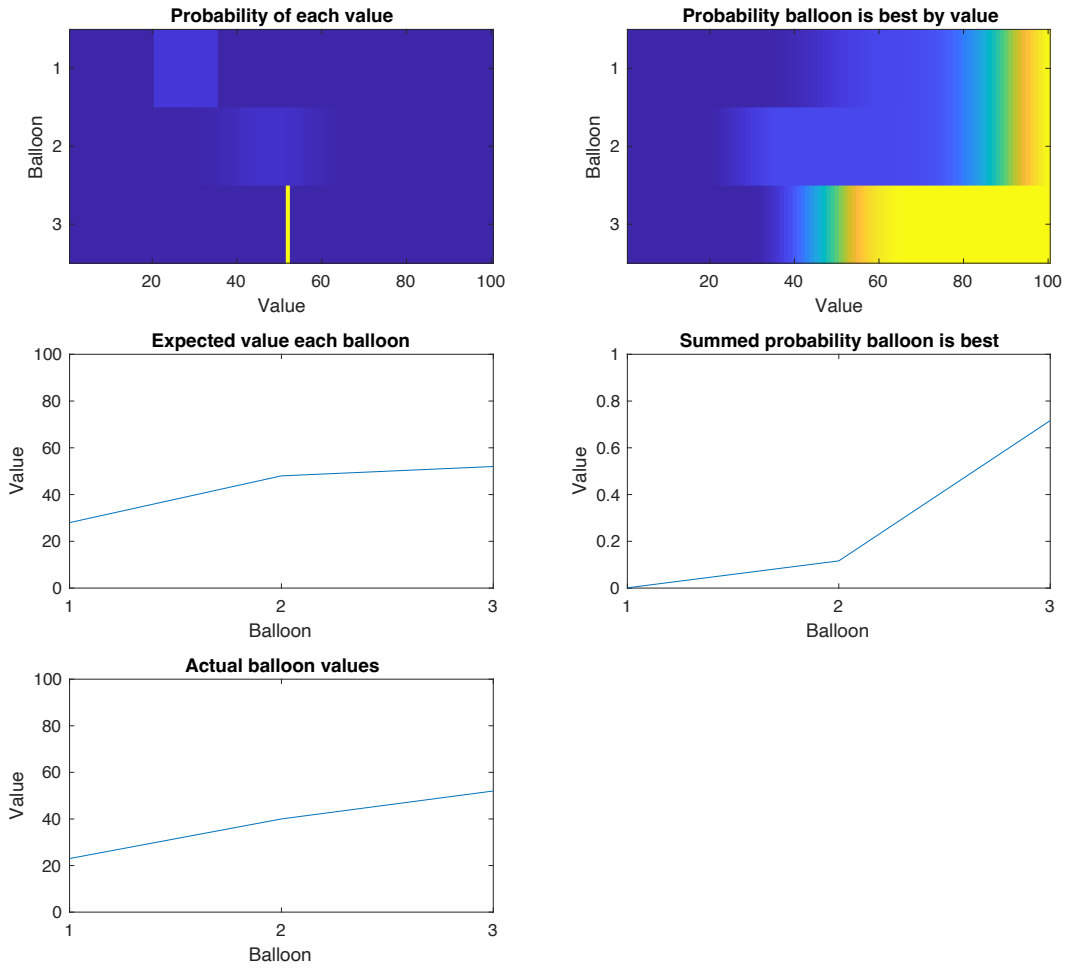
To find the value of checking for balloons that are *not* estimated to currently have the highest value, the agent needs to compute the probability that the estimated value of that balloon is in fact higher than the currently estimated best balloon, and the potential point gain from choosing this balloon instead of the currently estimated best balloon. The expected value of the currently estimated best balloon is used to compute the potential gain of switching from the best balloon to this other balloon. This potential gain can again be multiplied with the probability of the balloon's value and summing across all values higher than the expected value of the estimated best balloon this leads to the value of checking of this other balloon:

$$V_{check,balloon} = \sum_{n=EV_{best}}^{100} (P(V_{balloon} = n) \times (n - EV_{best}))$$

Again, here the value of checking exists because the agent would only find out a balloon had a higher value than the expected value of the estimated best balloon and switch to the former if it had checked it.

The agent uses the values of checking for each balloon in a softmax function to decide which balloon to sample on sampling trials. On choice trials, the agent decides which balloon to choose using the probabilities each balloon is best in a softmax function.

A


Probability of each value
Balloon
Value
Probability balloon is best by value
Balloon
Value
Expected value each balloon
Value
Balloon
Summed probability balloon is best
Value
Balloon
Actual balloon values
Value
Balloon

B

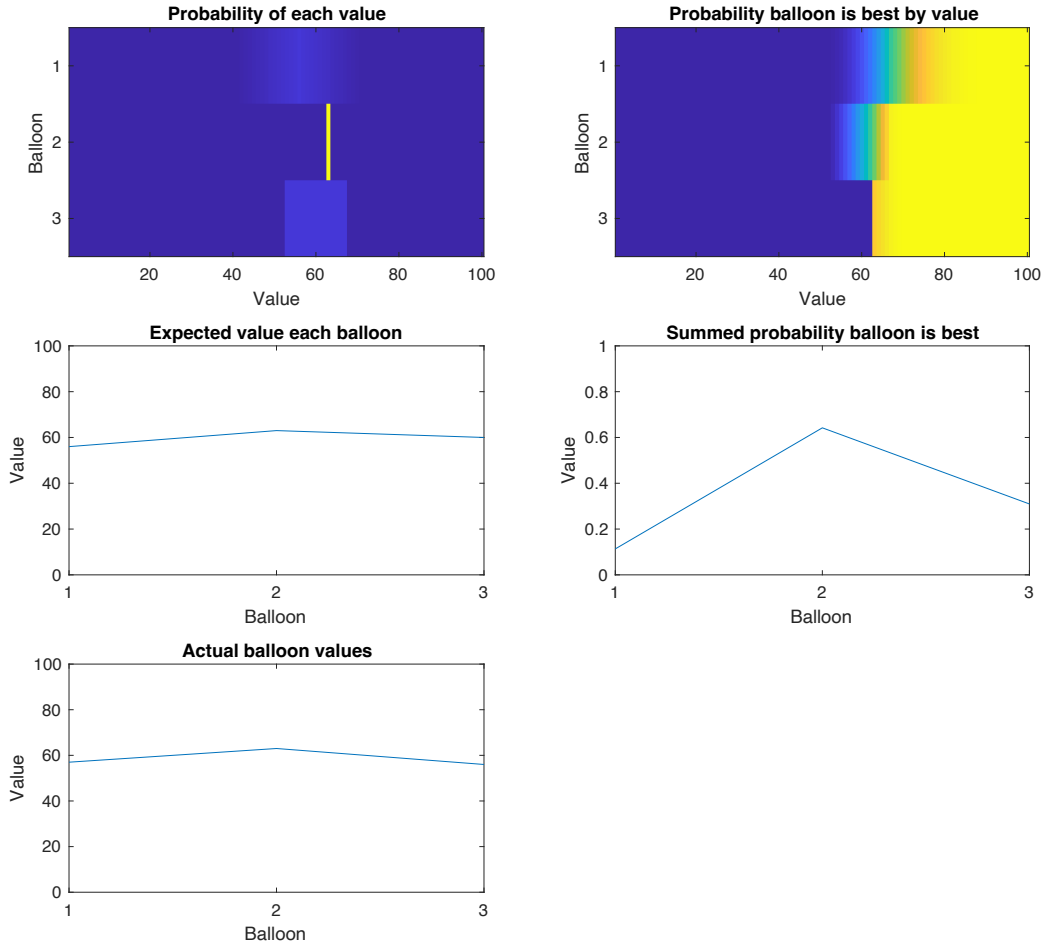


Figure S4.1 Simulation of two subsequent sampling trials (A: trial 1, B: trial 2) in an optimal model of task version 2. It can be seen that when a balloon has just been sampled, the probability is 1 for its actual value and all other probabilities are set to 0. This distribution becomes wider and its mean increases in future trials as the agent uses its knowledge of the range of possible value increases of each balloon, but it also becomes more uncertain over time. The probability that a balloon is best if it has a specific value can then be calculated by computing the product of the cumulative sum of the probabilities of each value of the other two balloons. To choose a balloon, the agent multiplies the probabilities of each value of each balloon with the probability that a balloon is best at each value to find the largest summed probability.

We made a second version of this optimal model that instead of just optimizing sampling behavior for the upcoming choice trial takes into account future trials. In other words, a model that uses *prospection*. In this model, the value of checking is a sum of the potential gain of a policy change across n future choice trials. This is computed by

simulating the probabilities of each value for each balloon in future trials. The value of checking the estimated best balloon on trial *t* then becomes:

$$V_{check,balloon,t} = \sum_{n=1}^{EV_{secondbest,t}} \left(P\left(V_{balloon,t} = n\right) \times \left(EV_{secondbest,t} - n\right)\right)$$

$$+ \sum_{n=1}^{EV_{secondbest,t+1}} \left(P\left(V_{balloon,t+1} = n\right) \times \left(EV_{secondbest,t+1} - n\right)\right)$$

$$+ \ . \ . \ .$$

$$+ \sum_{n=1}^{EV_{secondbest,t+k}} \left(P\left(V_{balloon,t+k} = n\right) \times \left(EV_{secondbest,t+k} - n\right)\right)$$

where *k* is the number of trials the agent can look ahead, $V_{balloon,t}$ is the value of a balloon at trial *t*, and $EV_{secondbest,t}$ is the expected value (calculated as outlined above) of the estimated second best balloon at trial *t.* The value of checking a balloon that is *not* estimated to be the best is:

$$V_{check,balloon,t} = \sum_{n=EV_{best,t}}^{100} \left(P\left(V_{balloon,t} = n\right) \times \left(n - EV_{best,t}\right)\right)$$

$$+ \sum_{n=EV_{best,t+1}}^{100} \left(P\left(V_{balloon,t+1} = n\right) \times \left(n - EV_{best,t+1}\right)\right)$$

$$+ \ . \ . \ .$$

$$+ \sum_{n=EV_{best,t+k}}^{100} \left(P\left(V_{balloon,t+k} = n\right) \times \left(n - EV_{best,t+k}\right)\right)$$

where $EV_{best,t}$ is the estimated expected value of the estimated best balloon at trial *t*. This is not strictly an optimal model, as it does not take into account intermediate future samples that would affect the probability distributions of the balloons, but we are using it as an approximate model to be able to make some behavioral predictions.

As we hypothesized that information sampling should depend on volatility of the information source as well as goal proximity, we looked at whether this optimal model had a preference for sampling any balloon in particular. In the model without prospection (where the value of checking only takes into account the potential gain of a policy change on the upcoming choice trial), the model most frequently samples the 'fast' balloon (with the largest value increase per trial), and least frequently samples the 'slow' balloon (Figure S4.2). Presumably this is because the rapid value increase of this balloon causes it to 'pop' more frequently, making its expected value more uncertain.

However, this pattern changes when prospection is introduced in the model. Specifically, the model starts sampling the medium and slow balloons more frequently, the more trials into the future it takes into account when computing a balloon's value of checking (Figure S4.2). This makes sense as while the 'fast' balloon has a higher value of checking due to its increased uncertainty, the other two balloons have a higher value of checking due to being more informative for future trials.

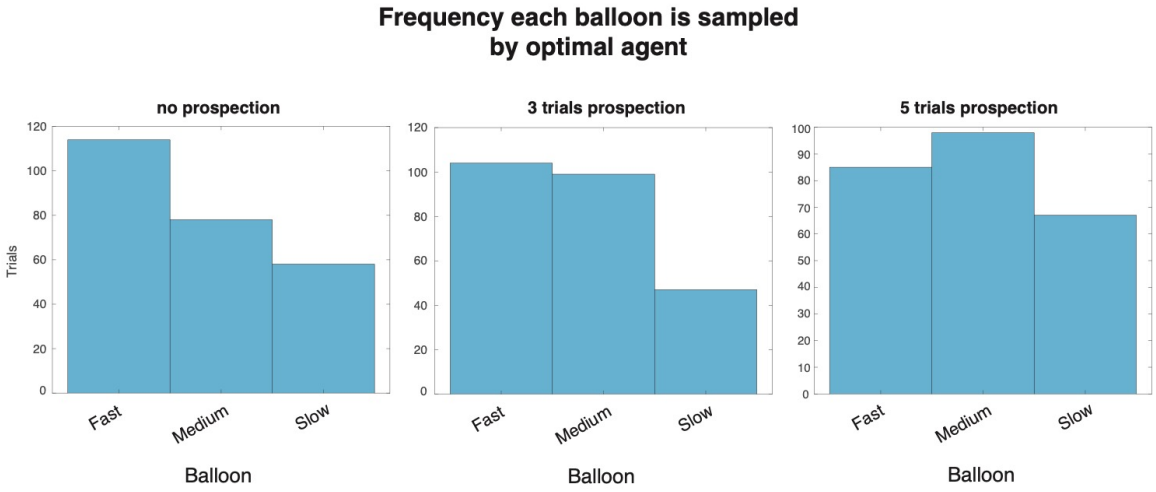


Figure S4.2 Simulations of samples from optimal model with no prospection, 3 trials of prospection or 5 trials of prospection. Plotted are the number of trials on which each of the three balloons were sampled by the model. It can be seen that the further the model looks into the future the more its preference moves from balloon 1 (the balloon increasing at a fast rate) to balloon 2 (the balloon increasing at a medium rate).

4.6.2 Modelling individual differences in behavior in task version 3

The observation that there were potentially distinct multiple sampling strategies that participants were using in this task, led us to develop a model that might disentangle those. Specifically, we developed a model including four distinct hypothesized strategies that could be weighted to produce an averaged behavior. The weights for each strategy could then be estimated based on each participant's data.

The first strategy we hypothesized participants might use is a random sampling strategy, where the probability of sampling a balloon $P(sample)$ is equal for the three balloons on each trial:

$$P(sample) = [\frac{1}{3}, \quad \frac{1}{3}, \quad \frac{1}{3}]$$

The second strategy is a 'sequential' sampling strategy where each balloon is sampled equally often in a sequential order:

$$P(sample) = [\, f(n_{balloon1}), f(n_{balloon2}), f(n_{balloon3})]$$

where *n* is the number of trials since a balloon was last sampled and $f(n_{balloon})$ indicates that the probability of sampling a balloon is some function of the number of trials that have passed since that balloon was last sampled. Many participants appeared to do this as evidenced by sampling being most frequent when a balloon was last sampled 3 trials ago (Figure 4.5).

The third strategy is a 'sample the best balloon' strategy in which the balloon currently estimated to be the best balloon continues being sampled on each trial until it is no longer the best estimated balloon (because it popped):

$$\begin{aligned} P(sample) = [\, & f\big(P(EV_{balloon1} = \max(EV))\big), \\ & f\big(P(EV_{balloon2} = \max(EV))\big), \\ & f\big(P(EV_{balloon3} = \max(EV))\big)\,] \end{aligned}$$

where the probability of sampling a balloon is some function of the probability that that balloon is currently estimated to be the best one ($f(P(EV_{balloon} = \max(EV))$). This strategy may explain 'greedy' sampling, as participants might repeatedly sample and choose a high valued balloon until it 'pops', at which point they try to find another high valued balloon.

Finally, the fourth strategy included in the model is a 'sample the chosen balloon' strategy, in which the balloon that was just chosen on the previous choice trial is sampled:

$$P(sample) = [\, f\big(Chosen_{balloon1}(t-1)\big),$$
$$f(Chosen_{balloon2}(t-1)),$$
$$f(Chosen_{balloon3}(t-1))\,]$$

where $Chosen_{balloon}$ is a dummy variable describing if a balloon was chosen on a trial (1) or not (0) and $f\big(Chosen_{balloon}(t-1)\big)$ indicates the probability of sampling a balloon is some function of whether or not that balloon was chosen on the previous trial *t-1*. We hypothesized participants might do this to know their pay-off on the previous choice trial as they received no feedback.

We combined these strategies by using the concept of value of checking described in supplemental materials section 4.6.1 and defining the value of checking a balloon on a certain trial *t* as a weighted function of the task parameters driving these four different strategies with five free parameters:

$$V_{check,balloon}(t) = \alpha + \beta \times n_{balloon}(t) + \gamma \times Unknown_{balloon}$$
$$+ \delta \times P(EV_{balloon} = \max(EV), t)$$
$$+ \zeta \times Chosen_{balloon}(t-1)$$

where α is a constant and thus describes the weight of the random sampling strategy. β weights the effect of how many trials passed since a balloon was last sampled and

captures the sequential sampling strategy. γ weights the effect of whether or not a balloon is unknown (described by dummy variable $Unknown_{balloon}$) and is included to account for samples of reset balloons, as these cannot be described in terms of trials passed $n_{balloon}(t)$. δ weights the effect of the estimated probability that a balloon has the highest value and captures the 'sample the best balloon' strategy. Finally, ζ weights the effect of prior choice, capturing the 'sample the chosen balloon' strategy, and $Chosen_{balloon}(t-1)$ is a dummy variable indicating whether the balloon was chosen on the previous trial.

To see if this approach indeed produces behavior in line with the hypothesized strategies, we simulated data setting the free parameters to values that predict the use of only a single strategy.

Firstly, we set α to 1 and the other free parameters to 0, so:

$$V_{check,balloon} = 1$$

Unsurprisingly, using this value of checking produced random samples with each balloon sampled roughly equally often, but in no particular order (Figure S4.3).

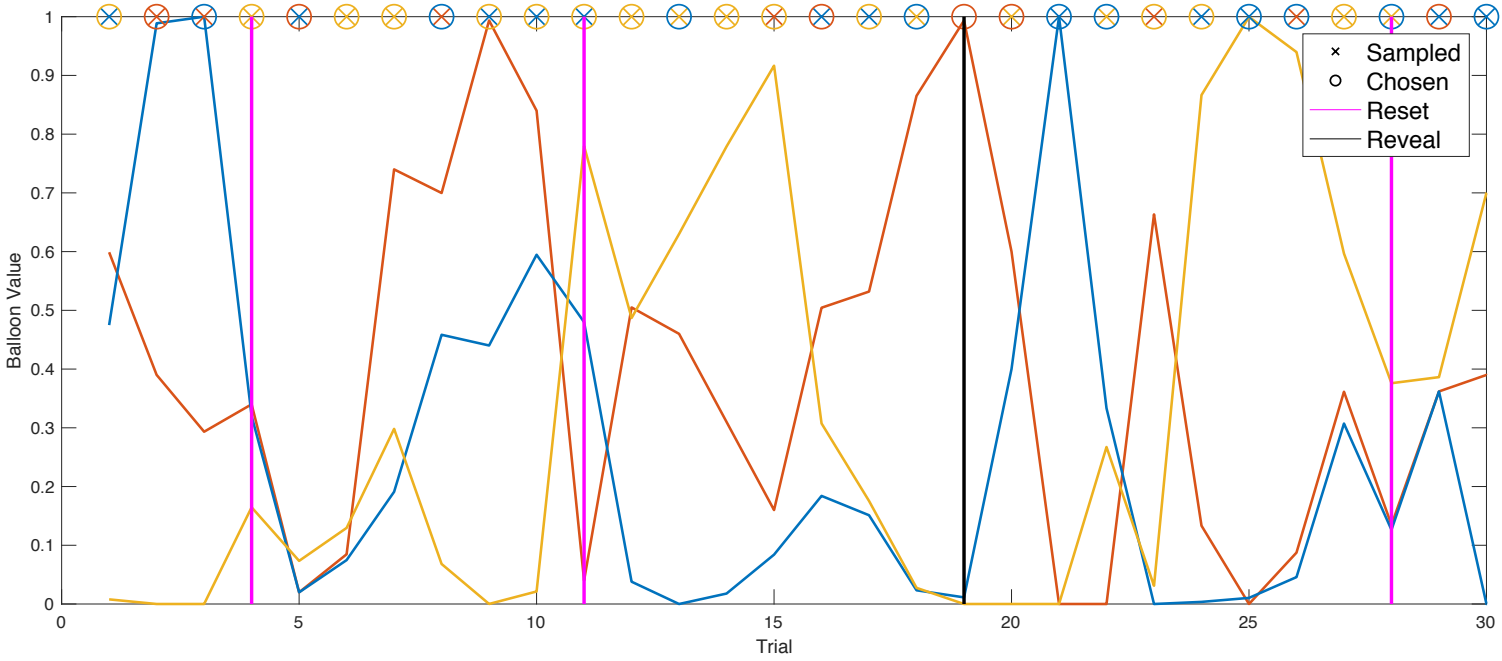


Figure S4.3 Simulated data from model with only free parameter α. The model is sampling each balloon equally but in no particular order. The fluctuating colored lines show the estimated probability that each balloon is best at each trial (blue: fast balloon; red: medium balloon; yellow: slow balloon). The vertical lines show trials where either all balloons were reset (pink) or revealed (black). The color of the crosses indicates which balloon was sampled on each trial, while the color of the circles indicates which balloon was chosen.

To capture the sequential sampling strategy we set β to 1 and γ to 3, while setting the other free parameters to 0, so:

$$V_{check,balloon} = n_{balloon}(t) + 3 \times Unknown_{balloon}$$

Samples simulated using this value of checking indeed have a sequential pattern (Figure S4.4), with the pattern being reset whenever there is a reset trial. Furthermore, it can be seen that the value of β affects the amount of 'greedy' sampling in simulated data (Figure S4.5).

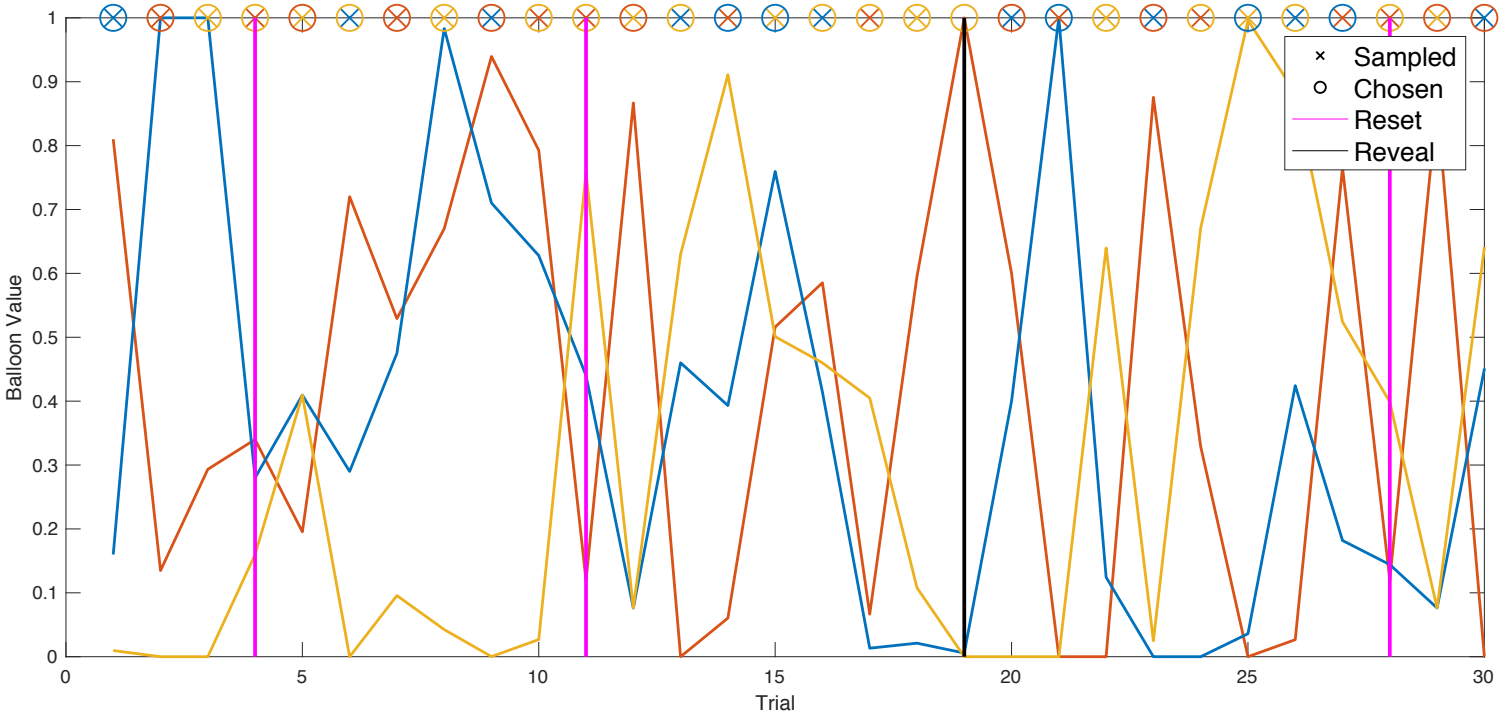


Figure S4.4 Simulated data from model with only free parameters β and γ. The model is sampling the balloons in a fixed sequence which is reset whenever there is a reset trial. The fluctuating colored lines show the estimated probability that each balloon is best at each trial (blue: fast balloon; red: medium balloon; yellow: slow balloon). The vertical lines show trials where either all balloons were reset (pink) or revealed (black). The color of the crosses indicates which balloon was sampled on each trial, while the color of the circles indicates which balloon was chosen.

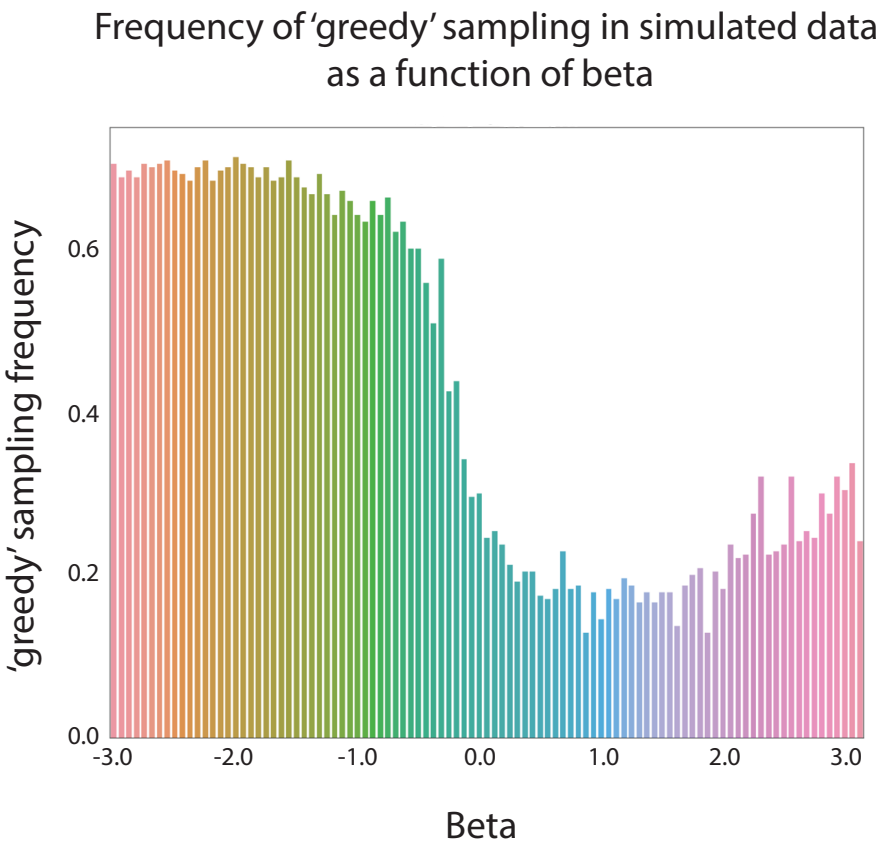


Figure S4.5 Proportion of 'greedy' samples as a function of value of β in simulated data. It can be seen that lower values of β lead to more 'greedy' samples, and that there is a small range of β centered around zero where 'greedy' sampling changes greatly as a function of β.

To get the model to sample the best balloon, we set δ to 1 and all other free parameters to 0:

$$V_{check,balloon} = P(EV_{balloon} = \max(EV), t)$$

Using this value of checking leads the model to almost always sample the balloon it currently estimates to have the highest probability of having the highest value of the three balloons (Figure S4.6). Furthermore, data simulation using many possible values of δ shows that higher δ leads to more sampling of the balloon with the highest value (Figure S4.7).

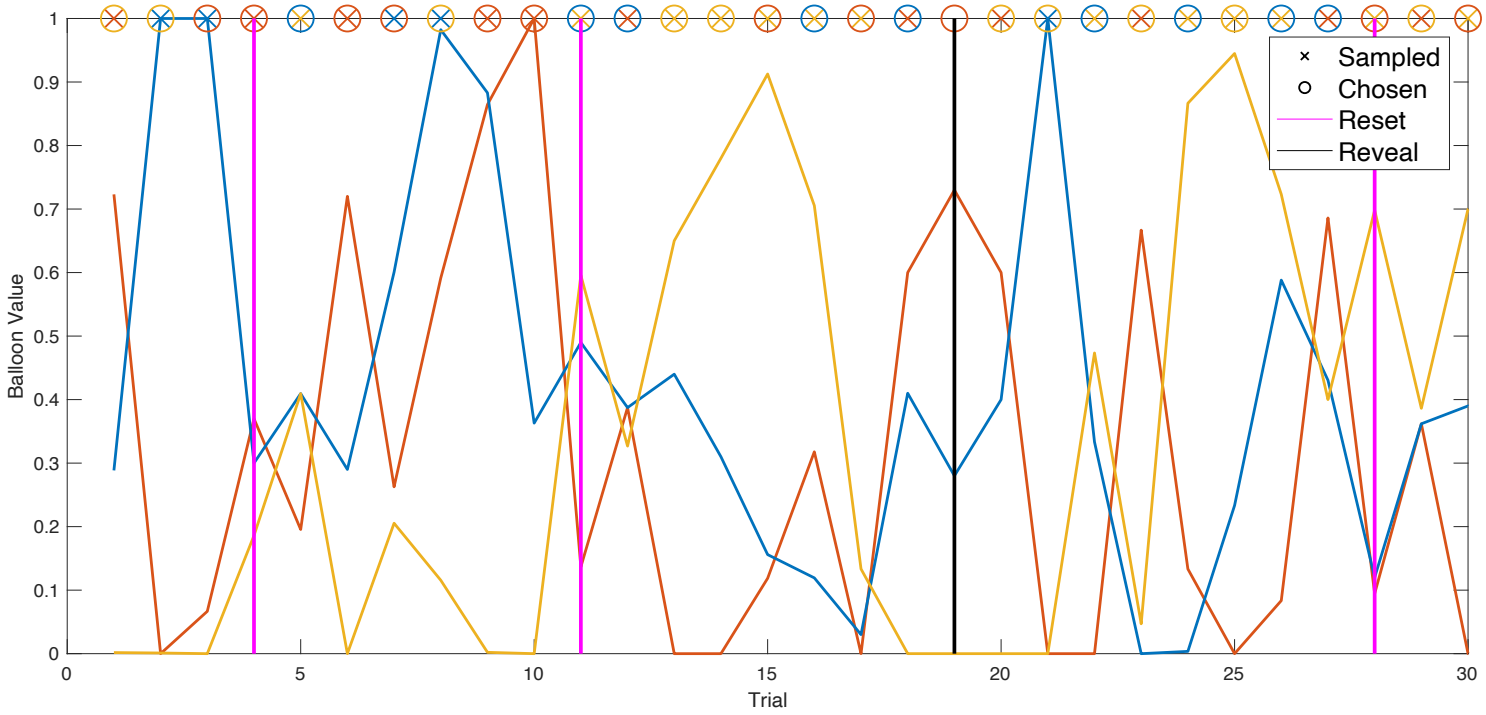


Figure S4.6 Simulated data from model with only free parameter δ. The model is almost always sampling the balloon with the current highest estimated probability of being the best balloon. The fluctuating colored lines show the estimated probability that each balloon is best at each trial (blue: fast balloon; red: medium balloon; yellow: slow balloon). The vertical lines show trials where either all balloons were reset (pink) or revealed (black). The color of the crosses indicates which balloon was sampled on each trial, while the color of the circles indicates which balloon was chosen.

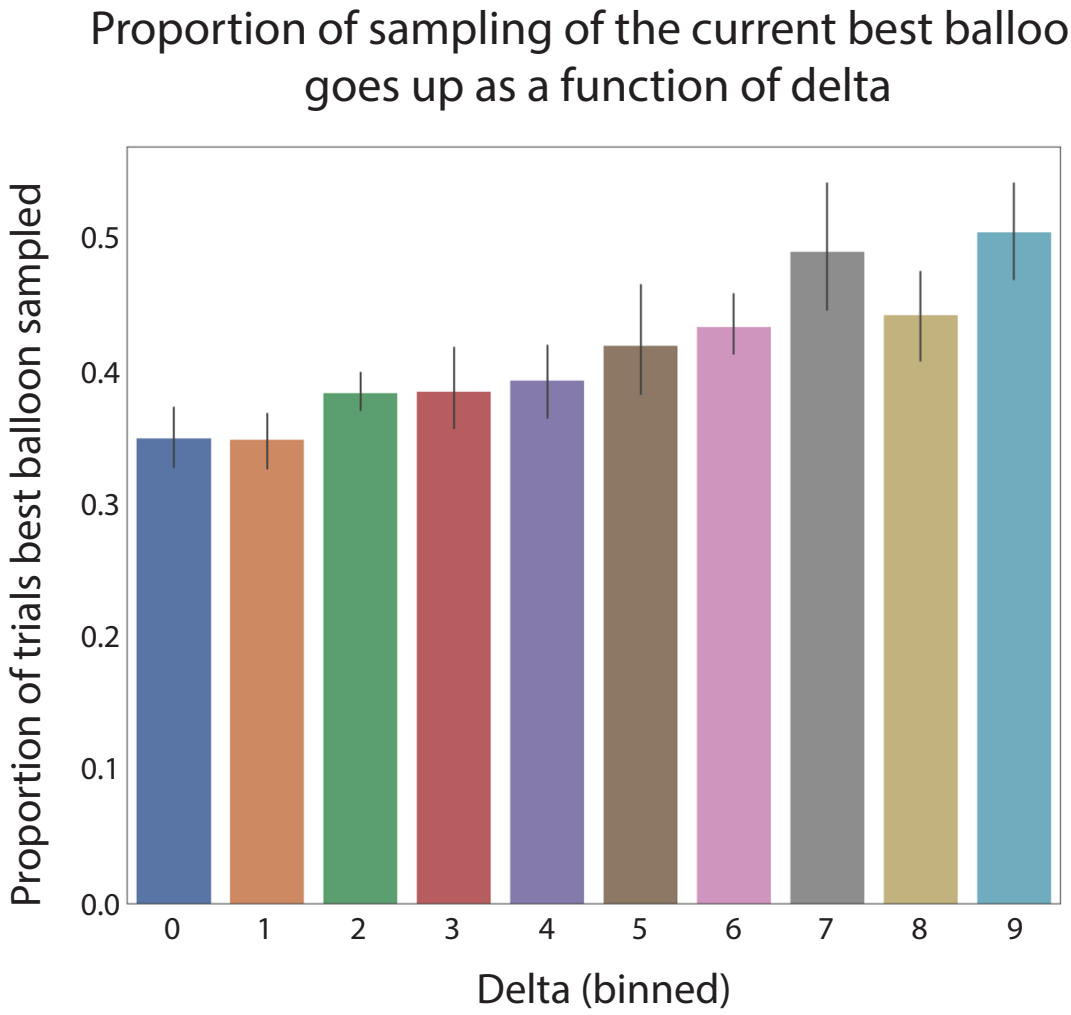


Figure S4.7 Proportion of trials where the balloon with the highest true value was sampled as a function of value of δ in simulated data. It can be seen that higher values of δ lead to more sampling of the best balloon.

Finally, the 'sampling the chosen balloon' strategy is estimated by setting free parameter ζ to 1 and the others to 0:

$$V_{check,balloon} = Chosen_{balloon}(t-1)$$

This version of the model indeed produces mostly samples of the balloon chosen on the previous trial (Figures S4.8-9).

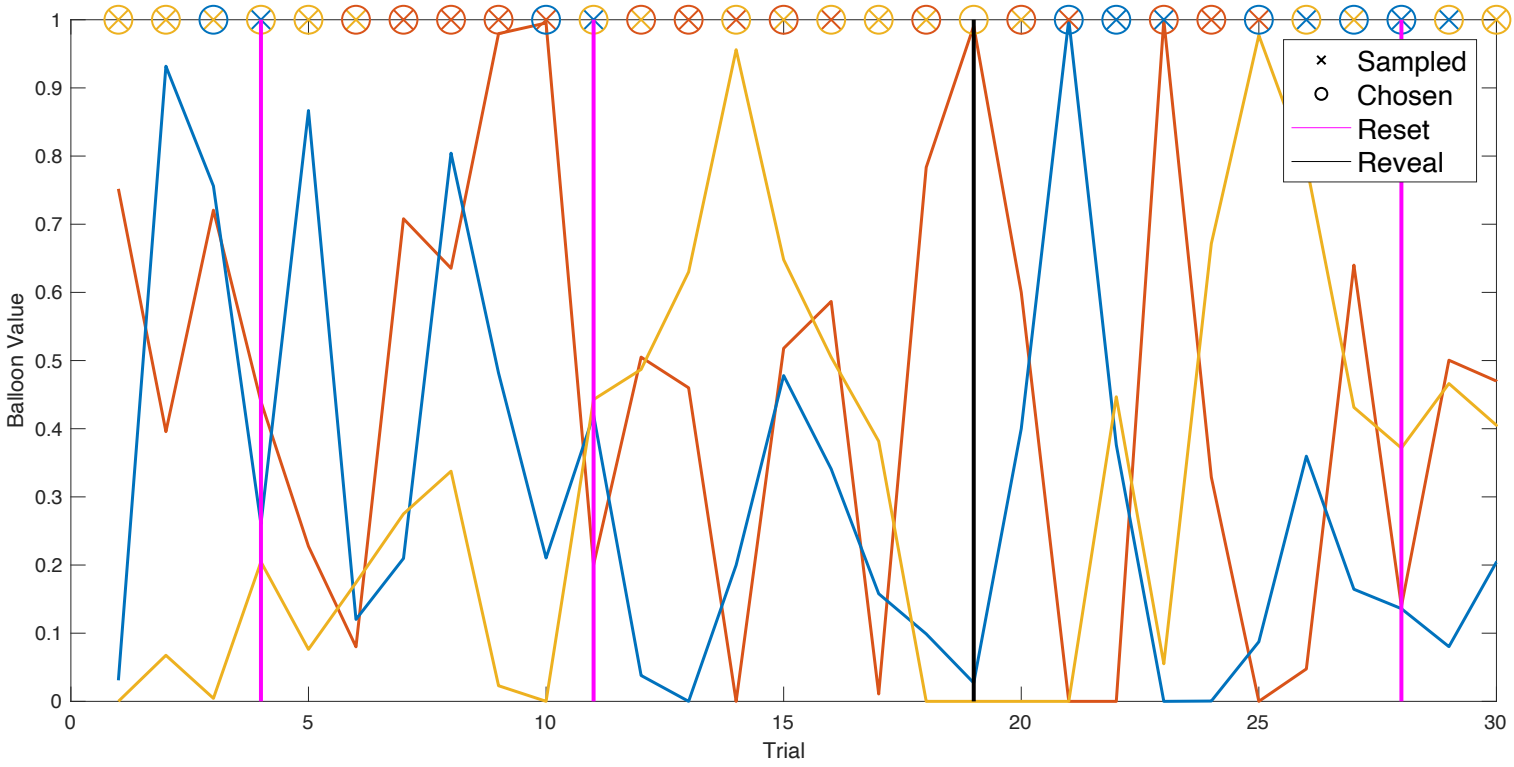


Figure S4.8 Simulated data from model with only free parameter ζ. The model is almost always sampling the balloon the model chose on the previous trial. The fluctuating colored lines show the estimated probability that each balloon is best at each trial (blue: fast balloon; red: medium balloon; yellow: slow balloon). The vertical lines show trials where either all balloons were reset (pink) or revealed (black). The color of the crosses indicates which balloon was sampled on each trial, while the color of the circles indicates which balloon was chosen.

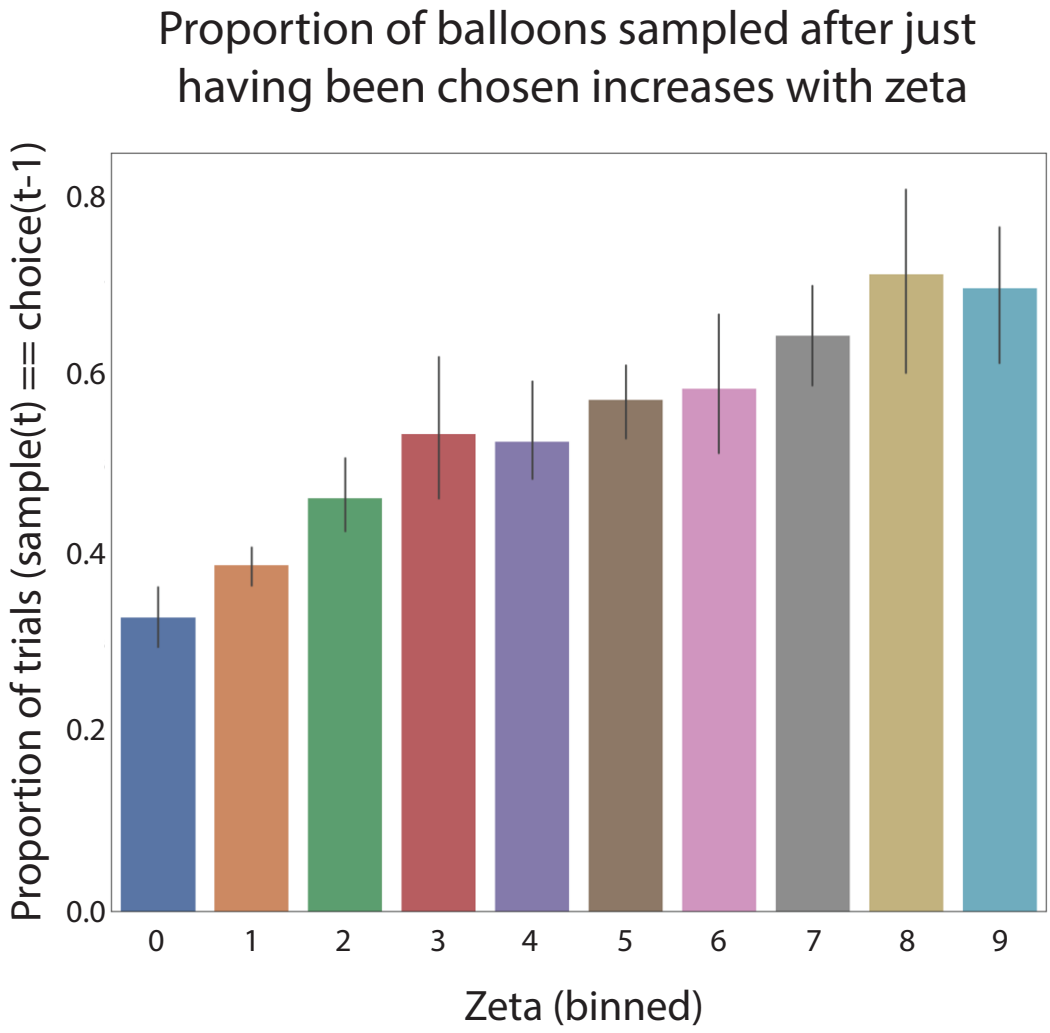


Figure S4.9 Proportion of trials where the balloon chosen on the previous trial was sampled as a function of value of ζ in simulated data. It can be seen that higher values of ζ lead to more sampling of the previously chosen balloon.

In summary, it looks like the task parameters included in the function of value of checking seem to accurately capture our hypothesized strategies.

We fit multiple versions of this model to the online data collected on task version 3 (see section 4.2.4). Specifically, we fit the model including the five free parameters described above along with inverse temperature and random noise as free parameters.

We also fit four versions of the model where one of β, γ, δ, or ζ are excluded. The full model had the lowest BIC score of all versions (Figure S4.10), but doing a Bayesian model comparison showed that there was individual variation in which model fit each participant best, with some participants' behavior being better fit by the models excluding δ or ζ (Figures S4.11-12). Besides this, we were able to recover the parameters from simulated data reasonably well, though this recovery was worse for higher parameter values (Figures S4.13).

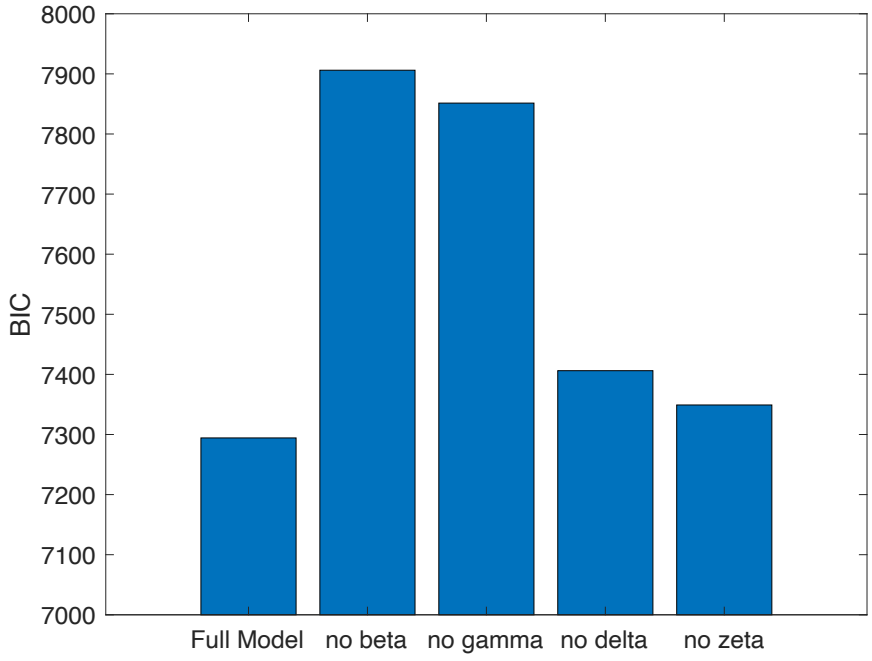


Figure S4.10 A model with all five free parameters fit the data better than the four versions of the model each excluding one of the parameters as indicated by its lower BIC score.

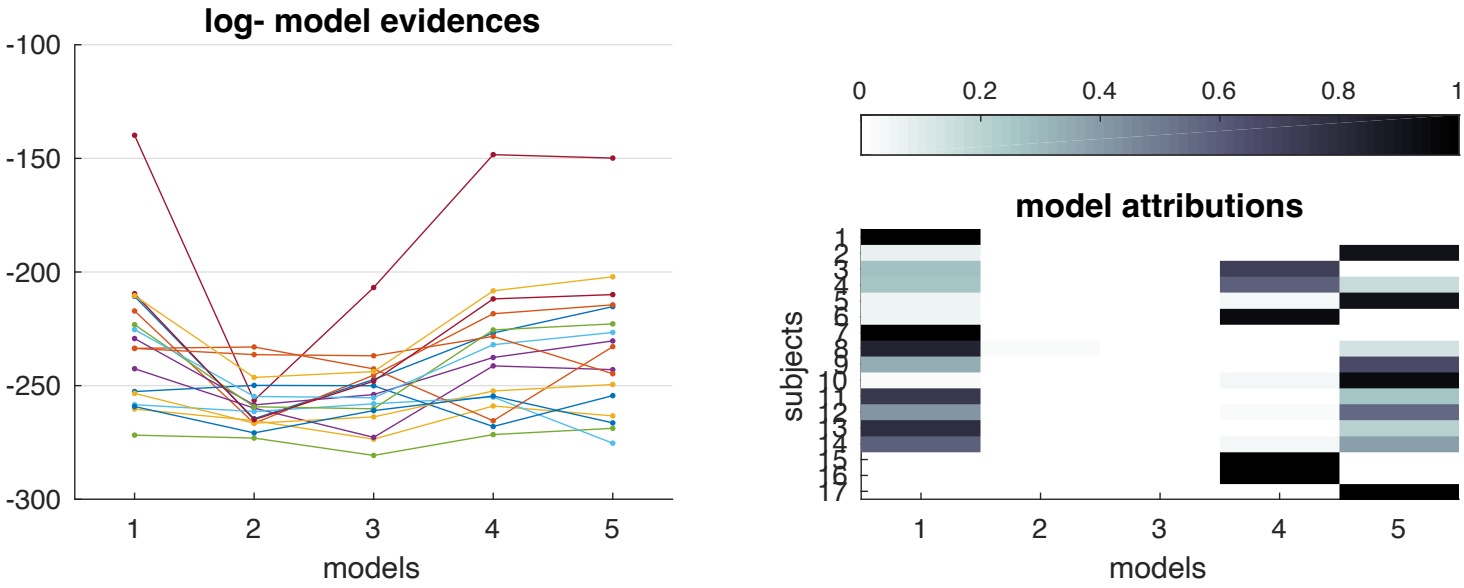


Figure S4.11 Bayesian Model Comparison of five described models using pilot data from task version 3. It can be seen that while overall the full model seems to fit the data best, there is quite a bit of individual variation, with the models without δ or ζ also explaining some participants' behavior better than those without β or γ. (A) Log-evidence over each model for each participant (each colored line represents one participant). (B) Posterior probability that a model best explains a participant for each of the 5 models and each of the participants.

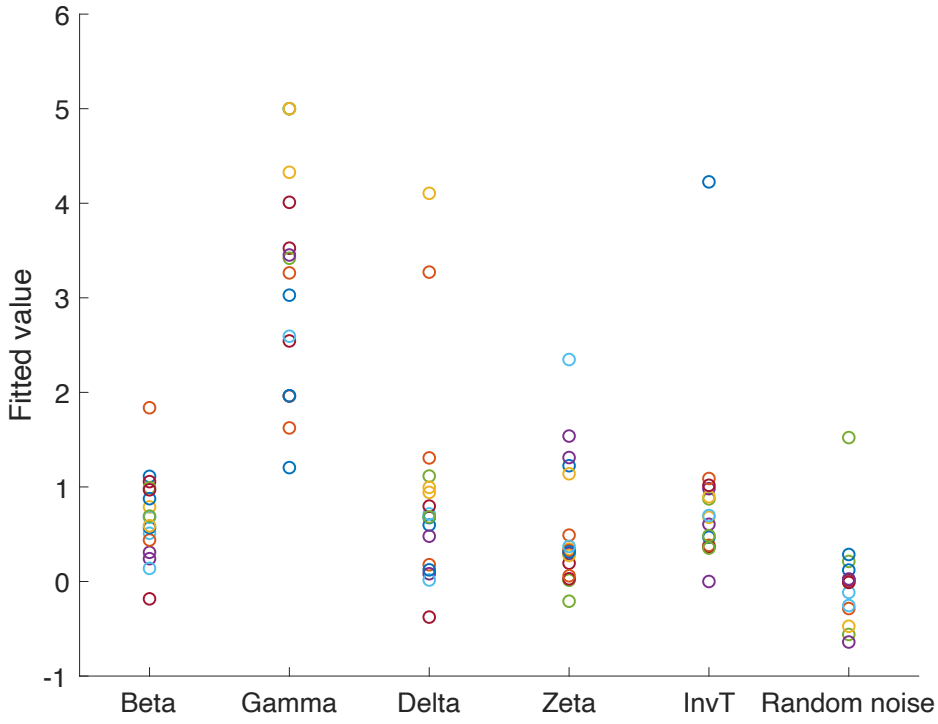


Figure S4.12 Fitted parameters for each of the participants using the full version of the model. Data collected using task version 3.

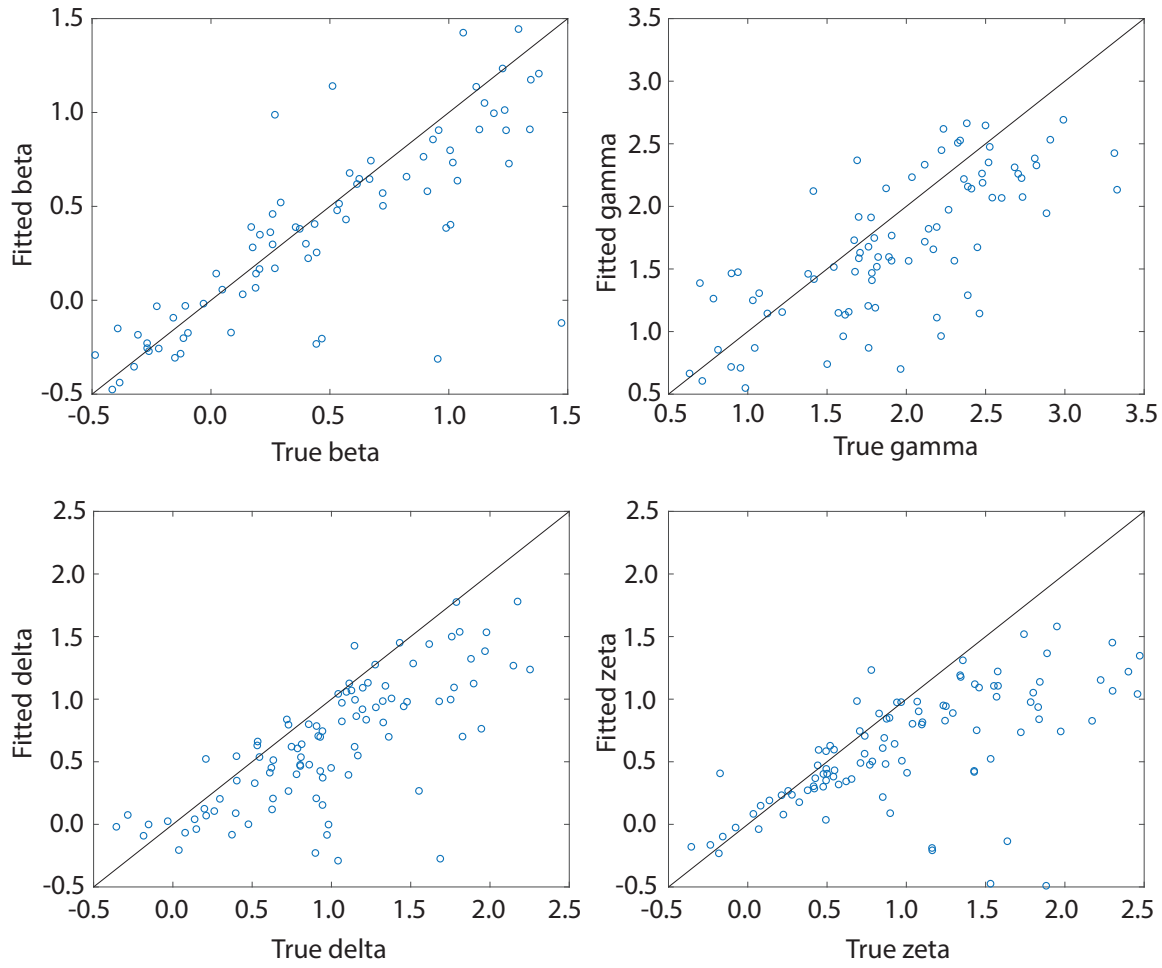


Figure S4.13 Parameter recovery of β, γ, δ, and ζ using the full model. Plotted are the true values of the parameters used to simulate data on the x-axis and the estimated values of the parameters by the model on the y-axis with each data point being a simulated 'participant'.

We made a final adaptation to this model by including an optimal driver of information sampling, that computes the value of checking as described in section 4.3.2. We added this value of checking weighted by a free parameter η to the model as follows:

$$V_{check,balloon}(t) = \alpha + \beta \times n_{balloon}(t) + \gamma \times Unknown_{balloon} + \delta \times Prob(EV_{balloon} = \max(EV), t)$$

$$+ \zeta \times Chosen_{balloon}(t-1) + \eta \times V_{opt,balloon}(t)$$

With:

$$V_{opt,balloon} = \sum_{n=1}^{EV_{secondbest}} (P(V_{balloon} = n) \times (EV_{secondbest} - n))$$

if the balloon in question is the currently best estimated balloon, and:

$$V_{opt,balloon} = \sum_{n=EV_{best}}^{100} (P(V_{balloon} = n) \times (n - EV_{best}))$$

if the balloon in question is not the currently best estimated balloon. (Both derived in section 4.6.1).

When fitting this model to our pilot data collected using task version 3, it can be seen that estimates for η are fairly low for all the participants compared to other free parameters, suggesting that participants probably do not sample like an optimal agent on this task (Figure S4.14). Furthermore, the best fitting models to the data were found to be those excluding the 'sampling the best balloon' or 'sampling the chosen balloon' strategies or excluding the optimal driver (Figure S4.15). Fitting the models to data simulated from each of the six models showed that each of the models could be recovered in most cases, except the model excluding β (Figure S4.16).

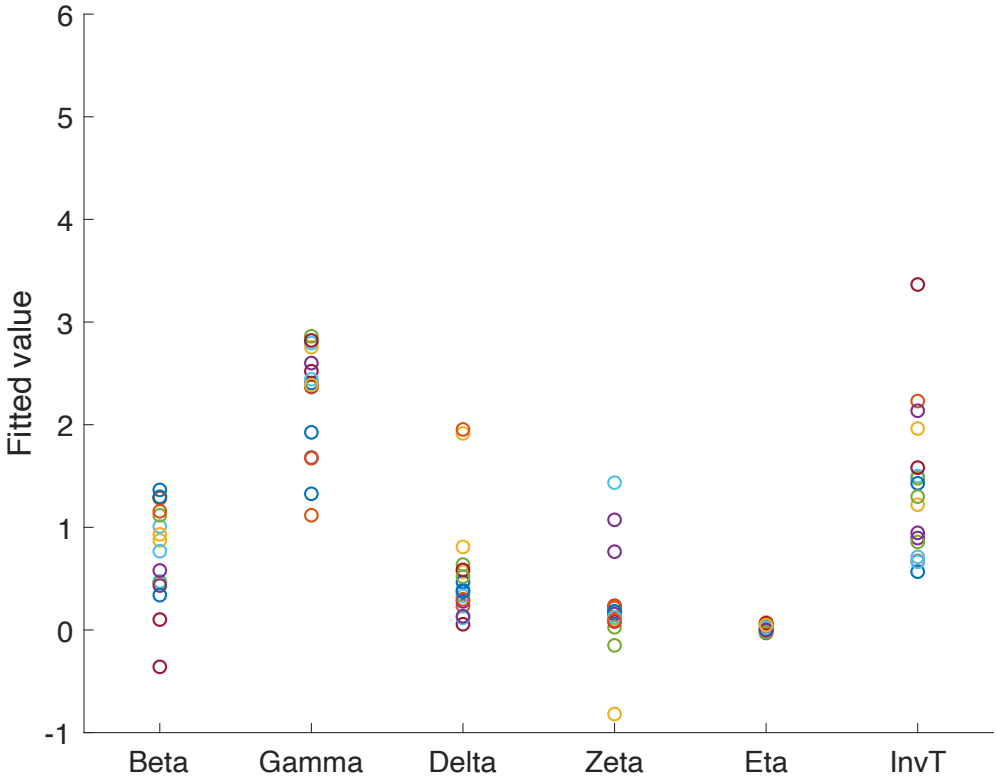


Figure S4.14 Fitted parameters for each of the participants using the version of the model including an optimal sampling driver weighted by η. It can be seen that the estimated values of η are low for each participant compared to the other weights, suggesting participants may not sample in an optimal fashion. Data collected using task version 3.

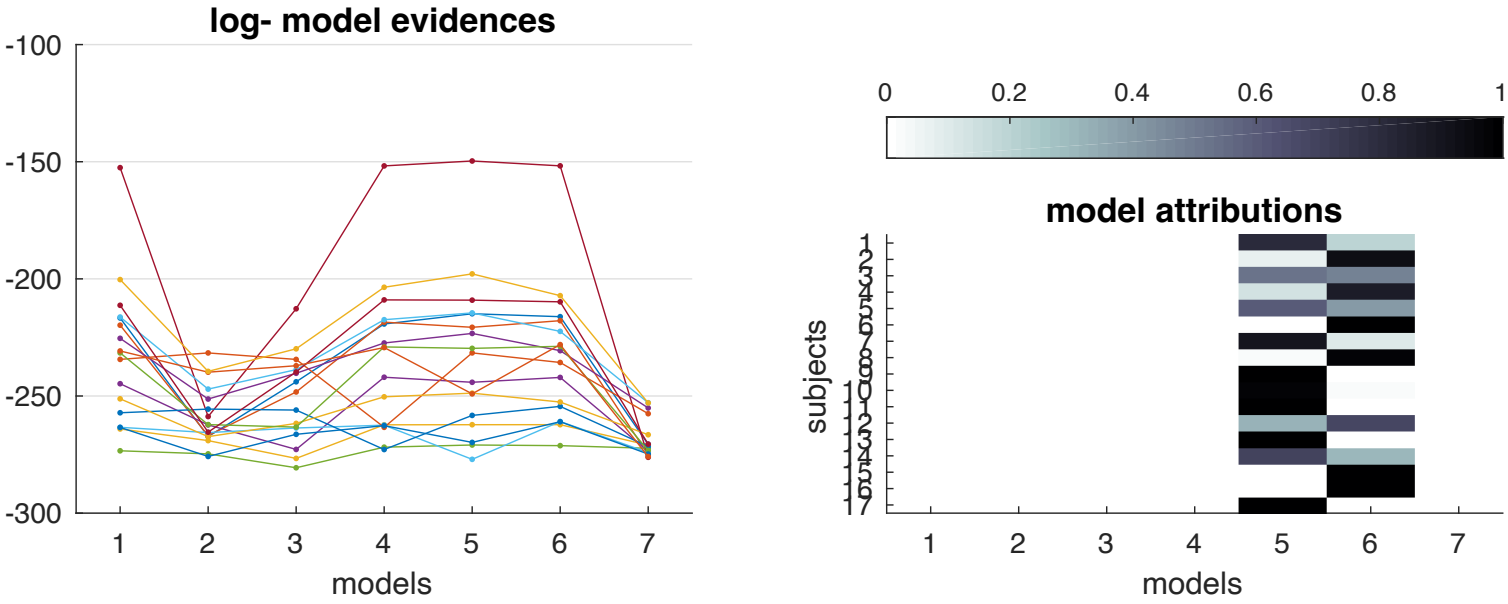


Figure S4.15 Bayesian model comparison of six models described above as well as a model consisting only of the optimal sampling driver using pilot data from task version 3. It can be seen that the models excluding the 'sampling the chosen balloon' strategy or excluding the optimal driver fit the data best. (A) Log-evidence over each model for each participant (each colored line represents one participant). (B) Posterior probability that a model best explains a participant for each of the seven models and each of the participants.

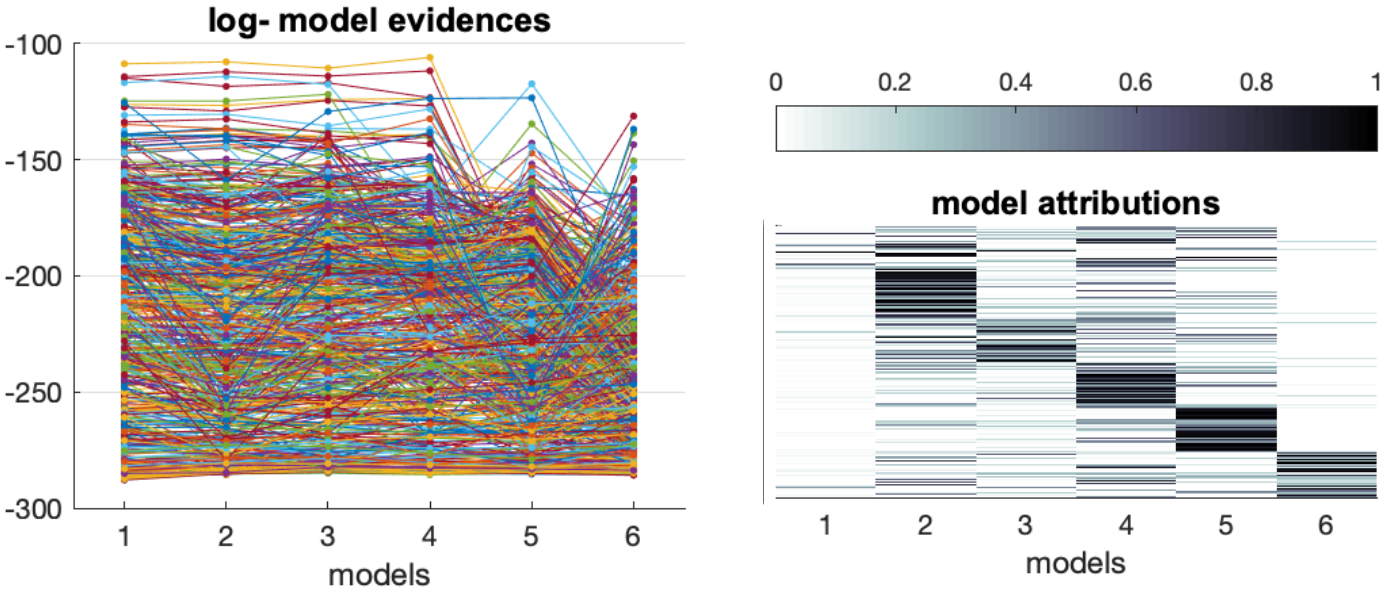


Figure S4.16 Model recovery of the six models described above. Data was simulated using each of the six models. The six models were then fit to each set of resulting data and a Bayesian model comparison was run. It can be seen that the full model is not recovered as well as the other models. (A) Log-evidence over each model for each set of simulated data (each colored line represents one simulated dataset). (B) Posterior probability that a model best explains a dataset for each of the six models and each of the participants. The datasets are ordered by model.

We fit the pilot data from task version 3 using the model with five free parameters (so excluding free parameter η) to investigate how many trials could be accurately predicted using this model. The amount of explained trials varied greatly from participant to participant, suggesting that this model is not sufficiently capturing all possible strategies participants are using to perform this task (Figure S4.17-19). Specifically, the wide variation in 'greedy' sampling observed in the pilot data, does not seem to be explained by participants using the 'sampling the best balloon' strategy to different extents. Behavior of participants who perform more 'greedy' samples is less well explained by the model (Figure S4.18) and there is no relationship between estimates of δ and 'greedy' sampling (Figure S4.20). β estimates, on the other hand, are strongly related with greedy sampling such that the higher β, the fewer greedy samples participants perform (S4.21). So it seems that perhaps participants' behavior may be best explained in terms of more or less sequential sampling. However, this still does not explain what feature of a balloon leads to repetitive greedy sampling in those participants that frequently engage in this.

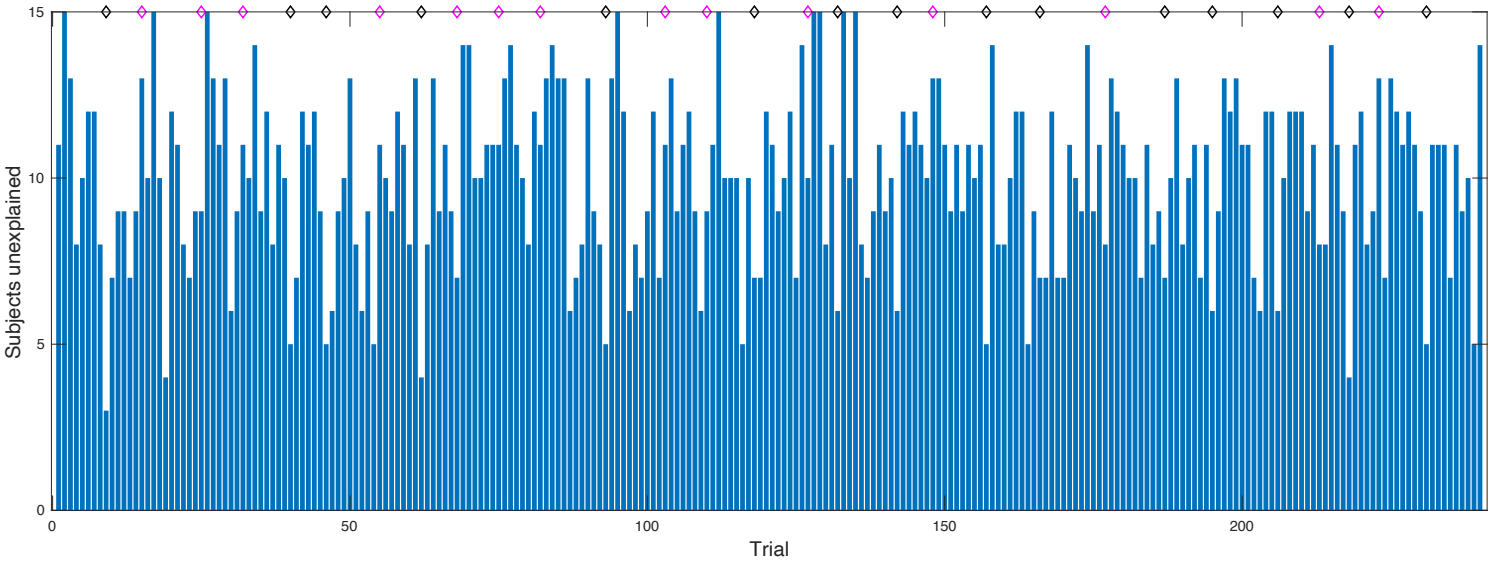


Figure S4.17 Number of participants' samples explained on each trial of the task by the model including five free parameters.

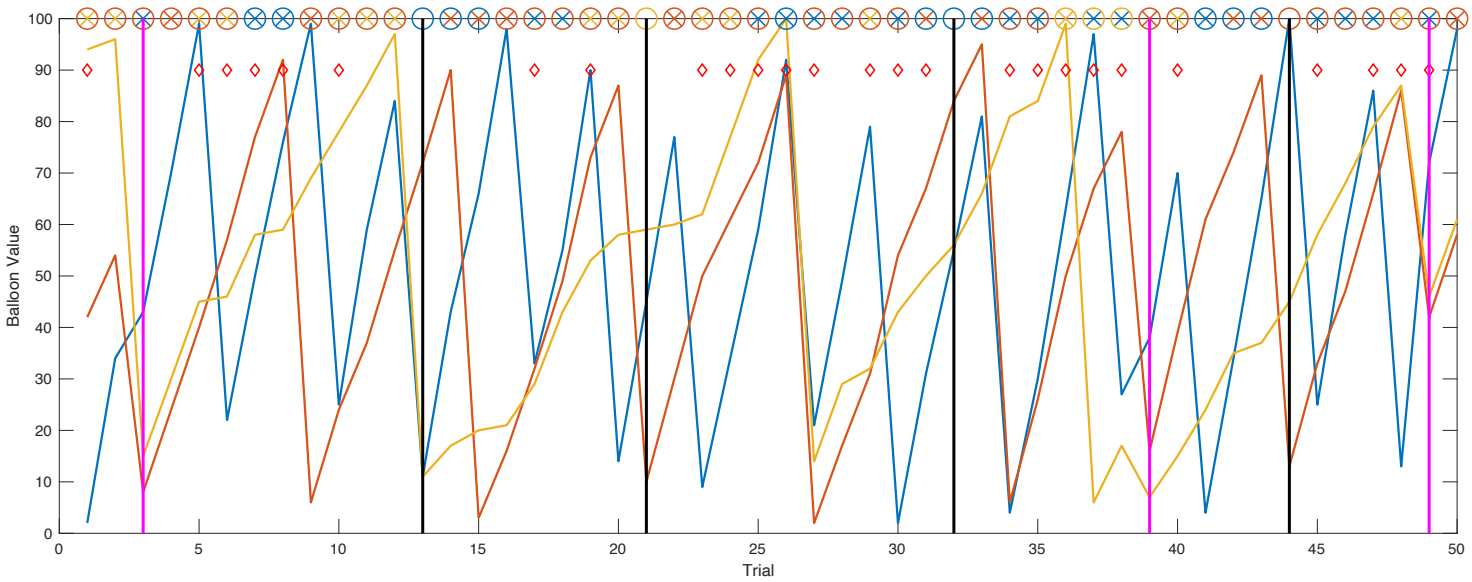


Figure S4.18 Sample and choice data from participant 25 from task version 3. It can be seen that the model was not able to predict this participant's behavior on many trials (indicated by red diamonds). This is also a participant who performed many 'greedy' samples, which apparently is not be explained by any of our hypothesized strategies included in the model. The fluctuating colored lines show the true value of each balloon at each trial (blue is fast balloon, red is medium balloon, yellow is slow balloon). The vertical lines show trials where either all balloons were reset (pink) or revealed (black). The color of the crosses indicates which balloon was sampled on each trial, while the color of the circles indicates which balloon was chosen.

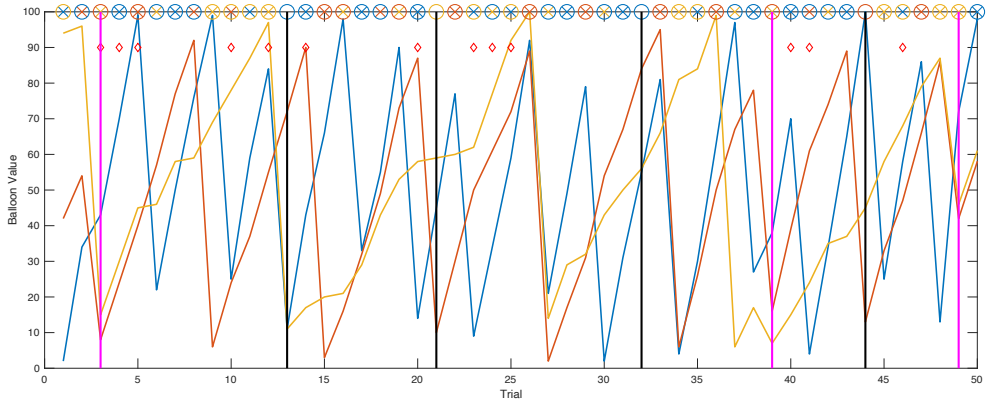


Figure S4.19 Sample and choice data from participant 13 from task version 3. It can be seen that the model was able to predict this participant's behavior on most trials (unpredicted trials indicated by red diamonds). The fluctuating colored lines show the true value of each balloon at each trial (blue is fast balloon, red is medium balloon, yellow is slow balloon). The vertical lines show trials where either all balloons were reset (pink) or revealed (black). The color of the crosses indicates which balloon was sampled on each trial, while the color of the circles indicates which balloon was chosen.

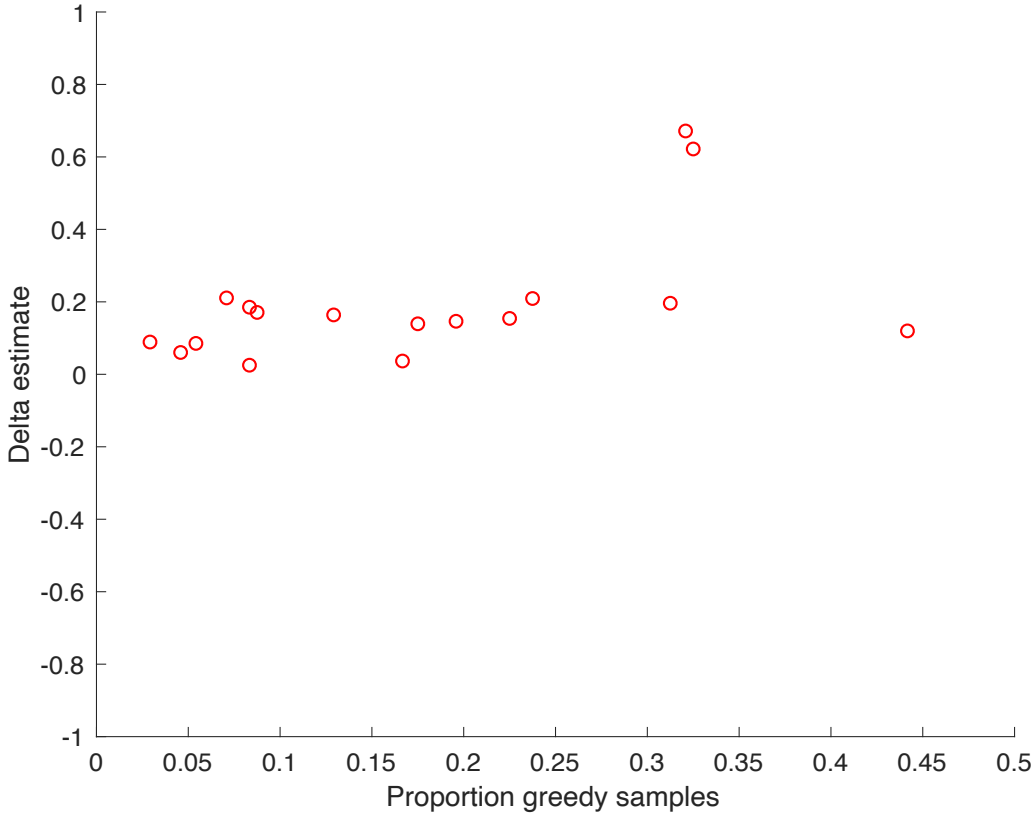


Figure S4.20 Estimated δ is not correlated with proportion of greedy samples in pilot data from task version 3. Each data point is a participant.

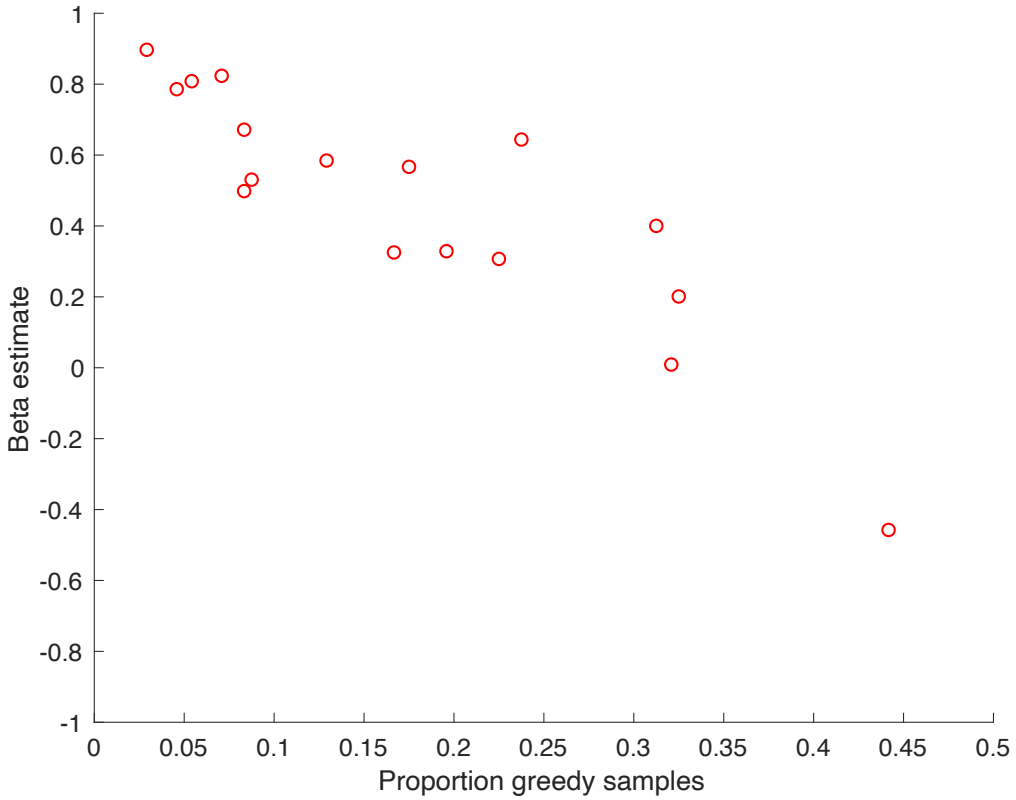


Figure S4.21 Estimated β is negatively correlated with proportion of greedy samples in pilot data from task version 3. Each data point is a participant.

Chapter 5: Compulsivity and information sampling

Abstract

Background: Compulsivity has both been linked to more 'habitual' decisions and increased information sampling. However, it is not clear to what extent increased sampling is another example of repetitive and inflexible behavior considered typical in OCD, or whether it is the result of altered evidence accumulation parameters that can flexibly respond to environmental demands.

Methods: We investigated the effect of compulsivity on information sampling in a novel information gathering task in a large online sample (discovery sample: n = 150, confirmation sample: n = 300), from which we collected responses to a set of psychiatric questionnaires to determine scores on three transdiagnostic dimensions, including compulsivity.

Results: We did not find a relationship between compulsivity scores and the amount of information sampled on the task. In the discovery sample, a modest effect was found of compulsivity on the interaction between value volatility and time passed since a previous sample on information sampling. This effect was not replicated in the confirmation sample.

Conclusion: Our task was potentially ill-suited to find an effect of compulsivity on information sampling, although it did contain substantial variability in information sampling across individuals with high reliability using split-half analysis. There have also been previous null findings in studies investigating this potential relationship. It is therefore unclear to what extent altered information sampling can be expected in large subclinical populations.

5.1 Introduction

The ability to make good decisions relies on the quality and quantity of the information used to guide that choice. Almost every psychiatric disorder has been associated with altered decision-making in some way (APA, 2013; Der-Avakian & Markou, 2012; Everitt & Robbins, 2005; Fine et al., 2007; Gillan et al., 2016; Moutoussis et al., 2011; Robbins et al., 2012; Sachdev & Malhi, 2005), but much less is known on how the information sampling process that precedes and drives choice might be affected by these disorders, and whether that may in fact explain part of the observed changes in decision-making.

So far, obsessive-compulsive disorder (OCD) is one of the disorders that has been most frequently linked to altered information sampling. Simple perceptual binary forced choice tasks have mostly been used to study this phenomenon, particularly the random dot motion task, using the drift diffusion model (DDM) to model how participants accumulate evidence leading up to a choice (Banca et al., 2015; Erhan et al., 2017). Using this approach, it has been repeatedly shown that OCD patients show increased decision thresholds, meaning they need to sample more evidence before being able to commit to a choice (Banca et al., 2015; Erhan et al., 2017). It has been argued that this is what produces the repetitive 'checking' behaviors and indecisiveness commonly found in OCD (Erhan et al., 2017; Sachdev & Malhi, 2005).

A downside of using this approach to study information sampling is that the degree of information sampling has to be inferred from the DDM. In most choice tasks used in these studies participants do not actively choose to sample information, but instead are passively presented with a continuous stream of evidence. The DDM is then fit to their data with the assumption that longer gazing times are synonymous with

more information sampling, or evidence accumulation, but the effect of compulsivity on information sampling is not directly tested. A few active information gathering tasks have been used to study OCD patients, with varying results: in some studies OCD patients showed increased information sampling (Fear & Healy, 1997; Milner et al., 1971; Pélissier & O'Connor, 2002), while in others there was no difference with healthy controls (Chamberlain et al., 2007; Jacobsen et al., 2012). Recently, Hauser et al. (2017) combined the two approaches outlined above by showing that compulsive individuals sampled more in an active information gathering task, and that this was due to a raised decision threshold, or: a reduced urgency signal, meaning they required more evidence to be able to commit to a choice, as shown before in perceptual binary forced choice tasks with passive information sampling.

These findings somewhat contrast with older accounts of compulsivity, which find that compulsive individuals tend to exhibit more habitual actions (Gillan et al., 2011). These actions are not guided by an internal model of the task and thus do not flexibly adapt to new evidence. We might expect compulsive individuals to show similar habitual behavior in information sampling tasks, such as sampling information irrespective of its informativeness or instrumentality. Hauser et al.'s (2017) findings, however, provide a slightly different account, suggesting that repetitive sampling behavior in compulsivity is in fact driven by the same decision-making process as in healthy controls, just with altered parameters (a higher decision threshold). Furthermore, compulsive individuals even performed better at the decision-making task used in this study than healthy controls, due to their increased sampling, challenging the notion that compulsivity exclusively drives suboptimal behavior (APA, 2013; Tolin et al., 2003).

In this chapter, I describe findings from a large-scale online study using a novel information gathering task combined with psychiatric questionnaires to further investigate the effect of compulsivity on information sampling and its underpinnings. In our task, participants are expected to use multiple task features to decide how much information to sample to guide upcoming reward-maximizing choices. The habitual account of compulsivity would suggest that high compulsivity leads to more repetitive sampling behavior that does not flexibly adapt to the specific need for information on a particular trial. By contrast, Hauser et al.'s (2017) findings would also predict that highly compulsive individuals sample more, but that this arises from a similar decision-making process as that in less compulsive individuals. Therefore, we would expect participants' information sampling to be affected by fluctuating demands for information throughout the task regardless of their compulsivity level.

5.2 Methods

5.2.1 Participants

We collected two separate samples, a discovery sample and confirmation sample. The discovery sample consisted of 150 participants, and the confirmation sample of 300 participants. These were recruited online via Prolific. Participants were only recruited to the study if they had a 95% prior approval rate on Prolific, had at least 10 previous Prolific study submissions, had British nationality and reported to be fluent in English.

5.2.2 Task

In a sequential sampling task, participants repeatedly made information sampling and reward-maximizing choices with the aim of earning as many points as possible (Figure 5.1). The available choice options were 3 balloons, each containing a value between 1-100 points at any given trial. Each of the balloons increased in value on each trial until their value reached more than 100 points, at which point the balloon would 'pop' and start increasing in value again from 1 point. Crucially, these values were hidden from the participant's view unless they chose to sample them during a sampling trial.

Sampling decisions to gather information were completely separated from choice decisions to maximize reward by alternating sampling rounds and choice rounds. Each trial consisted of one sampling round and one choice round. On sampling rounds, participants were allowed to sample as many of the three balloons as they liked (or none) through keypresses on their keyboard. When a balloon was sampled, participants had to wait for 3s, during which they could not respond, and then a value bar next to that balloon would update to reflect that balloon's current value. This value bar would stay at that value until the balloon was sampled again. In other words, the value bar did not reflect the balloon's true current value if that balloon was not sampled on that round. Participants were reminded of how many trials had passed since they last sampled a balloon by the number of white dots presented next to the value bar, to prevent working memory from confounding results in this study. When the participant felt they sampled enough, they could move on to the choice round using a keypress. If the participant sampled all the balloons on a sampling round, they were automatically moved on to the choice round.

Each sampling round was followed by a choice round. Here, participants were asked to choose the balloon they thought was best at that time by pressing the corresponding key on their keyboard. Participants received the number of points corresponding to their chosen balloon on that round, but did not receive feedback on the number of points they had received after choosing. In other words, participants could only receive information on the values of the balloons in the sampling rounds. On every five trials, though, participants did receive feedback on their performance by a message indicating what percentage of total possible points that could have been acquired in the previous five trials they won. After each choice round, all three balloons would increase in value, or 'pop' if their value went over 100 points, and a new sampling round started.

Additionally, the three balloons varied in how quickly they increased in value. Specifically, there was one 'fast, one 'medium' and one 'slow' balloon. The fast balloon increased at a rate of 21-35 points/trial, medium balloons at 11-25 points/trial and slow balloons at 1-15 points/trial. The increase on each trial was sampled from a uniform distribution with the corresponding range. The order of the balloons on the screen was 'fast', 'medium' 'slow' for half of the participants, and the reverse order was used for the other half to avoid ordering effects confounding our results. Furthermore, the trials were grouped into blocks of between 5-11 trials. At the start of a block, participants were presented with 'new' balloons, with starting values randomly sampled from a uniform distribution between 1-100 points. The start of a new block was clearly indicated on the screen. On half of the blocks, all the values of the new balloons were hidden to the participant ('reset' trials), and on the other half they were all revealed to the participant ('reveal' trials). In the latter case, participants were immediately taken

to a choice round, as there was no new information to sample. Participants received a bonus payment of £1 for every 3000 points earned on the task in addition to a participation fee of £7.50/hour.

The full task consisted of 150 trials. However, experiment time was restricted to 30min. This meant that if participants sampled often, the 3s waiting time after each sample could cause them to not reach the end of the experiment. As a result, their pay-off could be lower due to missed choice rounds. On the other hand, if participants would sample too little, it was more difficult for them to choose the best balloon in the choice trials, also leading to a lower pay-off as well. In other words, participants were asked to balance speed and accuracy on this task. See Supplemental Material S1 for an instructional video on the task.

5.2.3. Questionnaires

After completion of the main task, participants were directed to a questionnaire on Qualtrics consisting of 63 questions used in Wise & Dolan (2020). They devised this reduced set of questions to determine scores of the three transdiagnostic factors outlined in Gillan et al. (2016): compulsivity, anxiety/depression and social withdrawal. They showed that this reduced set of questions could accurately predict the true factor scores on the full set of questions. The questions originate from the following questionnaires: the Obsessive-Compulsive Inventory – Revised (OCI-R; Foa et al., 2002) the Self-Rating Depression Scale (Zung, 1965), the State-Trait Anxiety Inventory (Spielberger et al., 1983), the Alcohol Use Disorder Identification Test (AUDIT; Saunders et al., 1993), the Apathy Evaluation Scale (AES; Marin et al., 1991), the Eating Attitudes Test (EAT-26; Garner et al., 1982), the Barratt Impulsivity Scale

(BIS-10; Patton et al., 1995) and the Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987). As data was collected during the COVID-19 pandemic, participants were instructed to answer the questions as they would have prior to the pandemic. This was particularly relevant for questions about social interactions and activities, which would have been reduced for most participants due to the pandemic. See Supplementary Material S1 (section 5.5.1) for the set of questions used.

5.2.4 Participant exclusion

At the end of the main task, participants were asked two freeform questions: 'What strategy did you use during sampling rounds?' and 'What strategy did you use during choice rounds?' Based on participants' answers to these questions, it became apparent some participants had not understood the task properly (for example, some participants seemed to believe sampling balloons caused them to pop). 18 participants from the discovery sample and 36 participants from the confirmation sample were excluded because their answers to these questions made it clear they did not understand the task. 3 participants from the confirmation sample experienced technical issues during data collection due to which their data was not saved and could therefore not be used in the analysis. Finally, 7 participants from the confirmation sample were excluded from analysis due to not filling in the psychiatric questionnaire at the end. Overall, the data analysis was therefore done on data from 132 participants in the discovery sample and 254 participants in the confirmation sample.

5.2.5 Statistical analyses

Hierarchical regression models were conducted using the lme4 package in R (Bates et al., 2015; Gelman & Hill, 2006). All models allowed for random intercepts and slopes at the participant level. We computed degrees of freedom and p-values with the Kenward-Roger approximation, using the package pbkrtest (Halekoh & Hojsgaard, 2014). Degrees of freedom for ordinary least squares (OLS) regressions varied as a function of the number of predictors included in each model.

5.2.6 Transdiagnostic psychiatric dimension scores

We computed factor scores for the three psychiatric dimensions defined by Gillan et al. (2016; compulsivity, anxiety/depression, social withdrawal) from the participants' questionnaire responses using a classifier created by Wise & Dolan (2020).

5.2.7 Ethics

This study was approved by the Oxford University Central University Research Ethics Committee (CUREC; reference number R67699/RE002) and all participants provided informed consent prior to taking part.

5.3 Results

To investigate the effect of compulsivity on information sampling we collected a discovery sample of 150 participants and a confirmation sample of 300 participants on an economic choice task with active information sampling (Figure 5.1). In the task, participants were asked to choose between three balloons that could contain a value

of 1-100 points. On each trial, participants were first given an opportunity to sample as many of the balloons as they liked, at a cost of a 3s waiting time per sampled balloon, before choosing one of the balloons. Participants received no feedback on their individual choice, so the only information participants could gain about the balloon values was by sampling them. Crucially, the values of the balloons changed between each trial. Specifically, the three balloons increased at a fast, medium or slow rate, respectively, meaning they each increased by 21-35 points/trial, 11-25 points/trial and 1-15 points/trial (the exact value increase was sampled from uniform distributions with these ranges). When a balloon value reached more than 100 points, this balloon would 'pop' and start increasing again from 1 point. Participants would only know about this 'pop' if they sampled the balloon again after popping. In this way, balloons that were of high value on a previous trial, could be a bad option on the next. Whenever a balloon was sampled, its value was revealed on the screen, and this same value continued to be presented until the balloon was sampled again or reset at the start of a new block. Trials were grouped into block of 5-11 trials, after which all the balloons would reset to new values.

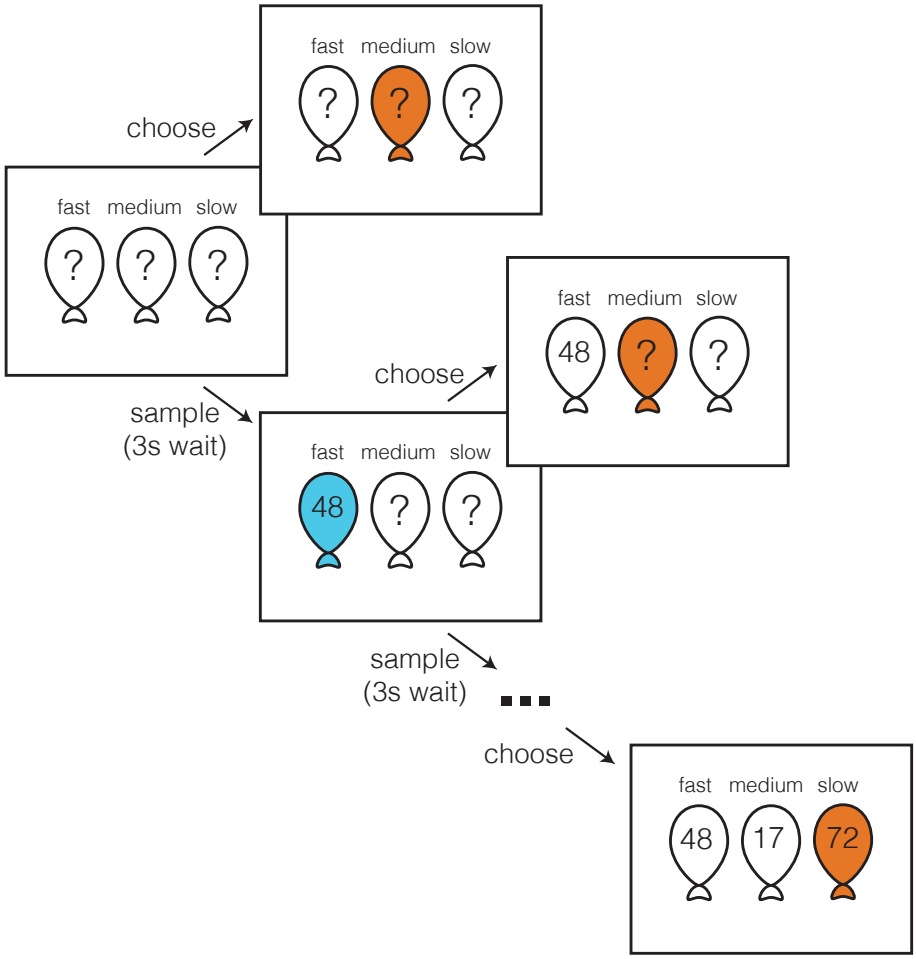


Figure 5.1. Task structure. Each trial started with the opportunity to sample as many of three balloons as the participants liked. Each balloon contained a value of 1-100 points, and this value was presented on the screen when a balloon was sampled. During this sampling phase, the participant could move on to the choice round by pressing a key on their keyboard when they were done sampling. When a participant had sampled all three balloons on a trial, they were automatically moved on to the choice round.

In addition to this, participants were asked to complete a set of 63 psychiatric questions derived from Wise & Dolan (2020), from which scores on three transdiagnostic psychiatric dimensions were computed: compulsivity, anxiety/depression and social withdrawal (Gillan et al., 2016; see supplementary materials section 5.5.1 for questions and dimension weights).

5.3.1 Discovery sample

We initially collected a sample of 150 online participants to investigate multiple potential effects of compulsivity on behavior in the task.

5.3.1.1 Task performance (discovery sample)

Participants showed high variability in the number of balloons sampled per trial (Figure 5.2A-B). The mean number of balloons sampled by participants on a randomly sampled half of trials was significantly correlated with the mean number sampled on the other trials, meaning the individual differences we observe are reliable (β = 0.94, $p < 0.001$). Besides this, the mean number of balloons participants sampled was negatively correlated with the number of trials they completed, meaning the more participants sampled, the less likely they were to complete all the trials due to the waiting cost associated with sampling (β = -0.02, $p < 0.001$).

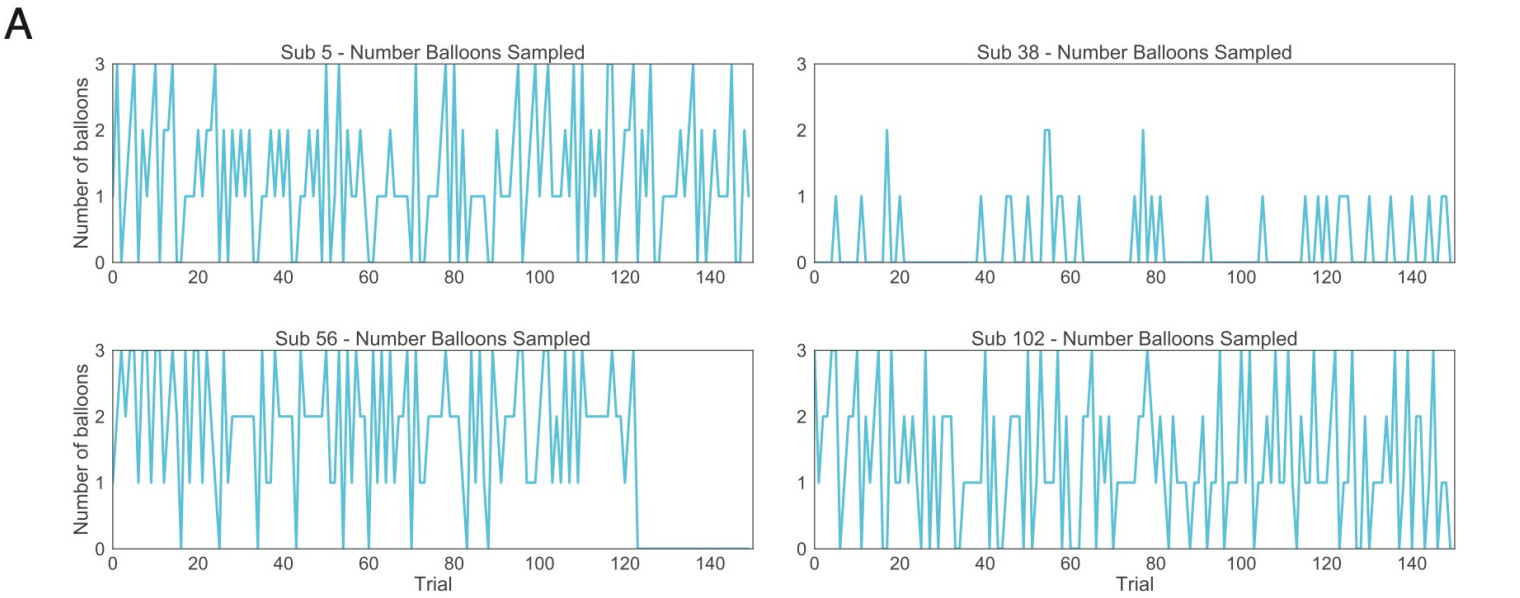
The value revealed through sampling negatively predicted further sampling on the same trial, meaning participants were more likely to skip to the choice round when they had just sampled a high-value balloon (Figure 5.2C; z = -15.17, $p < 0.001$). It can be seen in Figure 5.2C that the opposite effect is much less strong: participants are not much less likely to sample further if the balloon they just sampled was of low value.

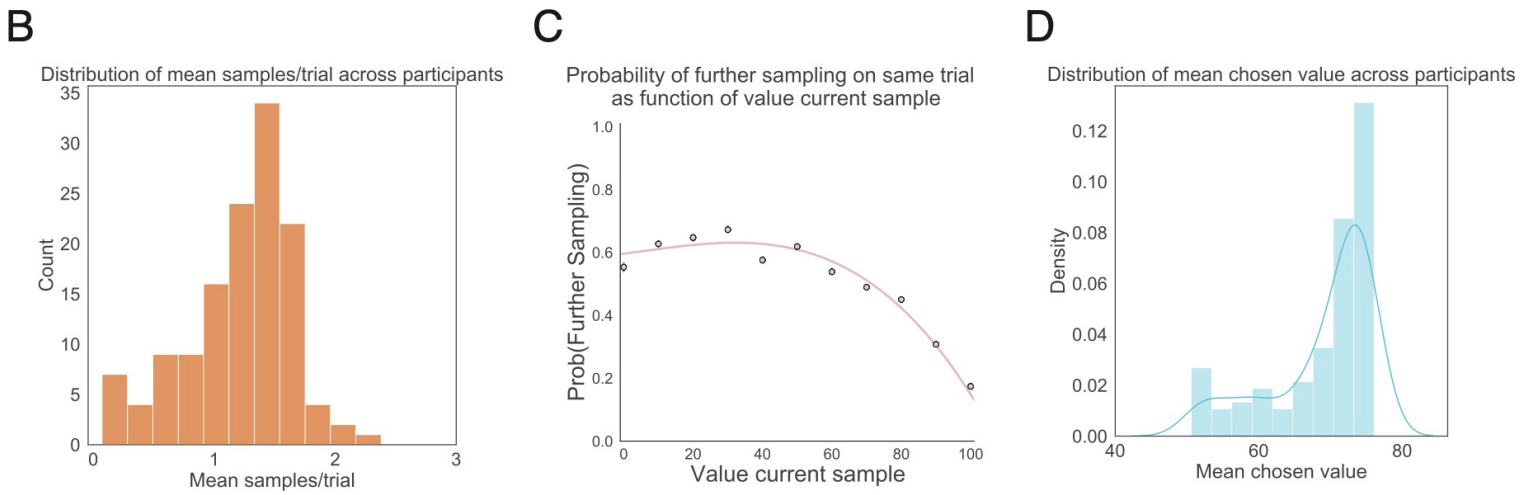
Participants effectively used sampling of the balloons to improve their choices (Figure 5.2D), such that participants' pay-off was significantly above chance ($t(131) = 30.14$, $p<0.001$). Furthermore, participants who sampled more on average, also chose balloons of higher value (β = 0.04, $p < 0.001$).

They were also less likely to sample balloons on a trial, the more balloons they had already sampled on that trial (Figure 5.2E; z = -8.83, $p<0.001$). We observed a

sequential sampling effect of balloon type, as participants were most likely to sample the ‘fast’ balloon first, the ‘medium’ balloon second, and the ‘slow’ balloon last (Figure 5.2E).

Scores on the psychiatric questionnaires were comparable to those found in Wise & Dolan (2020), except for the social withdrawal scores, which were considerably higher in both our samples (see supplemental materials S5.2, section 5.5.2). This might be because data collection took place during the Covid-19 pandemic.


A
Sub 5 - Number Balloons Sampled
Sub 38 - Number Balloons Sampled
Sub 56 - Number Balloons Sampled
Sub 102 - Number Balloons Sampled
Number of balloons
Trial


B
Distribution of mean samples/trial across participants
Count
Mean samples/trial
C
Probability of further sampling on same trial as function of value current sample
Prob(Further Sampling)
Value current sample
D
Distribution of mean chosen value across participants
Density
Mean chosen value

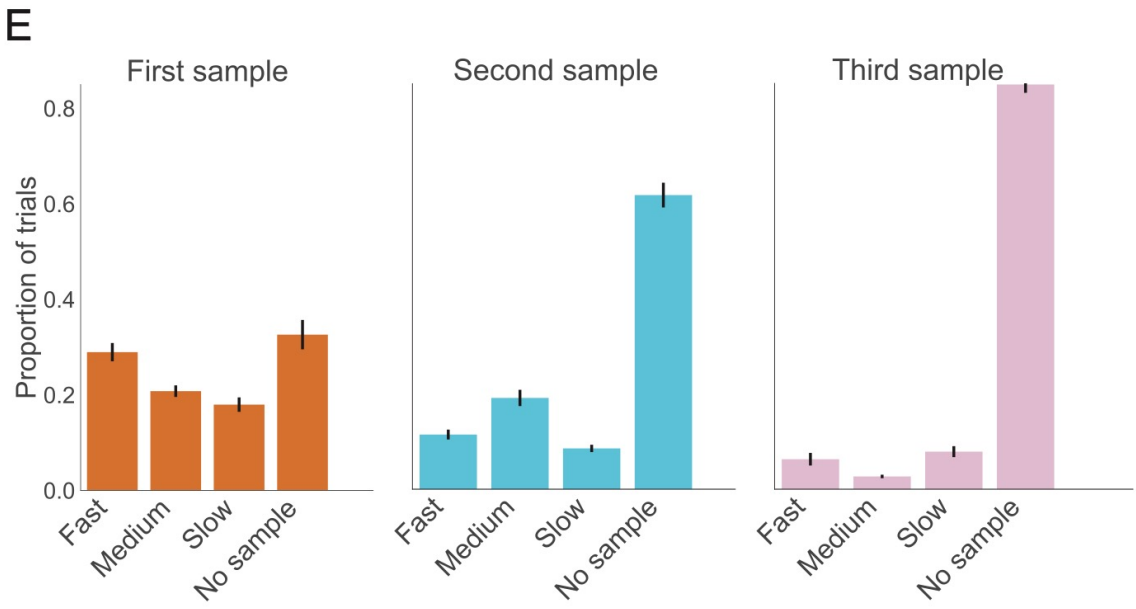

E
First sample
Second sample
Third sample
Proportion of trials
Fast
Medium
Slow
No sample

Figure 5.2. Participant behavior in discovery sample. (A) Samples made per trial for four example participants. It can be seen that behavior on this task is very heterogeneous, ranging from sampling almost no balloons (e.g. participant 38) to often sampling all of them and thereby running out of time before finishing all trials (e.g. participant 56). (B) Individual participants show a wide variety in mean number of balloons sampled per trial. (C) Participants were more likely to stop sampling on a trial when the most recently revealed sample on that trial was of high value. Plotted is the probability of continuing to sample on a trial as a function of the value of the balloon that was just revealed. (D) Participants used the sampling phase to gain information to improve their pay-off on the choice rounds. A participant performing at chance level would have a mean chosen value of 50 points. (E) Sampling became less likely the more balloons participants had already sampled on a trial. Furthermore, the rate of change of the balloons affected their likelihood of being sampled first, second or last in a trial. Specifically, the 'fast' balloon was the most frequent first sample on a trial, the 'medium' balloon the most frequent second sample, and the 'slow' balloon the most frequent third sample. This effect is not due to a left-to-right order effect, as the order of the balloons was reversed for half the participants. Shown are the mean proportion of trials across participants where each of the three balloons were sampled first, second and third, as well as the mean proportion of trials across participants where no balloon was sampled first, second or third, meaning that participants had already moved on to the choice round. Error bars represent the 95% confidence interval (CI). All data shown is from discovery sample.

5.3.1.2 Effect of compulsivity on sampling frequency (discovery sample)

Firstly, we wanted to see if there was a main effect of compulsivity on information sampling as seen in Hauser et al. (2017). In our task, the amount of information sampling done by the participant can be represented as the mean number of balloons sampled per trial by that participant. We did not find a main effect of compulsivity on the mean number of balloons sampled per trial in a linear regression (Figure 5.3A; $t(128) = 0.77$, $p = 0.44$). We did find a positive effect of the anxiety/depression score on the mean number of balloons sampled in the discovery sample (Figure 5.3B; $t(128) = 3.25$, $p < 0.01$). There was no effect of social withdrawal on information sampling (Figure 5.3C; $t(128) = 1.17$, $p = 0.25$). Furthermore, we did not find a main effect of any of the transdiagnostic dimension scores on task performance (compulsivity: $t(128) = 0.30$, $p = 0.76$; anxiety/depression: $t(128) = 0.11$, $p = 0.91$; social withdrawal: $t(128) = -0.07$, $p = 0.94$). We found a significant positive effect of anxiety/depression on response times (RT) on the sampling rounds ($t(128) = 2.18$, $p<0.05$), meaning more anxious/depressed participants took longer to respond

on sampling rounds. There was no effect of the other dimensions on sampling RT (compulsivity: t(128) = -0.44, p = 0.66; social withdrawal: t(128) = 1.12, p = 0.27). We did not find an effect of any of the psychiatric dimensions on choice round RT (compulsivity: t(128) = -1.08, p = 0.28; anxiety/depression: t(128) = 0.85, p = 0.40; social withdrawal: t(128) = 1.82, p = 0.07).

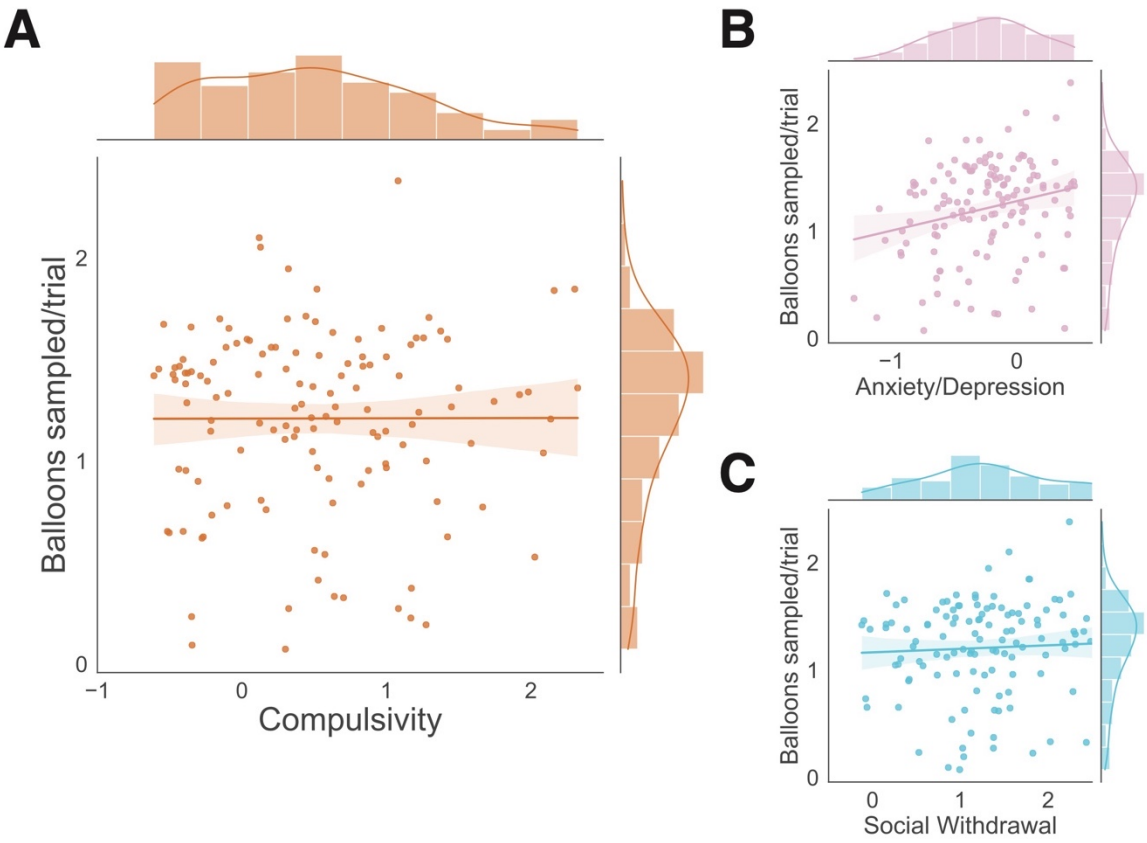


Figure 5.3. Only anxiety/depression scores were positively related to balloons sampled per trial. (A) Individual compulsivity scores plotted against the mean number of balloons sampled per trial. (B) Individual anxiety/depression scores plotted against the mean number of balloons sampled per trial. (C) Individual social withdrawal scores plotted against the mean number of balloons sampled per trial. (A-C) Each plotted data point is a participant. All data shown is from discovery sample.

We hypothesized that while there might not be a main effect of compulsivity on sampling frequency, it is possible that more compulsive individuals might have been

less likely to choose a balloon they had not sampled on that trial. To investigate this, we regressed the three transdiagnostic dimensions scores on the proportion of trials in which no balloons were sampled across participants. We found a significant negative effect of anxiety/depression, meaning that more anxious/depressed participants sampled zero balloons on sampling rounds less frequently ($t(128) = -3.50$, $p < 0.01$). There was no effect of compulsivity ($t(128) = -1.05$, $p = 0.30$) or social withdrawal ($t(128) = -0.57$, $p = 0.57$). We also investigated whether more compulsive individuals were more likely to sample a balloon before choosing it using a hierarchical logistic regression, in which random effects were estimated for each participant. We found that whether a balloon would be chosen on the subsequent choice round was a significant positive predictor of sampling ($z = 17.31$, $p < 0.001$). Using a median split on the compulsivity scores we identified 'low compulsive' and 'high compulsive' participants. We found no significant difference between the random effect coefficients in this regression between the two groups ($t(130) = -1.02$, $p = 0.31$).

5.3.1.3 Effect of compulsivity on interactions between task features used to guide sampling decisions (discovery sample)

As participants sampled on average 1.21 balloons per trial, we decided to focus our analyses on the first sampling decision of each sampling round, as the probability of further sampling on the same sampling round was substantially reduced after that. We found an interaction effect of balloon type ('fast', 'medium' or 'slow' value increase) and the number of trials since a balloon was last sampled ('trials passed') on sampling of the first balloon. This means that the effect of 'trials passed' on sampling decisions was different for each of the three balloons. Specifically, there is a quadratic effect of

trials passed since the previous sample on sampling, where the likelihood of a balloon being sampled first on a trial is highest when it was sampled last 3 trials ago, and lower when it was sampled more or less recently ('fast' balloon: linear effect – $z = 10.78$, $p < 0.001$; quadratic effect – $z = -9.20$, $p < 0.001$; 'medium' balloon: linear effect – $z = 8.82$, $p < 0.001$; quadratic effect – $z = -5.55$, $p < 0.001$; 'slow' balloon: linear effect – $z = 10.24$, $p < 0.001$; quadratic effect – $z = -6.65$, $p < 0.001$). This main effect of 'trials passed' is strongest for the 'fast' balloon and weakest for the 'slow' balloon (Figure 5.4). The main effects of balloon type and trials passed on sampling together with their interaction can be summarized as the area under the curve (AUC) describing the relationship between trials passed and the probability of the first sample for each balloon type. As can be seen in Figure 5.4, this AUC is highest for the 'fast' balloon, and lowest for the 'slow' balloon, and the AUC for each balloon type is significantly different from the others (fast versus medium: $t(130) = 2.40$, $p < 0.05$; medium versus slow: $t(130) = 4.46$, $p < 0.001$).

We hypothesized that the quadratic effect of 'trials passed' on sampling probability of the first sample in a round could be caused by a specific sampling strategy. Specifically, participants may sometimes repeatedly sample and choose the same balloon for a few trials, which would mean 'trials passed' would negatively predict sampling probability for that balloon. For the other non-sampled balloons, however, the relationship would be positive. To test this, we plotted the same relationship excluding balloons which were just chosen on the previous trial or about to be chosen on the current trial (see supplementary material S5.3, section 5.5.3). This did not affect the nature of the relationship between 'trials passed' and sampling

probability, meaning repeatedly sampling and choosing the same balloon as a sampling strategy did not strongly affect this.

While compulsivity does not appear to affect overall sampling frequency in this task, it is possible that it does affect how features of the task are used to make sampling decisions. To investigate this, we analyzed the interaction effect described above as a function of participants' compulsivity scores. We found an effect of compulsivity on the interaction effect between task features as described by the AUC. We used a median split to categorize participants as either 'low compulsive' or 'high compulsive'. We found that the AUC differences between the 'fast' and 'medium' balloons, as well as between the 'fast' and 'slow' balloons were significantly smaller in high compulsive participants (fast versus medium: $t(130) = -2.69$, $p < 0.01$; fast versus slow: $t(132) = -3.24$, $p < 0.01$). Only the AUC difference between the medium and slow balloon was not significantly different between the compulsivity groups ($t(130) = 1.63$, $p = 0.11$). As participants did not often sample more than 1 balloon on each trial, including all samples into this analysis produced similar results (see supplementary material S5.4, section 5.5.4). We collected a second participant sample in an attempt to replicate this effect.

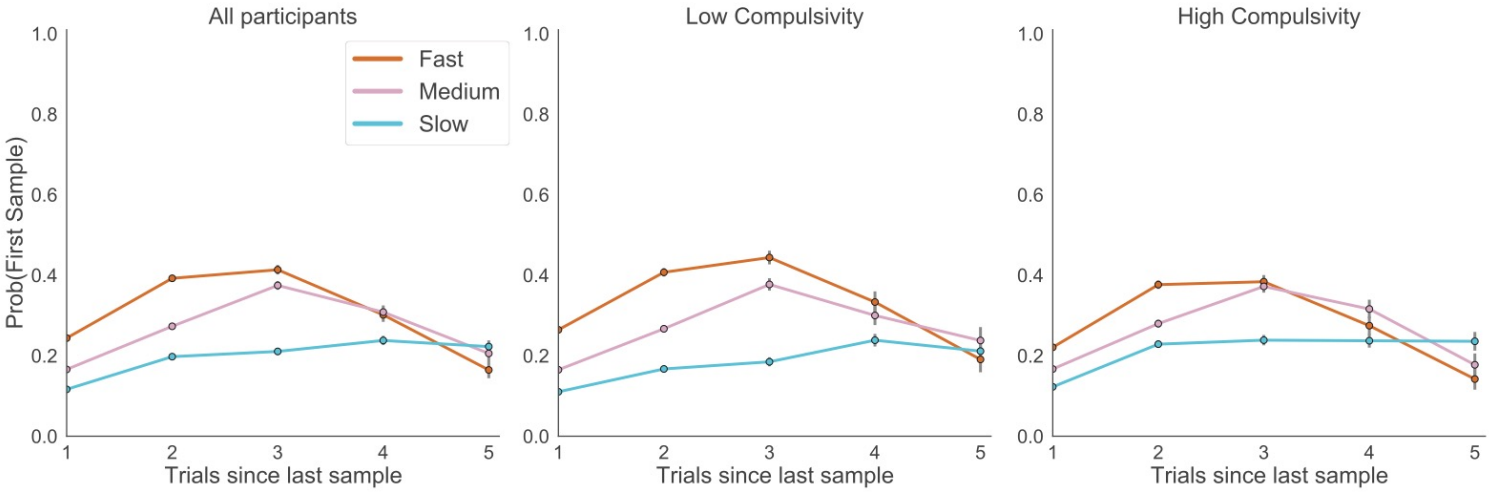


Figure 5.4. There is an interaction effect between balloon type and trials passed since the previous sample on likelihood of the first sample on a trial, and this may be mediated by compulsivity. Highly compulsive participants showed a stronger mediation effect of balloon type on the effect of trials passed on sampling. The quadratic effect of 'trials passed' on the probability of a first sample is highest for the 'fast' balloon and lowest for the 'slow' balloon. Plotted are the mean proportion of trials where a balloon was first sampled as a function of how many trials ago that balloon was last sampled for each balloon type. Error bars are standard errors of the mean. All data shown here is from the discovery sample.

5.3.2 Confirmation sample

Having found an effect of compulsivity on the interaction effect of balloon type and trials since a previous sample on sampling behavior as well as an effect of anxiety/depression on sampling frequency and the number of 'zero sample' trials in the discovery sample, we proceeded to collect a confirmation sample of an additional 300 participants. If it was the case that these effects are genuine, we expected to find similar results in the confirmation sample.

5.3.2.1 Task performance (confirmation sample)

Participants' behavior in this sample was comparable to that in the discovery sample: there again was much heterogeneity in sampling behavior (Figure 5.5A-B). A

split-half analysis showed significant correlation between the mean number of balloons sampled in two randomly sampled halves of the task for each participant, indicating the individual differences are reliable (β = 0.95, p < 0.001). Participants who sampled more on average also completed fewer trials due to the waiting costs associated with sampling in the task (β = -0.02, p < 0.001).

Participants effectively used the sampling rounds to make better decisions on the choice rounds (Figure 5.5D), and as such showed above chance performance (t(253) = 32.71, p < 0.001). Mean number of balloons sampled was also significantly correlated with the mean value of the chosen balloon, such that participants who sampled more on average, picked balloons of higher value (β = 0.04, p < 0.001). The value of the most recently sampled balloon negatively predicted further sampling on that trial (Figure 5.5C; z = -21.95, p < 0.001) and the number of balloons already sampled on a trial was again a negative predictor of further sampling (Figure 5.5E; z = -11.58, p < 0.001). Finally, we observed the same sequential sampling effect in this sample, where participants were most likely to sample the 'fast' balloon first, the 'medium' balloon second, and the 'slow' balloon last (Figure 5.5E). Scores on the psychiatric questionnaires were also comparable to the discovery sample (see supplementary material S5.1 in section 5.5.1).

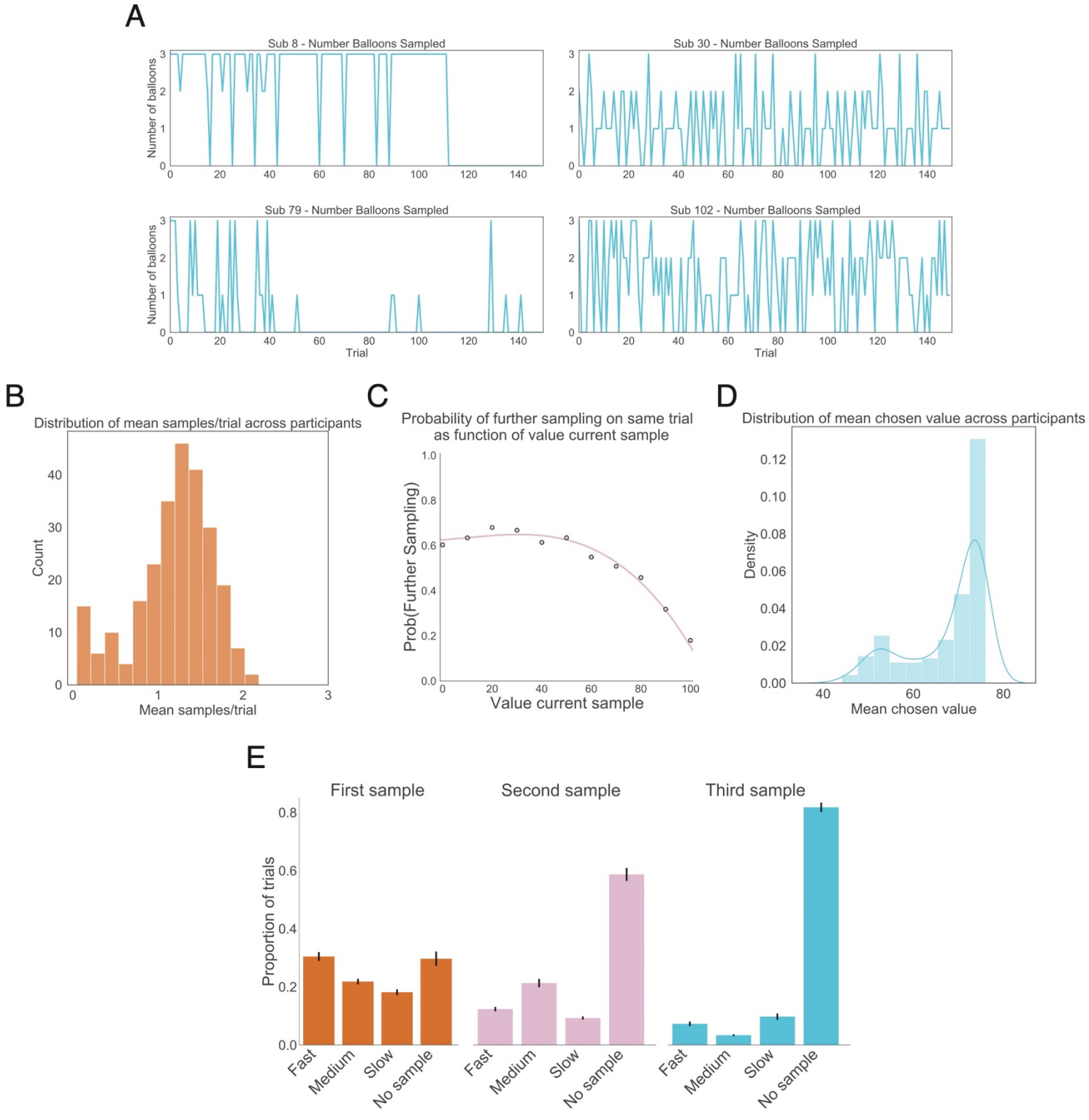


Figure 5.5. Participant behavior in the confirmation sample. (A) Samples made per trial for four example participants. It can be seen that behavior on this task is very heterogeneous, ranging from sampling almost no balloons (e.g. participant 79) to often sampling all of them and thereby running out of time before finishing all trials (e.g. participant 8). (B) Individual participants show a wide variety in mean number of balloons sampled per trial. (C) Participants were more likely to stop sampling on a trial when the most recently revealed sample on that trial was of high value. Plotted is the probability of continuing

to sample on a trial as a function of the value of the balloon that was just revealed. (D) Participants used the sampling phase to gain information to improve their pay-off on the choice rounds. A participant performing at chance level would have a mean chosen value of 50 points. (E) Sampling became less likely the more balloons participants had already sampled on a trial. Furthermore, the rate of change of the balloons affected their likelihood of being sampled first, second or last in a trial. Specifically, the 'fast' balloon was the most frequent first sample on a trial, the 'medium' balloon the most frequent second sample, and the 'slow' balloon the most frequent third sample. This effect is not due to a left-to-right order effect, as the order of the balloons was reversed for half the participants. Shown are the mean proportion of trials across participants where each of the three balloons were sampled first, second and third, as well as the mean proportion of trials across participants where no balloon was sampled first, second or third, meaning that participants had already moved on to the choice round. Error bars represent the 95% confidence interval (CI). Data shown is from confirmation sample.

5.3.2.2 Effect of compulsivity on sampling frequency (confirmation sample)

As before, we did not find a main effect of compulsivity on information sampling (Figure 5.6A; $t(250) = -0.28$, $p = 0.78$). Furthermore, while we found a positive effect of anxiety/depression on sampling in the discovery sample, this was not replicated in this sample (Figure 5.6B; $t(250) = -0.91$, $p = 0.37$). There was no effect of social withdrawal on information sampling (Figure 5.6C; $t(250) = -0.82$, $p = 0.41$). We did not find a main effect of any of the transdiagnostic dimension scores on task performance (compulsivity: $t(250) = -0.71$, $p = 0.48$; anxiety/depression: $t(250) = -0.31$, $p = 0.76$; social withdrawal: $t(250) = 0.85$, $p = 0.40$), response times on the sampling rounds (compulsivity: $t(250) = 0.67$, $p=0.50$; anxiety/depression: $t(250) = 1.15$, $p=0.25$; social withdrawal: $t(250) = -0.96$, $p = 0.34$), or response times on the choice rounds (compulsivity: $t(250) = -0.31$, $p = 0.76$; anxiety/depression: $t(250) = -0.10$, $p = 0.92$; social withdrawal: $t(250) = -0.49$, $p = 0.63$). We also did not replicate the effect of anxiety/depression on the amount of 'zero sample' trials in this sample ($t(250) = 0.40$, $p = 0.69$), and there was no effect of compulsivity ($t(250) = 0.06$, $p = 0.95$) and social withdrawal either ($t(250) = 0.77$, $p = 0.44$).

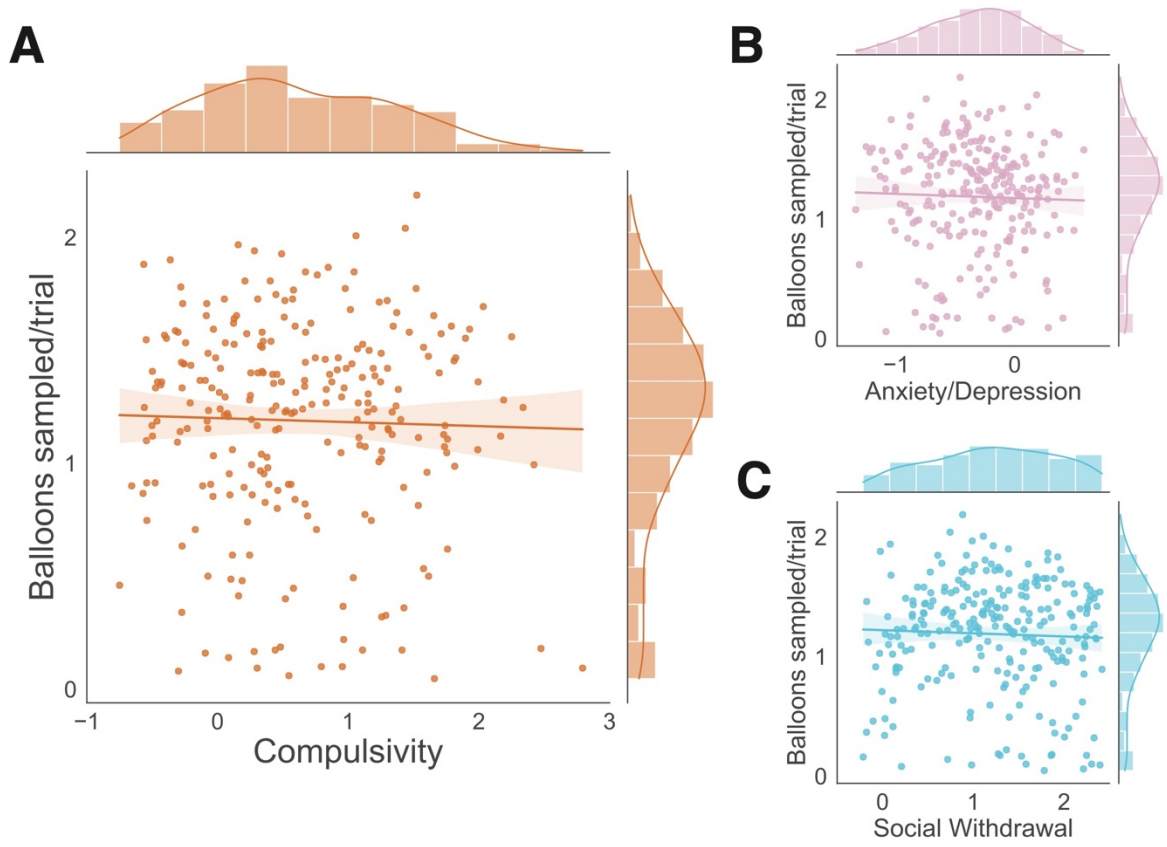


Figure 5.6. Transdiagnostic psychiatric dimension scores did not affect the number of balloons participants sampled on the task. (A) Individual compulsivity scores plotted against the mean number of balloons sampled per trial. (B) Individual anxiety/depression scores plotted against the mean number of balloons sampled per trial. (C) Individual social withdrawal scores plotted against the mean number of balloons sampled per trial. (A-C) Each plotted data point is a participant. Data shown is from confirmation sample.

5.3.2.3 Effect of compulsivity on interactions between task features used to guide sampling decisions (confirmation sample)

In the confirmation sample, there was also an interaction effect between balloon type and the number of trials since a balloon was last sampled on sampling of the first balloon (Figure 5.7; 'fast' balloon: linear effect – $z = 12.36$, $p < 0.001$; quadratic effect – $z = -10.14$, $p < 0.001$; 'medium' balloon: linear effect – $z = 11.49$, $p < 0.001$; quadratic effect – $z = -7.79$, $p < 0.001$; 'slow' balloon: linear effect – $z = 10.83$, $p < 0.001$; quadratic effect – $z = -6.44$, $p < 0.001$). The AUCs for the three balloon types were all

significantly different from one another ('fast' versus 'medium': $t(252) = 2.62$, $p < 0.01$; 'fast' versus 'slow': $t(252) = 9.55$, $p < 0.001$; 'medium' versus 'slow': $t(252) = 6.86$, $p < 0.001$). However, unlike in the discovery sample, these differences were not affected by compulsivity (high versus low compulsivity: 'fast' AUC - 'medium' AUC: $t(252) = 0.22$, $p = 0.83$; 'fast' AUC – 'slow' AUC: $t(252) = 0.49$, $p = 0.63$; 'medium' AUC – 'slow' AUC: $t(252) = 0.40$, $p = 0.69$). The same analysis was run predicting the second and third samples, but no effects of compulsivity were seen there either (see supplemental materials S5.5-6, sections 5.5.5-6).

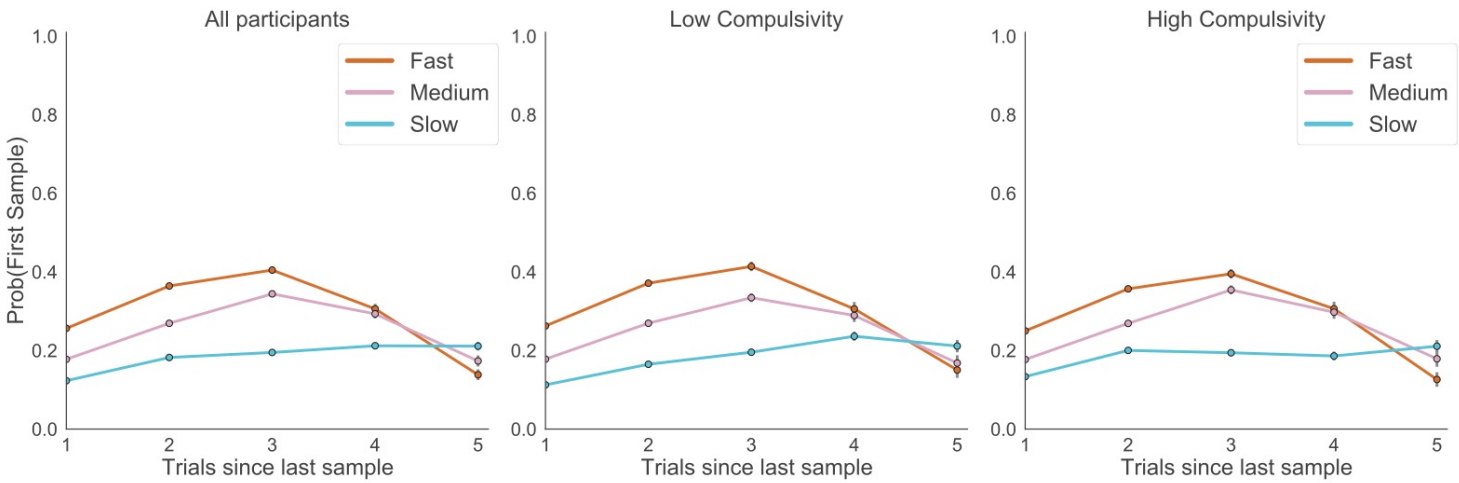


Figure 5.7. There is an interaction effect between balloon type and trials passed since the previous sample on likelihood of the first sample on a trial, but this is not mediated by compulsivity. The quadratic effect of 'trials passed' on the probability of a first sample is highest for the 'fast' balloon and lowest for the 'slow' balloon. Plotted are the mean proportion of trials where a balloon was first sampled as a function of how many trials ago that balloon was last sampled for each balloon type. Error bars are standard errors of the mean. Data shown is from the confirmation sample.

5.4 Discussion

Compulsivity is associated with repetitive and inflexible behaviors that have been described as 'habitual' (Gillan et al., 2011), but also with increased decision

thresholds in decision-making tasks (Banca et al., 2015; Erhan et al., 2017), due to which it is unclear to what extent highly compulsive individuals can flexibly use information to guide choices and future sampling. In this chapter, I have presented findings from a large-scale online study combining a novel information gathering task with the collection of transdiagnostic dimension scores showing no effect of compulsivity on information sampling, or on the mediation of task features on the link between compulsivity and sampling.

Recently, the field of computational psychiatry has been moving towards a transdiagnostic approach: trying to isolate behaviorally relevant symptoms, rather than using clinically defined disorders (Gillan et al., 2016). It has been argued that the difficulty in investigating the neurobiological basis of different clinical disorders is that many of the symptoms overlap across disorders (APA, 2013; Poyurovsky & Koran, 2005; Ruscio et al., 2010). This means that when the behavior of a single patient group is studied it is often unclear to what extent findings are specific to that patient group and clinical disorder (Gillan et al., 2016). This also means that patient groups defined as having the same clinical disorder often exhibit much heterogeneity (i.e. different combinations of symptoms; Cuthbert & Kozak, 2013), which can make it unclear which disordered behavior is being studied and lead to difficulties replicating findings across samples.

This approach has been used here as well, by using a set of questionnaires made up of multiple diagnostic questionnaires to extract scores on three transdiagnostic dimensions. This contrasts with the other studies discussed that investigated the link between compulsivity and information sampling, in which a sample of individuals scoring highly on an OCD questionnaire (usually patients) were

compared to a sample of healthy controls. As such, this paper presents the first attempt at identifying a link between compulsivity as a transdiagnostic dimension and information sampling in a subclinical population. The fact that we have not been able to find such a link, may mean that there is no linear relationship between information sampling and compulsivity, but that instead information sampling is only altered in individuals with unusually high levels of compulsivity. As this transdiagnostic approach is increasingly being used, it would be useful for future studies to directly compare the relationships between psychiatric diagnoses and behavior to those between transdiagnostic dimensions and behavior (Gillan et al., 2016), to directly test how much can be gained from this approach.

While participants in our task did not pay a monetary cost for sampling, they did pay a time cost of 3s every time they sampled. Because of this, it is possible that our task design was in fact more comparable to Hauser et al.'s (2017) 'decreasing' task condition in which reward decreased as a function of the number of samples participants took to make a choice. In this condition, they actually found no difference in the amount of information sampling between participants with low and high compulsivity, suggesting that compulsive individuals are able to flexibly adjust their strategy to maximize reward. Other studies have also found a difference in compulsive participants' behavior under different sampling costs (Banca et al., 2015), but Chamberlain et al. (2007) found no effect of compulsivity on sampling regardless of sampling cost.

Furthermore, as each sampling round was followed by a choice round and balloons changed across trials in a predictable way, sampling may not have been as consequential to the participant as in Hauser et al.'s (2017) task. In our task,

participants were able to use their knowledge of how the balloons increased in value over time to predict their future values to guide their choices, while in most other information sampling tasks used to study compulsivity, trials were independent and sampling on each trial was essential to guide decisions. It is therefore possible that compulsive individuals spent more time mentally simulating the potential values of the balloon based on their knowledge of the different increase rates of the balloons, rather than sampling more from the balloons themselves. If this were the case, however, we would expect response times to be longer for more compulsive participants, which is not what we found.

It could also be the case that information sampling is simply not altered in compulsive individuals. There have been previous null results in studies trying to demonstrate a link (Chamberlain et al., 2007), as well as results suggesting other traits correlated with compulsivity may instead be associated with increased information sampling (Volans, 1976). There is also strong evidence for distrust in memory in OCD patients (Boschen & Vuksanovic, 2007; Hermans et al., 2008; Olley et al., 2007; Tolin et al., 2001), suggesting increased information sampling could arise from down-weighting earlier evidence due to distrust in the accuracy of that memorized evidence. In our task, memory confounds were eliminated by presenting all necessary information on the screen: rates of value increase for each of the balloons ('fast', 'medium' or 'slow'), how many trials ago a balloon was last sampled and what value the balloon had when it was last sampled. This means that if increased information sampling in compulsive individuals is due to distrust in memory, we might not see increased information sampling in a task where participants do not have to rely on their memory of past samples for evidence accumulation. However, the information

gathering task used by Hauser et al. (2017) also did not require memorization of information samples, and they did find an effect of compulsivity.

In summary, we did not find a relationship between compulsivity and information sampling in a novel information gathering task. We also did not find a difference in how task features mediated this relationship. It is possible that our task was not well-suited to find this effect, but a split-half analysis showed that the individual differences in behavior seen on the task were reliable. There have been some null results in other previous studies studying this effect as well, suggesting the effect is perhaps weak or only present in certain subpopulations.

5.5 Supplementary materials

5.5.1 Psychiatric questions and dimension weights

Items	Dimension weights		
	Comp	**AD**	**SW**
Self-rating Depression Scale (Zung, 1965)			
My mind is as clear as it used to be.	0	0.030	0
I find it easy to do the things I used to.	0	0.020	0
I am restless and can't keep still.	0.093	0	0
I feel hopeful about the future.	0	0.029	0
I find it easy to make decisions.	0	0.011	0
I feel that I am useful and needed.	0	0.120	0
My life is pretty full.	0	0.035	0
I still enjoy the things I used to do.	0	0.064	0
State-Trait Anxiety Inventory (Spielberger, 1970)			
I feel pleasant.	0	0.012	0
I feel satisfied with self.	0	0.114	0
I feel like a failure.	0	0.023	0
I feel that difficulties are piling up so that I can't overcome them.	0	0.017	0
I worry too much over something that really doesn't matter.	0.014	0	0
I am happy.	0	0.104	0
I lack self-confidence.	0	0.020	0.023
I feel secure.	0	0.01	0
I am content.	0	0.034	0
I am a steady person.	0	0.051	0
I get in a state of tension or turmoil as I think over my recent concerns and interests.	0.016	0	0
Barratt Impulsivity Scale (Patton et al., 1995)			
I plan tasks carefully.	0	0.028	0
I have racing thoughts.	0.073	0	0
I concentrate easily.	0	0.059	0
I plan for job security.	0	0.021	0
I say things without thinking.	0.018	0	0
I like to think about complex problems.	-0.030	0	0
I act on impulse.	0.090	0	0
I am a steady thinker.	0	0.014	0
I buy things on impulse.	0.004	0	0
I spend or charge more than I earn.	0.045	0	0
I often have extraneous thoughts when thinking.	0.041	0	0
Apathy Evaluation Scale (Marin et al., 1991)			
I get things done during the day.	0	0.064	0
I approach life with intensity.	0	0.002	0
I have initiative.	0	0.052	0
I have motivation.	0	0.151	0
Obsessive-Compulsive Inventory – Revised (Foa et al., 2002)			
I have saved up so many things that they get in the way.	0.024	0	0
I check things more often than necessary.	0.036	0	0
I feel compelled to count while I am doing things.	0.062	0	0
I find it difficult to control my own thoughts.	0.148	0	0
I collect things I don't need.	0.021	0	0

I get upset if others change the way I have arranged things.	0.047	0	0.019
I sometimes have to wash or clean myself simply because I feel contaminated.	0.025	0	0
I am upset by unpleasant thoughts that come into my mind against my will.	0.136	0	0
I avoid throwing things away because I am afraid I might need them later.	0.008	0	0
I feel that there are good and bad numbers.	0.025	0	0
I frequently get nasty thoughts and have difficulty in getting rid of them.	0.080	0	0
Eating Attitudes Test (Garner et al., 1982)			
I am terrified about being overweight.	0.127	0	0
I am preoccupied with a desire to be thinner.	0.039	0	0
I think about burning up calories when I exercise.	0.008	0	0
I am preoccupied with the thought of having fat on my body.	0.010	0	0
Liebowitz Social Anxiety Scale (Liebowitz, 1987)			
Participating in small groups.	0	0	0.017
Going to a party.	0	0	0.150
Working while being observed.	0	0	0.007
Calling someone you don't know very well.	0	0	0.012
Talking with people you don't know very well.	0	0	0.132
Meeting strangers.	0	0	0.131
Entering a room when others are already seated.	0	0	0.035
Being the center of attention.	0	0	0.168
Speaking up at a meeting.	0	0	0.173
Expressing a disagreement or disapproval to people you don't know very well.	0	0	0.031
Giving a report to a group.	0	0	0.072
Giving a party.	0	0	0.072
Resisting a high pressure salesperson.	0	0	0.022
Alcohol Use Disorders Identification Test (Saunders et al., 1993)			
How often do you have a drink containing alcohol?	0	0	-0.005

Table S1.1 63 psychiatric questionnaire items with classifier weights from (Wise & Dolan, 2020)

5.5.2 Transdiagnostic dimension score distributions

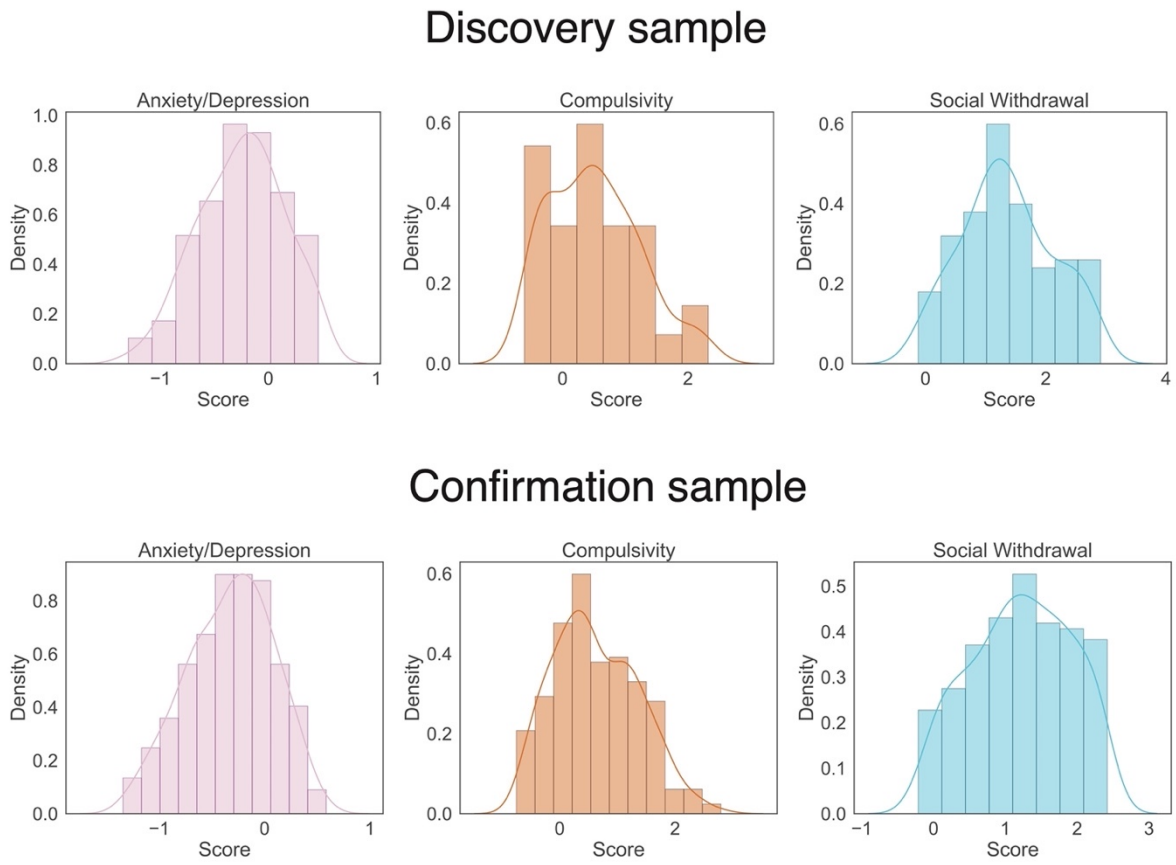


Figure S2.1. Distributions of the three transdiagnostic psychiatric dimension scores in both samples.

5.5.3 Excluding balloons chosen on previous and current trials

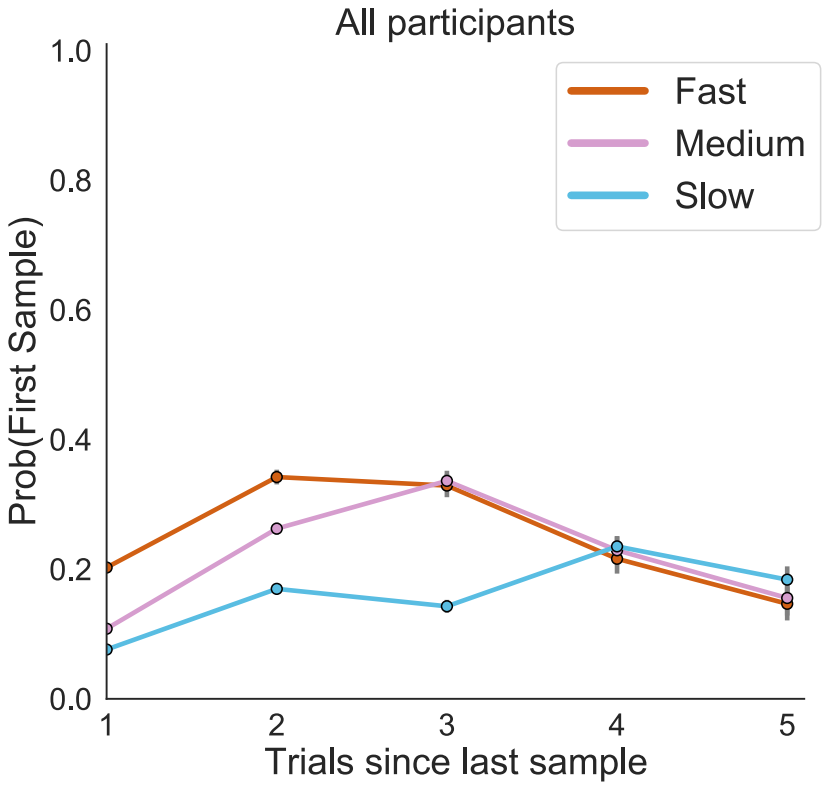


Figure S3.1. Removing balloons which were chosen on the previous choice round or are about to be chosen on the current choice round, does not affect the shape of the relationship between 'trials passed' and sampling probability on the first sample of a trial. Plotted are the mean proportion of trials where a balloon was first sampled as a function of how many trials ago that balloon was last sampled for each balloon type, excluding balloons that were chosen in the previous choice round or are about to be chosen on the current choice round. Error bars are standard errors of the mean. Data shown is from the discovery sample.

5.5.4 Including all samples in analysis of interaction effect between 'trials passed' and balloon type on sampling

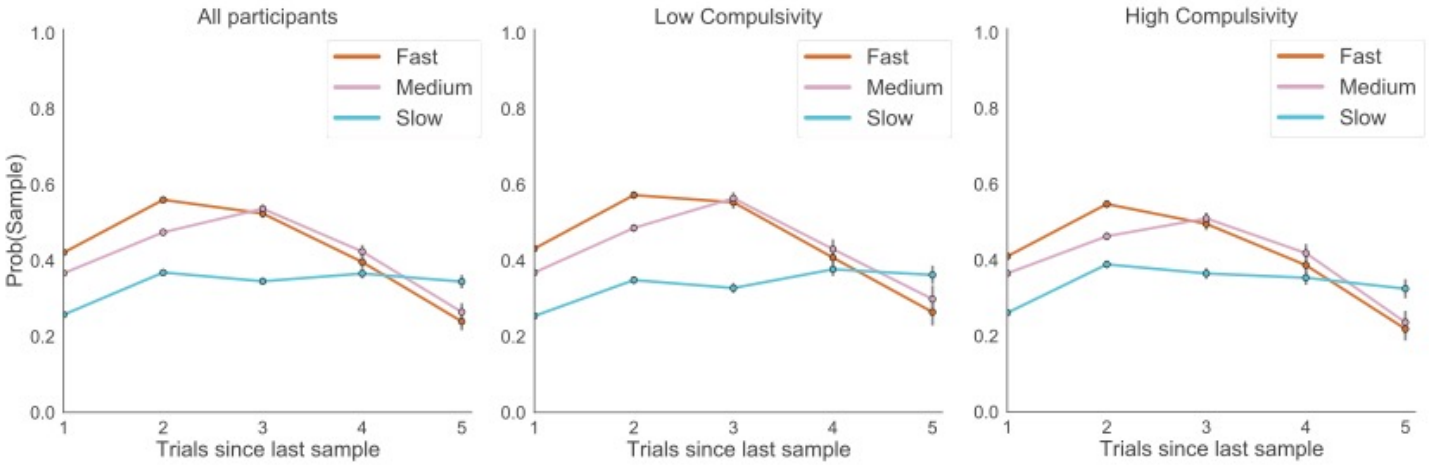


Figure S4.1 Including all samples (rather than just the first sample on each sampling round) into our analyses does not affect the shape of the relationship between 'trials passed' and sampling probability on the first sample of a trial. Plotted are the mean proportion of trials where a balloon was sampled as a function of how many trials ago that balloon was last sampled for each balloon type. Error bars are standard errors of the mean. Data shown is from the discovery sample.

5.5.5 Effect of task features on second sample

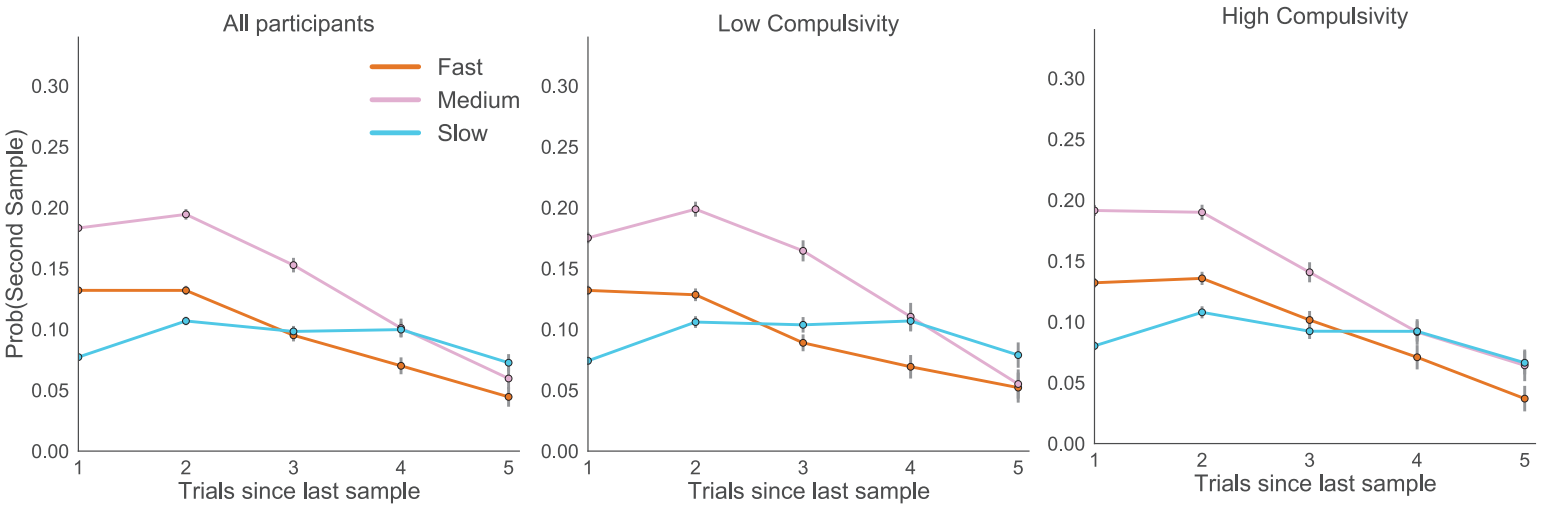


Figure S5.1 There is an interaction effect between balloon type and trials passed since the previous sample on likelihood of the second sample on a trial, but this does not seem strongly mediated by compulsivity. Plotted are the mean proportion of trials where a balloon was sampled second as a function of how many trials ago that balloon was last sampled for each balloon type. Error bars are standard errors of the mean. Data shown is from the confirmation sample.

5.5.6 Effect of task features on third sample

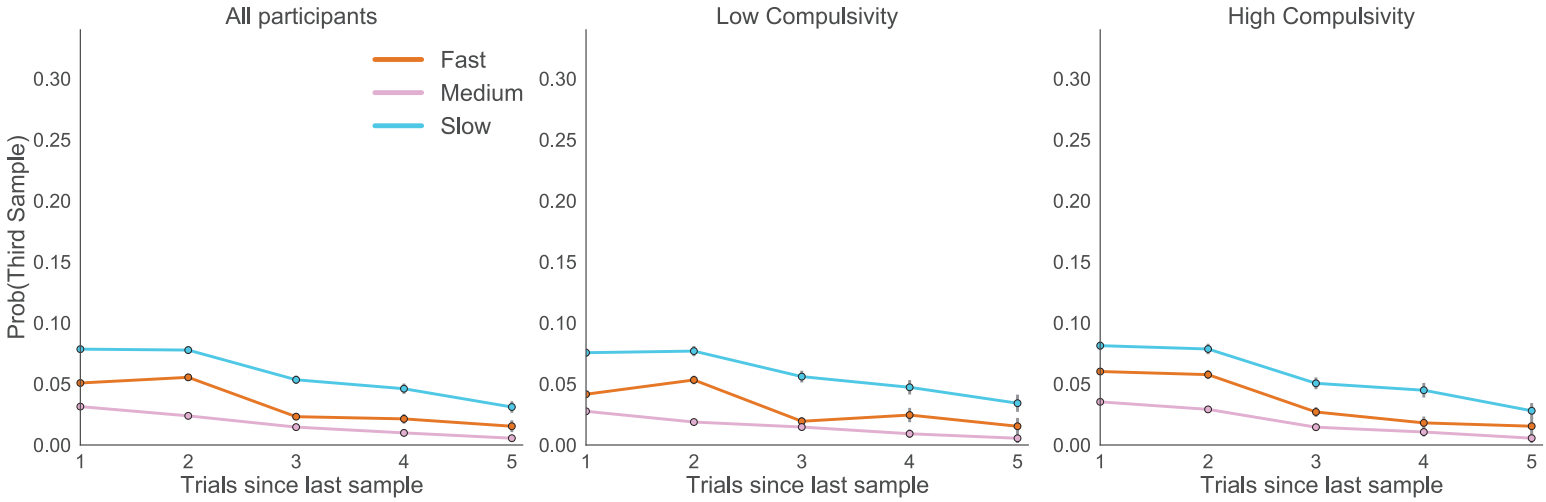


Figure S6.1 There is an interaction effect between balloon type and trials passed since the previous sample on likelihood of the third sample on a trial, but this does not seem strongly mediated by compulsivity. Plotted are the mean proportion of trials where a balloon was sampled third as a function of how many trials ago that balloon was last sampled for each balloon type. Error bars are standard errors of the mean. Data shown is from the confirmation sample.

Chapter 6: Goal proximity mediates activity in a widespread brain network involved in information sampling to guide economic choice

Abstract

Medial frontal and parietal cortex have been shown to co-activate in many different decision-making tasks. Their roles in attention and foraging behavior make them likely candidates to drive decisions to sample more information to guide future choice by representing the goal-relevance of available information. We have used functional MRI combined with a novel decision-making task with active sequential sampling decisions to investigate the neural signatures of the goal-relevance of available information. In this task, the goal-relevance of information is manipulated by varying goal proximity, information reliability and sampling cost. We find that all three features affect decisions to sample. Besides this, we show that activity in a large brain network encompassing anterior cingulate cortex, parietal cortex, frontal pole and dorsolateral prefrontal cortex represents only goal proximity. This suggests this widespread brain network may be involved in flexible behavior change as goals change over time.

6.1 Introduction

To be able to find information that can help us make decision in the future, we need an internal model of what information is likely to be relevant for those decisions. Indeed, recent studies have shown that attention is drawn to more goal-relevant information (Sepulveda et al., 2020; Tanner & Itti, 2019; Tatler et al., 2011), but to what extent this might drive information search to guide future actions is poorly understood. Here we present an fMRI study that investigates how brain networks involved in

sampling and decision-making represent this goal-relevance and then drive information sampling.

Several prefrontal and parietal brain regions have repeatedly been linked to information sampling in decision-making tasks. Firstly, activity in pre-supplementary motor area (pre-SMA) and dorsal anterior cingulate cortex (dACC), subregions of medial frontal cortex (MFC) have been shown to predict decisions to sample in a decision-making task (Blanchard & Gershman, 2018; Kaanders et al., 2021) as well as foraging behavior (Daw et al., 2006; Kolling et al., 2012, 2018; Trudel et al., 2021). A subregion of the superior parietal lobule is believed to be involved in planning visual allocation and might be a human homologue of the lateral intraparietal area (LIP) in non-human primates, which is known to drive visual attention (Mars et al., 2011; Rushworth et al., 2001). Activity in the intraparietal sulcus (IPS) has been found to predict sampling decisions or information gain in comparable economic choice tasks (Horan et al., 2019; Iigaya et al., 2020; Kaanders et al., 2021; van Lieshout et al., 2018). Indeed, we found increased activity in a large network, including pre-SMA and parietal cortex, predicted upcoming decisions to sample additional information in Kaanders et al. (2021; Figure 6.1; see **chapter 3**).

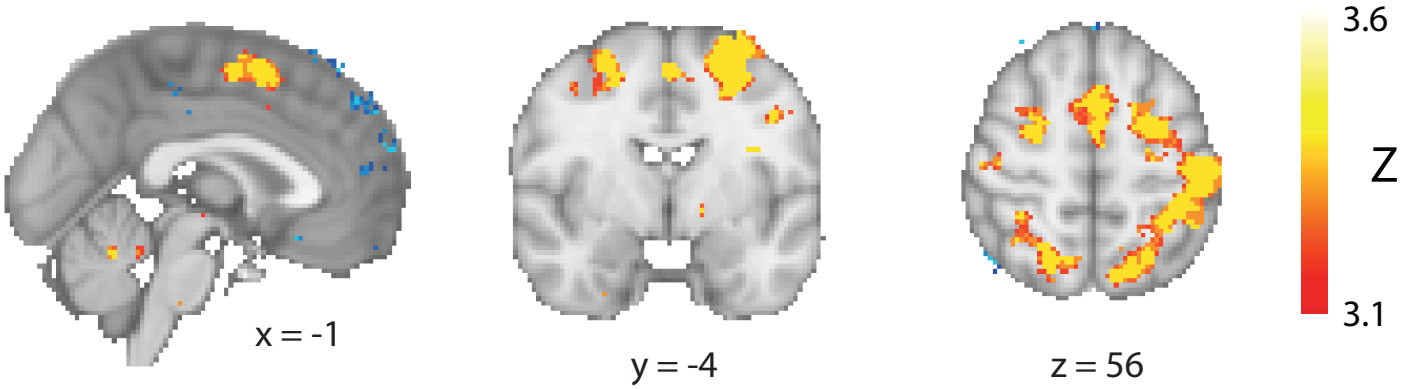


Figure 6.1 Decisions to sample information are predicted by increased brain activity in a large network including MFC and parietal cortex in Kaanders et al. (2021; see **chapter 3**). Parametric map is uncorrected and unthresholded.

In general, these regions also appear to frequently co-activate in decision-making tasks when information is presented that suggests something has changed about the environment, such as when a default choice option is no longer rewarding and it is better to switch to an alternative option (Akam et al., 2021; Domenech et al., 2020; Isoda & Hikosaka, 2007; M. P. Karlsson et al., 2012; Kolling et al., 2016; Molenberghs et al., 2007; O'Reilly et al., 2013; Schuck et al., 2015).

Furthermore, there is evidence that these same regions may encode the goal-relevance of information. Recently, Kobayashi et al. (2021) showed activity reflecting the instrumental value of information for upcoming decisions, i.e. the hypothesized gain in expected value as a result of sampling information, in a large brain network encompassing MFC and parietal regions. Crucially, they find activity in these areas in response to the value of information is seen close to the decision time, dissociating it from dorsolateral prefrontal cortex (dlPFC), which encodes the value of information

throughout the task. This makes this network a prime candidate to drive decisions to sample information to guide an upcoming choice.

If the activity in these regions is indeed driving sampling towards goal-relevant information, this activity is likely to also encode the probability that a decision to sample will provide goal-relevant information. To investigate this, we have designed a novel task in which participants make active decisions to sample information to guide future reward-maximizing choices during BOLD fMRI data collection. We manipulated the number of samples available before a choice must be made to investigate the effect of goal proximity on sampling behavior and sampling-related brain activity. Besides this, we varied the information reliability of different choice options as well as the sampling cost. We find that sampling behavior is affected by all these three task features that indicate the goal-relevance of available information. Of these task features, only goal proximity is reflected in the fMRI signal in a widespread network encompassing MFC, parietal cortex, frontal pole and dlPFC. This effect is partly explained away when including a co-regressor describing how far along the participant is in the block, but the same pattern of activity can still be seen in unthresholded parametric maps. We also see strong activation in this same set of regions in response to new information being revealed after they have decided to sample this information. Finally, we find significant activity in parietal cortex predictive of upcoming decisions to sample additional information, similar to that seen in Kaanders et al. (2021, see **chapter 3**). We do not find significant activity in pre-SMA as in **chapter 3**, but we do see sub-threshold activation in ACC associated with decisions to sample.

6.2 Results

30 healthy human participants (aged 18-65 years) performed a decision-making task including active decisions to sample information inside the MRI scanner (Figure 6.2A). In the task, participants' goal was to pick the balloon with the highest value out of three balloons in a choice round at the end of each block. Each block consisted of 3-8 sampling rounds in which participants could pay points to reveal the value of one of the three balloons. The number of rounds per block was varied in this way to investigate the effect of goal proximity on sampling decisions. On each sampling round, one of the three balloons was highlighted, and a sampling cost was displayed on a bar of between 1-10 points. At this point participants could decide to either 'accept' or 'reject' this sample at the displayed cost by pressing the corresponding button on a button box. If they accepted the sample, a value bar next to the sampled balloon would update to its current value (between 1-100 points). Each balloon had a probability of 'popping' on each trial and being replaced by a balloon with a new value. In other words, there was a possibility that a sampled balloon's value was no longer its true value by the time participants got to the choice round. This was done to investigate the effect of information reliability on sampling decisions. The only way for participants to find out if a balloon had popped since a previous sample was to sample it again. The three balloons had different probabilities of popping on a given trial: 12.5%, 25% and 50%, respectively. In the choice round at the end of the block, participants could choose one of the three balloons by pressing the corresponding button on a button box. After doing so, participants were shown how many points they had won on that block, which corresponded to the current value of the chosen balloon minus the total number of points spent on sampling in the block.

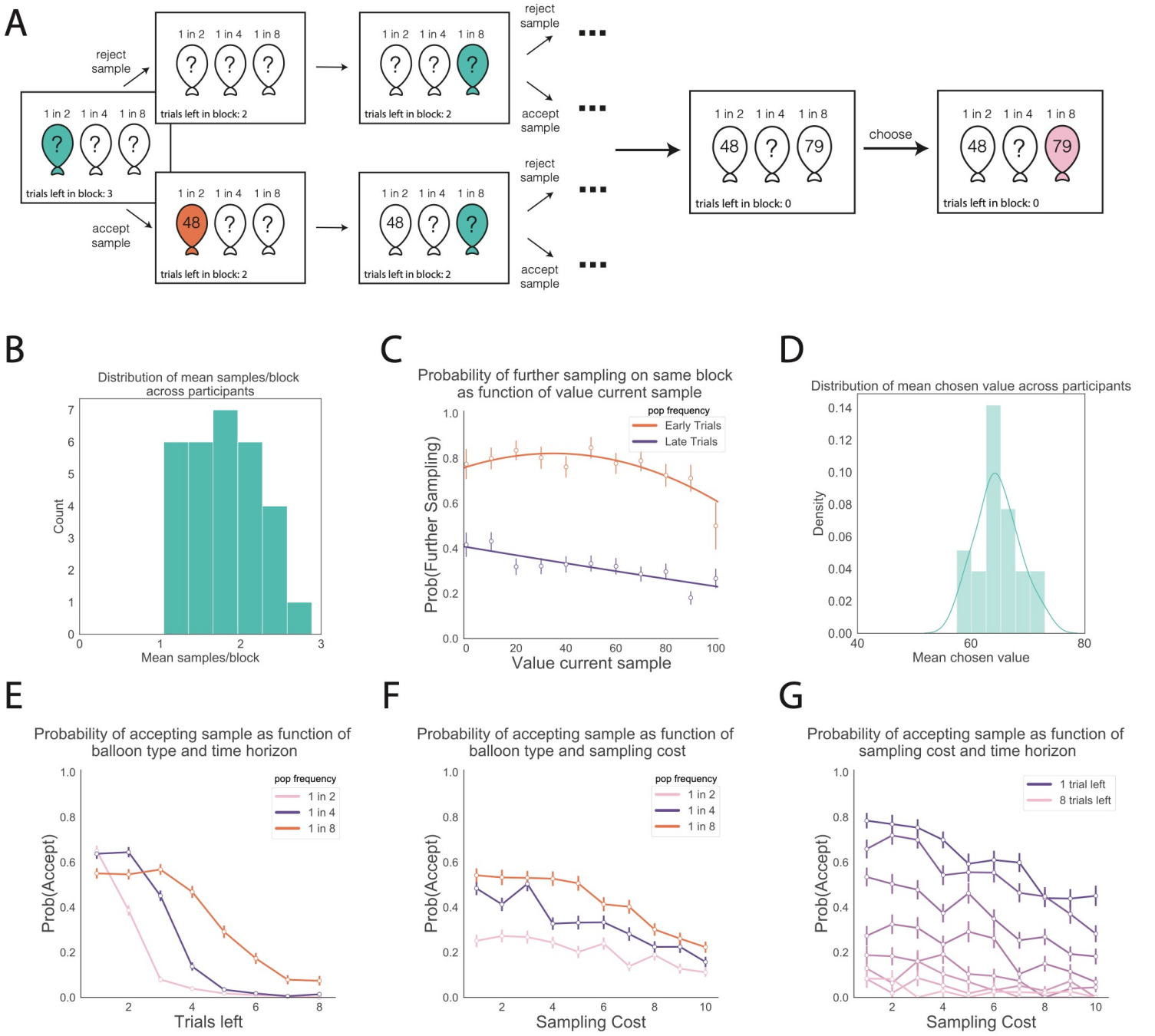


Figure 6.2 Experimental paradigm and participant behavior. (A) Task structure. Participants' goal in this task is to pick the balloon with the highest value out of three balloons on the choice round at the end of each block. Throughout a block, participants get multiple opportunities to sample one of the three balloons at a cost. On each sampling round, one of the balloons is made available for sampling. Participants could choose to accept or reject this sample using a button box. The three balloons varied in their likelihood of 'popping', meaning their value can be replaced by a new value throughout the block. Block lengths varied between 3-8 sampling rounds preceding the choice round. (B) Participants sampled on average ~1.2 balloons per block, but this varied substantially across participants. Plotted is the distribution of mean number of samples per block for each participant. (C) Participants were less likely to sample more balloons on the same block if the value of the most recent sample was high. Plotted is the probability of further samples on the same block as a function of the value of the most recent sampled balloon on the block across participant, split into late (last 3 sampling rounds before the

choice round) and early (all earlier sampling rounds) trials. (D) Participants effectively used information gained in the sampling rounds to perform above chance in the choice rounds, as seen by the mean value of the balloons they chose (at chance performance level, the mean value of chosen balloons would be 50 points). Plotted is the distribution of the mean value of chosen balloons for each participant. (E) The number of trials left until the choice round as well as the popping frequency of the balloon both affected the probability of sampling. Specifically, participants were less likely to sample the more trials were left before the choice round. However, this effect was mediated by popping frequency, such that participants were more likely to sample the balloon with the lowest popping frequency (1 in 8 rounds) earlier on, and more likely to sample the balloon with the highest popping frequency (1 in 2 rounds) close to the choice round. (F) Sampling was also affected by sampling cost, such that participants were less likely to accept expensive samples. (G) The effects of sampling cost and trials left until the choice round interact: the negative effect of sampling cost on sampling likelihood is stronger the closer participants were to the choice round. (C, E-G) Error bars are standard errors of the mean (SEM).

Participants sampled on average ~1.8 balloons per block (Figure 6.2B). They used the value of samples already made to decide whether or not to sample more information (Figure 6.2C). Specifically, the higher the value of a sampled balloon, the less likely participants were to sample more balloons on the same block ($z = -4.92$, $p < 0.001$). Furthermore, participants also effectively used information gained in the sampling rounds to improve their pay-off in the choice rounds, as the mean value of the chosen balloon was significantly higher than 50 points, which would be the expected reward when performing at chance level (Figure 6.2D; $t_{29} = 8.97$, $p < 0.001$).

All three manipulated task features – goal proximity, information reliability and sampling cost – affected sampling decisions on this task. We hypothesized that participants should be less likely to sample earlier on in the block, as information sampled on those trials was less likely to still hold true at the time of the choice round. The number of trials left until a choice round, i.e. goal proximity, indeed had a significant negative effect on sampling (Figure 6.2E; $z = -11.04$, $p < 0.001$), so participants were more likely to sample balloons the closer they were to the choice round. Furthermore, we hypothesized that participants would be less likely to sample balloons that were more likely to pop before the choice round. As predicted, popping

frequency was a negative predictor of sampling (Figure 6.2F; $z = -7.90$, $p < 0.001$). Finally, we predicted that sampling cost should also be a negative predictor of sampling, which was also the case (Figure 6.2F-G; $z = -6.97$, $p < 0.001$).

Interestingly, some of these effects of the task features on sampling decisions also interacted with each other. For example, participants were more willing to sample the balloon with the lowest popping frequency (1 in 8 sampling rounds) earlier on in the block (Figure 6.2E; interaction effect trials left and popping frequency: $z = -16.41$, $p < 0.001$). Presumably this is because the lower the popping frequency of a balloon, the less likely it was to pop between the sample and the choice round, and therefore the more likely the information gained from the sample is to still hold true at the time of the choice round. We might also expect there to be an interaction effect between popping frequency and sampling cost, such that participants would be more willing to spend more points for information that's likely to hold true for longer, i.e. from a balloon with a low popping frequency. However, this interaction was not significant (Figure 6.2F; $z = -1.00$, $p = 0.32$). Besides this, we predicted there would be a positive interaction effect between sampling cost and the number of trials left until the choice round, as we predicted participants would be willing to spend more points on sampling closer to the choice round, as information gained in these later rounds was more likely to still hold true at the choice round. However, while this trend appears to be there in Figure 6.2G, it did not reach significance ($z = 0.22$, $p = 0.83$).

As goal proximity, popping frequency and sampling cost were all seen to affect sampling behavior, we hypothesized these features to also be represented in brain activity. We performed a mass univariate analysis at the time of sampling trial onset on the fMRI data collected while participants performed the task. We included the task

features as predictors, as well as a predictor indicating whether the proposed sample on that trial would be accepted by the participant.

We found a widespread network of regions that responded to the *number of trials left* until the choice round, which we refer to as goal proximity (Figure 6.3A). These included MFC (peak at MNI coordinates [2, 32, 44] and peak Z = 6.19), parietal cortex (peak at MNI coordinates [10, -70, 54] and peak Z = 6.74), frontal pole (left FP: peak at MNI coordinates [-32, 60, 6] and peak Z = 4.76), dlPFC (left dlPFC: peak at MNI coordinates [-50, 18, 42] and peak Z = 5.17; right dlPFC: peak at MNI coordinates [52, 26, 38] and peak Z = 6.23), right anterior insula (AI; peak at MNI coordinates [32, 30, 2] and peak Z = 4.38), cerebellum (peak at MNI coordinates [-8, -76, -26] and peak Z = 4.76), and thalamus (peak at MNI coordinates [4, -20, -1] and peak Z = 4.53). Specifically, activity in these areas was decreased the more sampling trials participants were still removed from the choice round. Timeseries analysis on the averaged activity in ROIs specified from the whole-brain analysis showed that this main effect was mostly due to strongly increased activity in these regions during the last two sampling trials before the choice round (Figure 6.3B).

Even though they were seen to affect behavior (Figure 6.2), the other two task features, sampling cost and popping frequency, were not significant predictors of brain activity in this analysis.

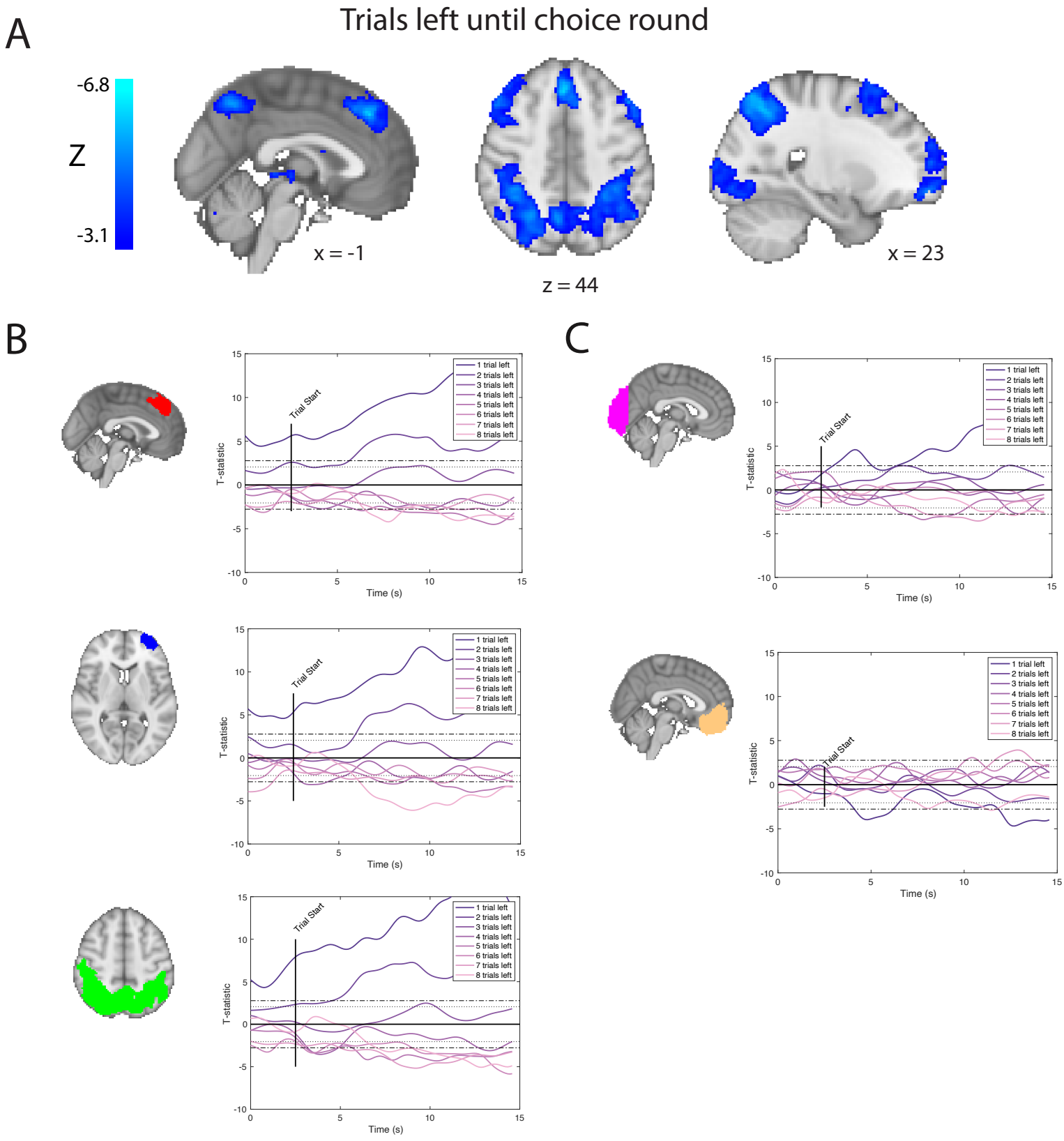


Figure 6.3 The number of sampling trials left before the choice round, i.e. goal proximity, is associated with decreased brain activity in a widespread brain network, including MFC, parietal cortex, dlPFC, and frontal pole. (A) Linear negative effect of trials left until choice in multiple brain regions in frontal and parietal cortex. Parametric maps are cluster-corrected and thresholded at z>3.1. (B) Brain activity increases in three regions of interest (ROIs), derived from the whole-brain analysis testing the effect of

goal proximity on brain activity, as a function of proximity to the choice round. Plotted are the evoked brain activity responses time-locked to trial onset (when a new balloon is highlighted and made available for sampling) for the 8 goal proximity conditions for three ROIs derived from the whole-brain analysis: MFC (red), left frontal pole (blue), and parietal cortex (green). (C) No consistent effect of number of trials left is seen on activity in occipital pole (pink) or ventromedial prefrontal cortex (beige).

We hypothesized that the strong response in the brain to goal proximity could simply be due to a general attention effect. It is possible that participants did not attend much to the task until they were close to the choice round. If this were the case, we might expect activity in other brain regions to also correlate with goal proximity, but this might not have survived cluster-correction and thresholding. Attention is known to modulate brain activity in sensory areas (Müller & Ebeling, 2008; Yantis, 2008), as well as in prefrontal regions thought to compute value during decision-making (Lim et al., 2011). Therefore we performed timeseries analyses in two control regions – primary visual cortex and ventromedial prefrontal cortex (vmPFC) – to see if activity in these regions was at all modulated by goal proximity. These areas are not known to specifically activate in response to goal-relevant information, so if we do see activity in these regions related to the number of trials left on the block, this may be due to a modulatory effect of attention. This did not appear to be the case (Figure 6.3C), which made us confident that the regions found to activate closer to the choice round were in fact representing the time horizon rather than increased attention. Another potential confound is how far along participants are in the block, which we address in a second whole-brain analysis discussed later in this chapter.

In the same whole-brain mass univariate analysis we also included a predictor indicating the upcoming sampling decision ('accept' or 'reject'). In a previous study (see **chapter 3**) we found that activity in MFC and parietal cortex significantly predicted

upcoming sampling decisions, with increased activity predicting further information sampling (Kaanders et al., 2021), and we thus hypothesized to find the same in this study. We only found one significant cluster predicting sampling decisions, which was in parietal cortex (Figure 6.4A; peak at MNI coordinates [-48, -38, 59], peak Z = 4.44), but none at MFC. However, we did find sub-threshold activity in MFC, though this was in a more anterior subregion of MFC than that found in Kaanders et al. (Figure 6.4C). There was some correlation between goal proximity and sampling (Figure 6.4B), but not so much that we expect the results of this analysis to be affected by collinearity.

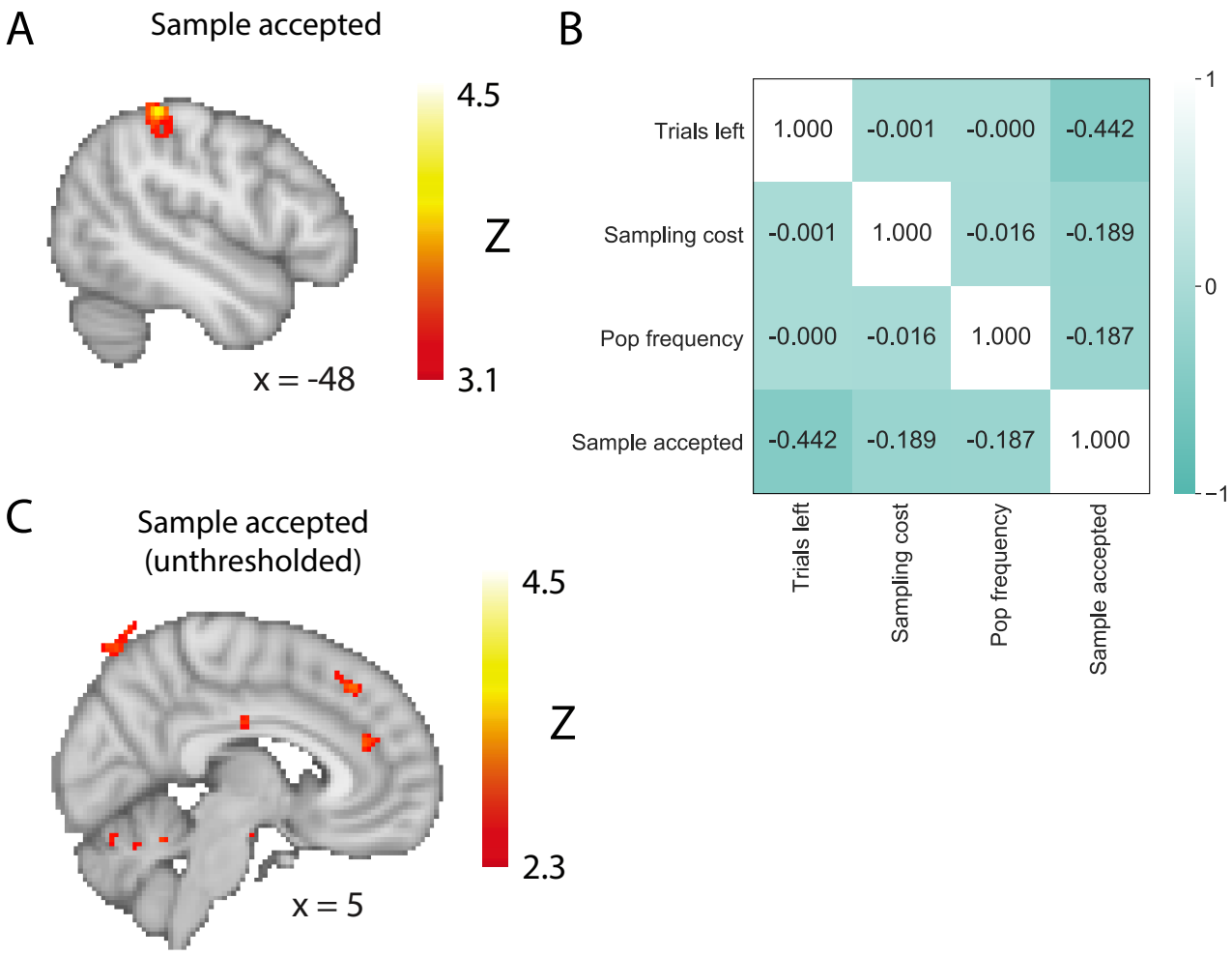


Figure 6.4 Activity in parietal cortex at trial onset predicted the upcoming sampling decision. (A) A small region within parietal cortex showed increased activity prior to decisions to sample information. Parametric map is cluster-corrected and thresholded at z>3.1. (B) The number of trials left was the strongest predictor of sampling compared to the other two manipulated task features, sampling cost and popping frequency.

The strongest predictor of brain activity at the onset of sampling trials was goal proximity, a task feature that indicates the goal-relevance of available information. We might therefore expect the same regions to respond to revelation of said information after a sample has been accepted. Indeed, in response to information revelation we find significant clusters in parietal cortex (Figure 6.5; peak at MNI coordinates [30, -62,

52], peak Z = 7.02), MFC (peak at MNI coordinates [0, 25, 46], peak Z = 6.13), dlPFC (left dlPFC: peak at MNI coordinates [-52, 10, 30], peak Z = 5.2; right dlPFC: peak at MNI coordinates [49, 41, 23], peak Z = 6.39), frontal pole (left FP: peak at MNI coordinates [-32, 58, 18], peak Z = 4.25; right FP: peak at MNI coordinates [36, 58, -4], peak Z = 4.13), orbitofrontal cortex (OFC; left OFC: peak at MNI coordinates [-26, 50, -14], peak Z = 4.27; right OFC: peak at MNI coordinates [23, 41, -17], peak Z = 3.76), caudate nucleus (peak at MNI coordinates [-16, 24, 6], peak Z = 4.61), posterior cingulate cortex (peak at MNI coordinates [4, -38, 22], peak Z = 5.00), cerebellum (peak at MNI coordinates [40, -66, -46], peak Z = 4.69), and thalamus (right thalamus: peak at MNI coordinates [16, -10, 7], peak Z = 4.13). We also included the value that was revealed as a parametric regressor in the same analysis, but only activity in visual cortex was significantly increased as a function of value (left: peak at MNI coordinates [-10, -100, 4], peak Z = 5.14; right: peak at MNI coordinates [18, -96, 14], peak Z = 5.36).

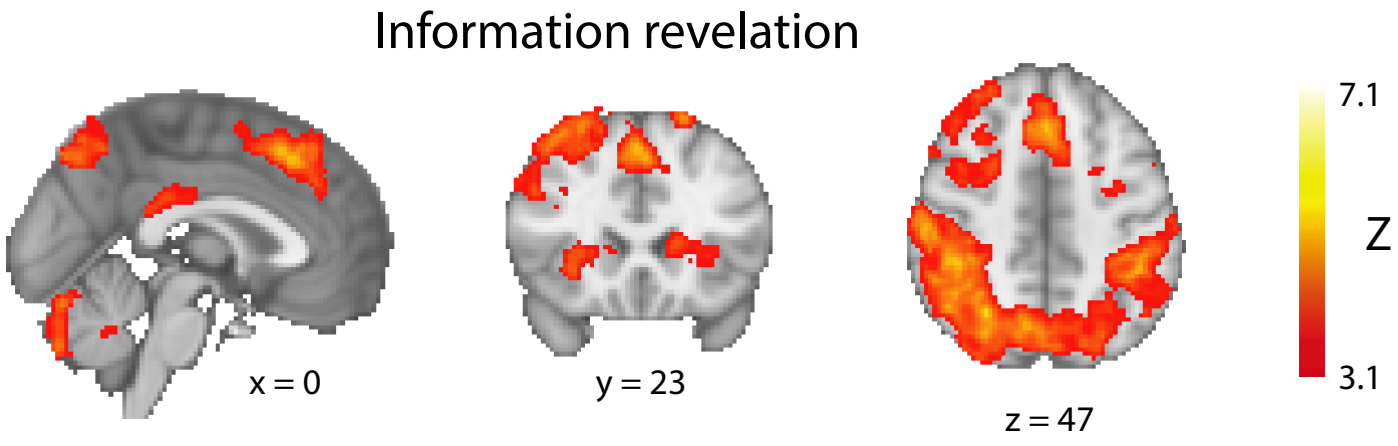


Figure 6.5 Increased activity in response to information revelation after choosing to sample in MFC, parietal cortex, dlPFC, posterior cingulate cortex, and cerebellum. Parametric maps are cluster-corrected and thresholded at z>3.1.

As the blocks were of varying lengths (between 3-8 sampling rounds), the number of trials left as a predictor is confounded by block length. For example, if the number of trials left is 3, this may be the first sampling round in some blocks, but the fifth in others. To account for this we ran a second whole-brain mass univariate analysis including the trial number within the block as a predictor. Furthermore, in our timeseries analysis we saw that activity increased across multiple regions in the last 3 sampling rounds before the choice round with respect to earlier sampling rounds. This suggest the effect of goal proximity on brain activity is non-linear. To try and better fit the model to the data we therefore included a contrast subtracting activity in response to earlier sampling rounds from that in response to the last 3 sampling rounds of the block, rather than a linear parametric regressor of the number of trials left, as previously. In this second analysis, we found that significant clusters for this contrast were much fewer than seen in the initial analysis (Figure 6.6A). Specifically, we only found two significant clusters in right parietal cortex (peaks at MNI coordinates [16, -68, 60] and [40, -46, 56], peak Z = 4.82, and Z = 4.26, respectively). Trial number within the block predicted activity in visual cortex (Figure 6.6C; left: peak at MNI coordinates [-26, -96, 0], peak Z = 5.01; right: peak at MNI coordinates [38, -80, 6], peak Z = 5.52), right dlPFC (peak at MNI coordinates [46, 12, 32], peak Z = 4.62), and right parietal cortex (peak at MNI coordinates [24, -76, 54], peak Z = 4.66).

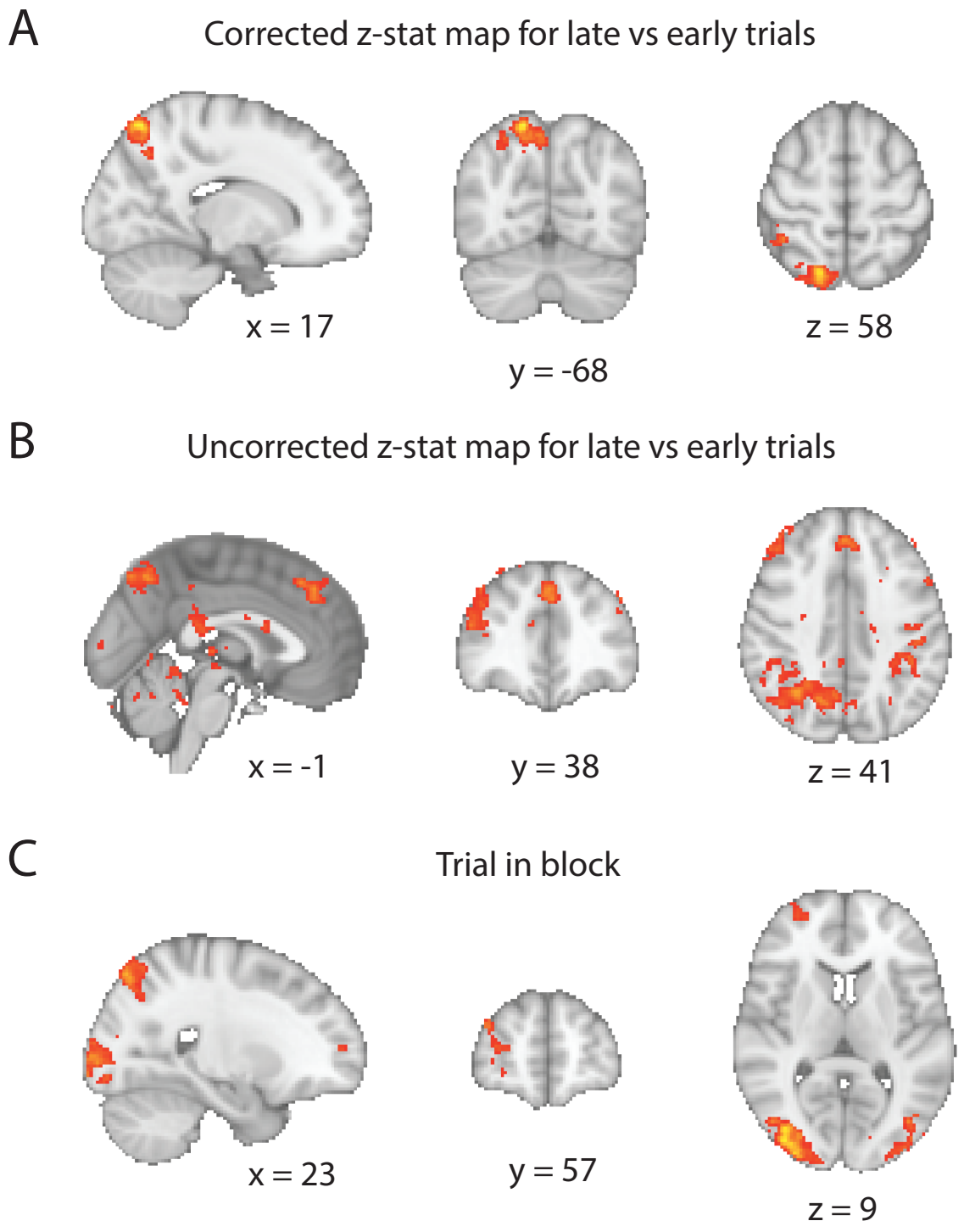


Figure 6.6 Including trial number in the whole-brain mass univariate analysis appears to remove the effect of goal proximity in most of the regions where an effect was seen before (Figure 6.2A). This may, however, only be due to statistical correction for multiple comparisons, as uncorrected activity in response to the number of trials left looks similar to that seen in the previous analysis. (A) Corrected activity in response to goal proximity as defined in a contrast comparing late (last 3 sampling rounds) to early (all other sampling rounds) trials. (B) Uncorrected activity in response to goal proximity. (C) Activity in response to trial number within the block.

However, a recent study has shown that statistical correction of multiple comparisons can introduce large variability in results in the form of thresholded z-statistic maps, which are typically reported in fMRI papers (Botvinik-Nezer et al., 2020). Looking at the unthresholded z-statistic map for the goal proximity contrast in our second mass univariate analysis, it can be seen the same set of brain regions is active (Figure 6.6B). To test whether thresholding could be producing the different responses to goal proximity seen in the two thresholded maps, we correlated the parameter estimates at each voxel for the goal proximity contrast of each analysis. These were significantly correlated ($r = 0.77$, $p < 0.001$), meaning that the same pattern of activity is observed in response to goal proximity in both analyses. This suggests that the activity seen in these regions cannot be explained away by trial number within the block.

6.3 Discussion

Here we have presented a novel decision-making task, in which human participants could make active information sampling decisions leading up to a choice, thereby separating decisions to seek out new information from reward-maximizing choice. In this task, we manipulated goal proximity, information reliability and sampling cost and found that all three affected sampling behavior as predicted. We have shown that a large network made up of multiple frontal and parietal brain regions encodes goal proximity, and that this activity is unlikely to be explained by a general attention increase throughout the task. The same set of regions activated in response to the revelation of information after decisions to sample. Besides this, we found activity in a region in parietal cortex predicts upcoming decisions to sample information, as well as

sub-threshold activity in MFC, which partly replicates our previous study on the neural signatures of information sampling (Kaanders et al., 2021; see **chapter 3**).

Brain activity was seen to increase the closer participants got to the choice round at the end of each block in MFC, parietal cortex, frontal pole, dlPFC and AI, among other regions. MFC activity has been shown to correlate with that of evidence accumulators in the drift diffusion model (DDM) used to model behavior in binary forced choice tasks (Gluth et al., 2012; Hare et al., 2011; Pisauro et al., 2017; Rodriguez et al., 2015; Venkatraman et al., 2009; Zhang et al., 2012), which led to our hypothesis that MFC could be well-positioned to drive sequential sampling decisions leading up to a choice, as well as the final reward-maximizing choice (Kaanders et al., 2021). It is therefore a possibility that in the current task, MFC is tracking goal proximity as evidence towards sampling more information. On the other hand, if this were the case, we would expect MFC activity to equally reflect other task features that provide evidence in favor or against sampling more information, such as sampling cost and information reliability, which we do not find.

We observed increased activity in response to goal proximity in a large parietal region. However, peak activation was observed in a subregion of the superior parietal lobule thought to be the human homologue of primate LIP (Mars et al., 2011; Rushworth et al., 2001). The large parietal cluster we found to represent goal proximity also included parietal regions shown to be involved in object grasping (Binkofski et al., 1999; Culham et al., 2003; Grefkes & Fink, 2005; Mars et al., 2011). This is of note, as besides visual attention allocation, reaching for and grasping objects is probably the second-most prominent way in which primates sample information from their environment. Therefore, parietal cortex is another region we might expect to be

involved in decisions to sample information through its role in driving attention using both eye and hand movements. Indeed, single-cell recordings of parietal cortex have revealed that parietal neurons encode reductions in uncertainty in advance of decisions to sample information (Foley et al., 2017; Horan et al., 2019). Goal proximity may mediate activity in this region as attention reallocation to enable the sampling of new information should be more urgent the closer an agent is to having to make a reward-maximizing choice.

Other regions found to be sensitive to goal proximity were frontal pole, dlPFC, and anterior insula (AI). Frontal pole has been shown to track the value of alternative choice options with respect to a default 'best' option in a multi-alternative forced choice task (Boorman et al., 2009), while complex planning behavior seen in primates is thought to critically depend on the highly specialized function of dlPFC (Burgess, 2000; Goel & Grafman, 1994; Hunt et al., 2021). Finally, AI activity has been associated with exploration and information sampling (Blanchard & Gershman, 2018; Kaanders et al., 2021).

The regions we observe to be sensitive to goal proximity in our task are regions that are frequently found to co-activate in decision-making tasks (Akam et al., 2021; Domenech et al., 2020; Isoda & Hikosaka, 2007; Karlsson et al., 2012; Kolling et al., 2016; Molenberghs et al., 2007; O'Reilly et al., 2013; Schuck et al., 2015; Kobayashi et al., 2021). Given their specialized roles in attention, action selection and decision-making discussed above, it is possible that activity in all these regions is sensitive to goal proximity because this determines how strongly an agent should be driven to sample information and plan for the upcoming choice. Mediating activity as a function

of goal proximity in this way might enable sampling behavior to flexibly change in response to current and upcoming goals and demands.

We did not find significant brain activation in response to the other task features that were shown to predict sampling decisions: popping frequency of the balloon and sampling cost. A possible explanation for this is that our mass univariate analysis does not account for the fact that the extent to which participants used each of the task features varied substantially (see **chapters 4 and 5**). For example, we may not see a representation of sampling cost in the brain, because we are currently equally weighting the brain activity in response to cost from each participant, regardless of how much they used cost to guide their sampling decisions. A next step in analyzing this data would be to perform a higher-level analysis that takes into account how well each task feature predicted sampling decisions as a covariate.

In summary, we have shown that activity in a widespread network of frontal and parietal brain regions is mediated by goal proximity during active information sampling decisions. This suggests that multiple regions known to be involved in information sampling and decision-making may guide flexible behavior change as goals come and go.

6.4 Methods

30 healthy human participants (aged 18-65) attended two study sessions: one online behavioral training session (1.5hr) and one fMRI session (2hr 15min).

6.4.1 Experimental design

In the main task, participants were asked to choose between three balloons of different values between 1-100 points (Figure 6.1A). At the start of a block, these values were hidden and participants were given between 3-8 sampling trials to sample the balloons. However, each of the balloons could 'pop' from one trial to the next and change to a new value. Specifically, each of the balloons was associated with a probability of 'popping' on any given trial (50%, 25%, and 12.5%, respectively), and this probability was displayed on the screen throughout the task.

Each block started with the screen displaying the number of sampling trials participants would have on that block for 5s. Then the three balloons were displayed on the screen, each accompanied by an empty value bar with a question mark, indicating that the balloon had not yet been sampled and its value was thus unknown. On each sampling trial, one of the three balloons was chosen at random and highlighted, and participants were asked whether they wanted to sample that balloon's value at a cost of between 1-10 points (randomly sampled from uniform distribution on each trial). This sampling cost was shown next to the highlighted balloon by a red bar. Participants could choose to either accept or reject this sample by pressing the corresponding buttons. This decision was followed by a 3-7s jitter. If the sample was accepted, a gold bar next to the balloon was then revealed to show that balloon's current value. If the sample was rejected, participants moved onto the next sampling

trial after a 1s delay. Throughout the sampling trials, the value bars next to the balloons continued to display the value of the balloon when it was last sampled. So in order to make sure that value was still the balloon's actual value, participants would have to sample it again on a future trial. The number of trials that had passed since a balloon was last sampled was displayed in the form of dots next to the balloons.

After the sampling trials the block would end with a single choice trial in which participants had to choose the balloon they thought contained the most points by using the corresponding buttons. The number of points won on that block equaled the number of points in the participant's balloon of choice minus however many points they spent sampling the balloons during the sampling trials. The values of all the balloons were revealed after participants made a choice and their reward on that block was displayed on the screen. All the trials were separated by a 1s inter-stimulus interval (ISI). Participants were only invited to take part in an fMRI session if on the online training session they chose to accept a sample between 20-80% of the time. This experiment was realized using PsychoPy (Peirce et al., 2019).

In the fMRI session, participants did 20 practice trials and five blocks of the main task (a total of approximately 250 trials; this varied per participant as the trial schedule was randomized for each participant).

At the end of the experiment, two blocks from the main task and one from the online training session were chosen at random and participants received the average number of points gained on those two blocks multiplied by 10p as a bonus payment. They also earned a £35 show-up fee for attending both study sessions.

6.4.2 fMRI data collection

Whole-brain fMRI measurements were made using a Siemens Prisma 3T scanner with a 2 x 2 x 2mm voxel size, repetition time (TR) = 1.235s, echo time (TE) = 20ms, flip angle = 65° with an axial orientation angled to AC-PC using a 64-channel head coil. The sequence used was MB3 PAT2. Participants performed five blocks of the main task inside the scanner with short breaks in between. For each new block a new run was initiated. Given that the task was not speeded and the number of trials varied per run, the runs were of variable length and so were the number of volumes collected per run. The average number of volumes collected per participant was 2517.3. T1-weighted structural images were obtained using an MPRAGE sequence with 1 x 1 x 1mm voxel size, on a 174x192x192 grid, TE = 3.97ms, TR = 1.9s. A field map with dual echo-time images was also acquired (TE=7.38ms, whole-brain coverage, 2.4 x 2.4 x 2.4 mm voxel size).

6.4.3 fMRI data analysis

fMRI analysis was carried out using the FMRIB Software Library (FSL). The pre-processing performed on the data was the following: motion correction using MCFLIRT (Jenkinson et al., 2002); non-brain removal using BET (S. M. Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 5mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; highpass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma=50.0s). Registration to high resolution structural and/or standard space images was carried out using FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001).

6.4.4 Behavioral data analysis

The effects of the three task features, goal proximity, popping frequency and sampling cost, were analysed using hierarchical logistic regression predicting whether or not participants 'accepted' or 'rejected' to sample on each trial.

6.4.5 Mass univariate analysis

GLMs were fit to the individual subject data including regressors describing the number of sampling rounds left before the choice round, sampling cost, popping frequency, sampling decision, button presses, value of the sampled balloon (if participants chose to sample information on that trial), and reward at trial onset and at information revelation (if participants chose to sample on that trial). All regressors were modelled as 2s boxcar functions, to match the duration for which the cue was presented. We report mass univariate results at the group-level using whole-brain family wise error (FWE) corrected statistical significance and a cluster significance threshold of $Z>3.1$ and $p<0.05$. We also present some uncorrected statistical maps, at thresholds of either $Z>3.1$ or $Z>2.3$.

Chapter 7: General Discussion

All our important decisions are determined by the information we have encountered or sought out during our lives up to that point. Despite this fact, much of decision-making research has treated human participants as passive recipients of information to simplify the decisions being studied. This approach has greatly improved our understanding of decision-making but has now begun to reach its limit in how much more it can teach us about human behavior. Luckily, many researchers are exploring ways in which to study the full complexity of information search and decision-making in humans. In this thesis I have made my own attempt at studying information search in a way that appreciates the active way in which humans seek out and use information to guide choices.

7.1 What have we learned?

Firstly, I have shown that MFC and subregions of parietal cortex may play a special role in driving information sampling to guide choice, using an existing decision-making task with active information sampling (**chapter 3**). I then developed a novel information sampling task in which participants could actively sample information, decisions to sample were separate from decisions to maximize reward, and task features could be manipulated to vary the goal-relevance of available information (**chapter 4**). This revealed strong heterogeneity in sampling strategies across participants. We used one version of this task to investigate information sampling as a function of transdiagnostic psychiatric dimensions and found no convincing effects (**chapter 5**). Finally, using another version of the same task, we have shown that a

multitude of brain regions, that consistently co-activate in response to various decision-making tasks, represent features of the environment that indicate how valuable information sampling is likely to be in light of future goals (**chapter 6**).

Allowing participants to actively sample information in our tasks has revealed that brain regions often found to activate in decision-making tasks also activate prior to decisions to sample information, making it likely regions involved in tracking decision evidence also drive sequential decisions to gather more evidence. We find parietal cortex and MFC activity are predictive of upcoming decisions to sample information in our first MRI study (**chapter 3**), but MFC activity does not reach significance in the second MRI study (**chapter 6**). However, the pattern of brain activity predictive of decisions to sample is similar across studies, suggesting the specific task and combination of co-regressors included in the GLM used to analyze the data strongly affects the results.

Indeed, in **chapter 3** we found that MFC activity encoded choice difficulty, but this effect disappeared when a regressor describing upcoming sampling choices was included, and MFC activity was instead found to be predictive of future sampling. In **chapter 6**, we instead found the strongest effect in the brain in response to goal proximity, one of three manipulated task features found to affect sampling behavior, and no effect of the other two. Besides this, brain activity predictive of upcoming sampling decisions was modest compared to that observed in **chapter 3**. An interesting follow-up study to this thesis would therefore be to design a task in which different drivers of sampling and decisions to sample can be better dissociated, for example by temporally separating the revelation of evidence known to affect sampling behavior.

In **chapter 5** we tried to link sampling behavior to transdiagnostic psychiatric dimensions but were not able to replicate an effect of compulsivity on the integration of task features in a confirmation sample. As effects of compulsivity on information sampling have been observed before in OCD patients and individuals scoring highly on compulsivity questionnaires, it is possible that effects of compulsivity on behavior are non-linear. In other words, behavior may only be affected when compulsivity is very high, and effects might thus only be observed in clinical participant samples. To investigate this further, it would be useful if future studies used the same information gathering task to both specifically test clinical samples and large online samples.

7.2 What should be addressed in future work?

The wide-ranging terminology used to describe information-seeking behaviors or curiosity reflects its many potential drivers. Many studies have shown humans and animals have a preference for new information, and seek it out to satisfy upcoming goals (Blanchard & Gershman, 2018; Cohen et al., 2007; Kakade & Dayan, 2002; Kolling et al., 2012; Navarro et al., 2016; Newell & Simon, 1972; Padoa-Schioppa, 2011; Shadlen & Shohamy, 2016). But agents are also known to sample information when it is not immediately clear how the information might be useful to guide future behavior (see **chapter 1**; Berlyne, 1955; Butler, 1953, 1957a, 1957b; Ennaceur & Delacour, 1988; Hughes, 1965, 1968, 2007; Kang et al., 2009; Nicki, 1970; Reed et al., 1996; Stolba & Wood-Gush, 1980; Wood-Gush & Vestergaard, 1991). The latter has been explained by the hypothesized benefit of sampling information that does not immediately lead to a reward in simulating long action sequences leading to reward that can only be discovered through random exploration (Bellemare et al., 2016).

Indeed, this type of information search greatly improves the performance of machine learning agents on hard exploration problems (Alet et al., 2020; Badia et al., 2020; Brafman & Tennenholtz, 2001; Burda et al., 2018; Ecoffet et al., 2019; Ermolov & Sebe, 2020; Ng et al., 1999; Raileanu & Rocktäschel, 2020; Sekar et al., 2020; Twomey & Westermann, 2018; Zheng et al., 2020). However, to fully understand the wide variety of rich information seeking behavior we observe in humans, we must begin to appreciate the wide range of potential rewards humans are hoping to find. In the study of decision-making within neuroscience, rewards have been typically limited to money or food, as they are easy to administer in an experimental context. It is likely, though, that humans search information to discover paths to a much broader range of rewards, such as improved social relationships and status, aesthetic pleasure, or better health. Furthermore, explaining all information search as intrinsically rewarding ignores the finding that humans in fact avoid some types of information, in particular information about bad outcomes (Charpentier et al., 2018; N. Karlsson et al., 2009; Lerman et al., 1998). This suggests that the rewarding nature of information might be increased for information that leads the agent to hold more positive beliefs (Golman & Loewenstein, 2018). Finally, we know that humans use their prior knowledge of the world to judge what information sources are likely to be useful, rather than treating all new information as equally rewarding (Dubey et al., 2018; Spelke & Kinzler, 2007; Tsividis et al., 2017). To be able to investigate information search in its full complexity, experimental tasks must be developed that reflect this complex reward space.

In the current work it has been difficult to dissociate how the computations in different brain regions contribute to the process of sampling information and how they are used to make decisions. To do so, novel tasks must be designed that can better

dissociate different drivers of information sampling. Besides this, brain data must be collected using a variety of imaging modalities to investigate both the spatial and temporal properties of brain regions involved in these tasks. While fMRI studies can help us understand what brain areas are likely to play a role in information sampling, to dissociate their specific computations causal studies, using pharmacology or virtual lesions, will be essential.

Furthermore, from collecting behavioral data on many task versions during task development and executing a large-scale online study it became apparent that there is large heterogeneity in people's information sampling strategies. It was not the focus of the work presented here to explore this heterogeneity, but characterizing these individual differences and learning why they exist is most likely essential to beginning to understand the wide range of human behavior observed in real life. We need to develop models that account for this variation and can predict upcoming decisions for a variety of behavioral strategies that people might employ. The field of computational psychiatry is making some strides in linking such behavioral variation to psychiatric symptoms (Gillan et al., 2016; Hauser et al., 2017), but as we have shown in our online study this approach has mixed results studying variation in subclinical populations, i.e. the majority of humans.

7.3 What are the implications of this work?

As many psychiatric disorders are characterized by altered decision-making in some way, understanding how this may be driven by specific information sampling behavior can improve our understanding of the mechanisms behind these disorders, and to find treatments that may improve symptoms by targeting such behaviors. For

example, it has been suggested antidepressants may not directly boost mood, but in fact alter emotional processing such that information is interpreted more positively, eventually leading to more positive beliefs and better mood in the long-term (Harmer et al., 2009). It is possible that many symptoms relating to beliefs and decisions in mental illness arise through disordered information sampling or processing in a similar way. If so, it may be better to develop treatments that target how information is sought out and weighted by the individual, rather than directly targeting the symptom itself.

7.4 Conclusion

I have used active information sampling tasks to separate sampling decisions from reward-maximizing choice and have shown the involvement of multiple prefrontal and parietal brain regions in the making of decisions to gather information, but it is clear more sophisticated experimental tasks and methodologies are needed to investigate how different regions interact and contribute to these decisions. I have also shown the great heterogeneity present in sampling behavior in a novel information sampling task, which is a promising avenue for future study, as this behavioral variety seen in experimental settings may be predictive of the great range of sampling behaviors present in humans navigating real life.

References

Akam, T., Rodrigues-Vaz, I., Marcelo, I., Zhang, X., Pereira, M., Oliveira, R. F., Dayan, P., & Costa, R. M. (2021). The Anterior Cingulate Cortex Predicts Future States to Mediate Model-Based Action Selection. *Neuron*, *109*(1), 149-163.e7. https://doi.org/10.1016/j.neuron.2020.10.013

Alet, F., Schneider, M. F., Lozano-Perez, T., & Kaelbling, L. P. (2020). Meta-learning curiosity algorithms. *ArXiv:2003.05325 [Cs, Stat]*. http://arxiv.org/abs/2003.05325

APA. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Association.

Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex*, *80*, 125–140. https://doi.org/10.1016/j.cortex.2015.11.024

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, *7*(5), 358–366. https://doi.org/10.1038/nrn1888

Badia, A. P., Sprechmann, P., Vitvitskyi, A., Guo, D., Piot, B., Kapturowski, S., Tieleman, O., Arjovsky, M., Pritzel, A., Bolt, A., & Blundell, C. (2020). Never Give Up: Learning Directed Exploration Strategies. *ArXiv:2002.06038 [Cs, Stat]*. http://arxiv.org/abs/2002.06038

Banca, P., Vestergaard, M. D., Rankov, V., Baek, K., Mitchell, S., Lapa, T., Castelo-Branco, M., & Voon, V. (2015). Evidence Accumulation in Obsessive-Compulsive Disorder: The Role of Uncertainty and Monetary Reward on

Perceptual Decision-Making Thresholds. *Neuropsychopharmacology*, *40*(5), 1192–1202. https://doi.org/10.1038/npp.2014.303

Bardo, M., Donohew, R., & Harrington, N. (1996). Psychobiology of novelty seeking and drug seeking behavior. *Behavioural Brain Research*, *77*, 23–43.

Barnett, S. A. (1958). Exploratory behaviour. *British Journal of Psychology*, *49*(4), 289–310. https://doi.org/10.1111/j.2044-8295.1958.tb00667.x

Barron, H. C., Garvert, M. M., & Behrens, T. E. J. (2016). Repetition suppression: A means to index neural representations using BOLD? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1705), 20150355. https://doi.org/10.1098/rstb.2015.0355

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences*, *107*(50), 21767–21772. https://doi.org/10.1073/pnas.0908104107

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious Mixed Models. *ArXiv*. http://arxiv.org/abs/1506.04967

Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2003). General multilevel linear modeling for group analysis in FMRI. *NeuroImage*, *20*(2), 1052–1063. https://doi.org/10.1016/S1053-8119(03)00435-X

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, *100*(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Bellemare, M., Srinivasan, S., Ostrovski, G., Schaul, T., Saxton, D., & Munos, R. (2016). Unifying Count-Based Exploration and Intrinsic Motivation. *NIPS*, 1471–1479.

Belujon, P., & Grace, A. A. (2017). Dopamine System Dysregulation in Major Depressive Disorders. *International Journal of Neuropsychopharmacology*, *20*(12), 1036–1046. https://doi.org/10.1093/ijnp/pyx056

Berlyne, D. (1955). The arousal and satiation of perceptual curiosity in the rat. *Journal of Comparative and Physiological Psychology*, *48*(4), 238–246.

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behavior. *British Journal of Psychology*, *41*(1–2), 68–80.

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology. General Section*, *45*(3), 180–191. https://doi.org/10.1111/j.2044-8295.1954.tb01243.x

Berlyne, D. E. (1966). Curiosity and Exploration. *Science*, *153*, 25–33.

Bhandari, A., Benna, M., Rigotti, M., Fusi, S., & Badre, D. (2019). Measuring prefrontal representational geometry: FMRI adaptation vs pattern analysis. *2019 Conference on Cognitive Computational Neuroscience*. 2019 Conference on Cognitive Computational Neuroscience, Berlin, Germany. https://doi.org/10.32470/CCN.2019.1162-0

Bhandari, A., Gagne, C., & Badre, D. (2018). Just above Chance: Is It Harder to Decode Information from Prefrontal Cortex Hemodynamic Activity Patterns?

Journal of Cognitive Neuroscience, *30*(10), 1473–1498. https://doi.org/10.1162/jocn_a_01291

Binkofski, F., Buccino, G., Stephan, K. M., Rizzolatti, G., Seitz, R. J., & Freund, H.-J. (1999). A parieto-premotor network for object manipulation: Evidence from neuroimaging. *Experimental Brain Research*, *128*(1–2), 210–213. https://doi.org/10.1007/s002210050838

Blanchard, T. C., & Gershman, S. J. (2018). Pure correlates of exploration and exploitation in the human brain. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(1), 117–126. https://doi.org/10.3758/s13415-017-0556-2

Bogerts, B., Häntsch, J., & Herzer, M. (1983). A morphometric study of the dopamine-containing cell groups in the mesencephalon of normals, Parkinson patients, and schizophrenics. *Biological Psychiatry*, *18*(9), 951–969.

Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. *Neuron*, *62*(5), 733–743. https://doi.org/10.1016/j.neuron.2009.05.014

Boschen, M. J., & Vuksanovic, D. (2007). Deteriorating memory confidence, responsibility perceptions and repeated checking: Comparisons in OCD and control samples. *Behaviour Research and Therapy*, *45*(9), 2098–2109. https://doi.org/10.1016/j.brat.2007.03.009

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari,

M., Bault, N., Beaton, D., Beitner, J., & Benoit, R. G. (2020). *Variability in the analysis of a single neuroimaging dataset by many teams*. 37.

Bradberry, C. W., Gruen, R. J., Berridge, C. W., & Roth, R. H. (1991). Individual differences in behavioral measures: Correlations with nucleus accumbens dopamine measured by microdialysis. *Pharmacology Biochemistry and Behavior*, *39*(4), 877–882. https://doi.org/10.1016/0091-3057(91)90047-6

Brafman, R. I., & Tennenholtz, M. (2001). R-MAX - A general polynomial time algorithm for near-optimal reinforcement learning. *IJCAI International Joint Conference on Artificial Intelligence*, 953–958.

Broniatowski, D. A., Jamison, A. M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S. C., & Dredze, M. (2018). Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. *American Journal of Public Health*, *108*(10), 1378–1384. https://doi.org/10.2105/AJPH.2018.304567

Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T., & Efros, A. A. (2018). Large-Scale Study of Curiosity-Driven Learning. *ArXiv:1808.04355 [Cs, Stat]*. http://arxiv.org/abs/1808.04355

Burgess, P. W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research*, *63*, 279–288.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*(3), 432–459. https://doi.org/10.1037/0033-295X.100.3.432

Butler, R. (1953). Discrimination learning by rhesus monkeys to visual-exploration motivation. *Journal of Comparative and Physiological Psychology*, *46*(2), 95–98.

Butler, R. (1957a). Discrimination learning by rhesus monkeys to auditory incentives. *Journal of Comparative and Physiological Psychology*, *50*(3), 239–241.

Butler, R. (1957b). The effect of deprivation of visual incentives on visual exploration motivation in monkeys. *Journal of Comparative and Physiological Psychology*, *50*(2), 177–179.

Buzsáki, G. (2015). Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus*, *25*(10), 1073–1188. https://doi.org/10.1002/hipo.22488

Callaway, F., Rangel, A., & Griffiths, T. L. (2021). Fixation patterns in simple choice reflect optimal information sampling. *PLOS Computational Biology*, *17*(3), e1008863. https://doi.org/10.1371/journal.pcbi.1008863

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, *143*(4), 1476–1488. https://doi.org/10.1037/a0035813

Chamberlain, S. R., Fineberg, N. A., Blackwell, A. D., Clark, L., Robbins, T. W., & Sahakian, B. J. (2007). A neuropsychological comparison of obsessive–compulsive disorder and trichotillomania. *Neuropsychologia*, *45*(4), 654–662. https://doi.org/10.1016/j.neuropsychologia.2006.07.016

Chambon, V., Théro, H., Vidal, M., Vandendriessche, H., Haggard, P., & Palminteri, S. (2020). Information about action outcomes differentially affects learning

from self-determined versus imposed choices. *Nature Human Behaviour*, *4*(10), 1067–1079. https://doi.org/10.1038/s41562-020-0919-5

Charpentier, C. J., Bromberg-Martin, E. S., & Sharot, T. (2018). Valuation of knowledge and ignorance in mesolimbic reward circuitry. *Proceedings of the National Academy of Sciences*, *115*(31), E7255–E7264. https://doi.org/10.1073/pnas.1800547115

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*(3), 455–462. https://doi.org/10.1038/nn.3635

Cloninger, C. R. (1987). A Systematic Method for Clinical Description and Classification of Personality Variants: A Proposal. *Archives of General Psychiatry*, *44*(6), 573. https://doi.org/10.1001/archpsyc.1987.01800180093014

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1481), 933–942. https://doi.org/10.1098/rstb.2007.2098

Culham, J. C., Danckert, S. L., Souza, J. F. X. D., Gati, J. S., Menon, R. S., & Goodale, M. A. (2003). Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Experimental Brain Research*, *153*(2), 180–189. https://doi.org/10.1007/s00221-003-1591-5

Cuthbert, B. N., & Kozak, M. J. (2013). Constructing constructs for psychopathology: The NIMH research domain criteria. *Journal of Abnormal Psychology*, *122*(3), 928–937. https://doi.org/10.1037/a0034028

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879. https://doi.org/10.1038/nature04766

Der-Avakian, A., & Markou, A. (2012). The neurobiology of anhedonia and other reward-related deficits. *Trends in Neurosciences*, *35*(1), 68–77. https://doi.org/10.1016/j.tins.2011.11.005

Deroche, V., Piazza, P. V., Le Moal, M., & Simon, H. (1993). Individual differences in the psychomotor effects of morphine are predicted by reactivity to novelty and influenced by corticosterone secretion. *Brain Research*, *623*(2), 341–344. https://doi.org/10.1016/0006-8993(93)91451-W

Diba, K., & Buzsáki, G. (2007). Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, *10*(10), 1241–1242. https://doi.org/10.1038/nn1961

Domenech, P., Rheims, S., & Koechlin, E. (2020). Neural mechanisms resolving exploitation- exploration dilemmas in the medial prefrontal cortex. *Science*, *369*(1076), 11.

Dong, M.-X., Chen, G.-H., & Hu, L. (2020). Dopaminergic System Alteration in Anxiety and Compulsive Disorders: A Systematic Review of Neuroimaging Studies. *Frontiers in Neuroscience*, *14*, 608520. https://doi.org/10.3389/fnins.2020.608520

Duan, H., Fernández, G., van Dongen, E., & Kohn, N. (2020). The effect of intrinsic and extrinsic motivation on memory formation: Insight from behavioral and imaging study. *Brain Structure and Function*, *225*(5), 1561–1574. https://doi.org/10.1007/s00429-020-02074-x

Dubey, R., Agrawal, P., Pathak, D., Griffiths, T. L., & Efros, A. A. (2018). Investigating Human Priors for Playing Video Games. *ArXiv:1802.10217 [Cs]*. http://arxiv.org/abs/1802.10217

Dubois, J., de Berker, A. O., & Tsao, D. Y. (2015). Single-Unit Recordings in the Macaque Face Patch System Reveal Limitations of fMRI MVPA. *Journal of Neuroscience*, *35*(6), 2791–2802. https://doi.org/10.1523/JNEUROSCI.4037-14.2015

Duncan, J. (2013). The Structure of Cognition: Attentional Episodes in Mind and Brain. *Neuron*, *80*(1), 35–50. https://doi.org/10.1016/j.neuron.2013.09.015

Ecoffet, A., Huizinga, J., Lehman, J., Stanley, K. O., & Clune, J. (2019). Go-Explore: A New Approach for Hard-Exploration Problems. *ArXiv:1901.10995 [Cs, Stat]*. http://arxiv.org/abs/1901.10995

Ekelund, J., Lichtermann, D., Järvelin, M.-R., & Peltonen, L. (1999). Association Between Novelty Seeking and the Type 4 Dopamine Receptor Gene in a Large Finnish Cohort Sample. *Am J Psychiatry*, *156*(9), 1453–1455.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, *113*(28), 7900–7905. https://doi.org/10.1073/pnas.1602413113

Ennaceur, A., & Delacour, J. (1988). A new one-trial test for neurobiological studies of memory in rats. 1: Behavioral data. *Behavioural Brain Research*, *31*(1), 47–59. https://doi.org/10.1016/0166-4328(88)90157-X

Erhan, C., Bulut, G. Ç., Gökçe, S., Ozbas, D., Turkakin, E., Dursun, O. B., Yazgan, Y., & Balcı, F. (2017). Disrupted latent decision processes in medication-free

pediatric OCD patients. *Journal of Affective Disorders*, *207*, 32–37. https://doi.org/10.1016/j.jad.2016.09.011

Ermolov, A., & Sebe, N. (2020). Latent World Models For Intrinsically Motivated Exploration. *ArXiv:2010.02302 [Cs, Stat]*. http://arxiv.org/abs/2010.02302

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. *Nature Neuroscience*, *8*(11), 1481–1489. https://doi.org/10.1038/nn1579

Fear, C. F., & Healy, D. (1997). Probabilistic reasoning in obsessive–compulsive and delusional disorders. *Psychological Medicine*, *27*(1), 199–208. https://doi.org/10.1017/S0033291796004175

Filimon, F., Nelson, J. D., Sejnowski, T. J., Sereno, M. I., & Cottrell, G. W. (2020). The ventral striatum dissociates information expectation, reward anticipation, and reward receipt. *Proceedings of the National Academy of Sciences*, *117*(26), 15200–15208. https://doi.org/10.1073/pnas.1911778117

Fine, C., Gardner, M., Craigie, J., & Gold, I. (2007). Hopping, skipping or jumping to conclusions? Clarifying the role of the JTC bias in delusions. *Cognitive Neuropsychiatry*, *12*(1), 46–77. https://doi.org/10.1080/13546800600750597

Flesch, T., Juechems, K., Dumbalska, T., Saxe, A., & Summerfield, C. (2021). *Rich and lazy learning of task representations in brains and neural networks* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.04.23.441128

Foa, E., Huppert, J., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. (2002). The obsessive-compulsive inventory: Development and validation of a short version. *Psychological Assessment*, *14*, 485–496.

Foley, N. C., Kelly, S. P., Mhatre, H., Lopes, M., & Gottlieb, J. (2017). Parietal neurons encode expected gains in instrumental information. *Proceedings of the National Academy of Sciences*, *114*(16), E3315–E3323. https://doi.org/10.1073/pnas.1613844114

Folke, T., Jacobsen, C., Fleming, S. M., & Martino, B. D. (2017). Explicit representation of confidence informs future 2 value-based decisions. *Nature Human Behaviour*, *1*(1), 1–8.

Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology*, *37*, 66–74. https://doi.org/10.1016/j.conb.2016.01.010

Garner, D., Olmsted, M., Bohr, Y., & Garfinkel, P. (1982). The eating attitudes test: Psychometric features and clinical correlates. *Psychological Medicine*, *12*, 871–878.

Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *ELife*, *5*, e11305. https://doi.org/10.7554/eLife.11305

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., & de Wit, S. (2011). Disruption in the Balance Between Goal-Directed Behavior and Habit Learning in Obsessive-Compulsive Disorder. *American Journal of Psychiatry*, *168*(7), 718–726. https://doi.org/10.1176/appi.ajp.2011.10071062

Glimcher, P. W., & Fehr, E. (2014). *Neuroeconomics: Decision making and the brain* (2nd ed.). Elsevier/AP.

Gluth, S., Rieskamp, J., & Buchel, C. (2012). Deciding When to Decide: Time-Variant Sequential Sampling Models Explain the Emergence of Value-Based Decisions in the Human Brain. *Journal of Neuroscience*, *32*(31), 10686–10698. https://doi.org/10.1523/JNEUROSCI.0727-12.2012

Goel, V., & Grafman, J. (1994). Are the frontal lobes implicated in 'planning' functions? Interpreting data from the Tower of Hanoi. *Neuropsychologia*, *33*(5), 623–642.

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, *30*, 535–574.

Goldt, S., & Seifert, U. (2017). Thermodynamic efficiency of learning a rule in neural networks. *New Journal of Physics*, *19*(11), 113001. https://doi.org/10.1088/1367-2630/aa89ff

Golman, R., & Loewenstein, G. (2018). Information gaps: A theory of preferences regarding the presence and absence of information. *Decision*, *5*(3), 143–164. https://doi.org/10.1037/dec0000068

Golman, R., Loewenstein, G., Molnar, A., & Saccardo, S. (2022). The Demand for, and Avoidance of, Information. *Management Science*, *68*(9), 6355–7064.

Gopnik, A. (2020). Childhood as a solution to explore–exploit tensions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1803), 20190502. https://doi.org/10.1098/rstb.2019.0502

Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, *19*(12), 758–770. https://doi.org/10.1038/s41583-018-0078-0

Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, *391*(6666), 481–484. https://doi.org/10.1038/35135

Grefkes, C., & Fink, G. R. (2005). REVIEW: The functional organization of the intraparietal sulcus in humans and monkeys: Intraparietal sulcus in monkeys and humans, C. Grefkes and G. R. Fink. *Journal of Anatomy*, *207*(1), 3–17. https://doi.org/10.1111/j.1469-7580.2005.00426.x

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*(1–3), 293–321. https://doi.org/10.1016/S0001-6918(01)00019-1

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of Curiosity Modulate Hippocampus-Dependent Learning via the Dopaminergic Circuit. *Neuron*, *84*(2), 486–496. https://doi.org/10.1016/j.neuron.2014.08.060

Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage*, *62*(1), 154–166. https://doi.org/10.1016/j.neuroimage.2012.04.024

Haber, S. N., & Behrens, T. E. J. (2014). The Neural Network Underlying Incentive-Based Learning: Implications for Interpreting Circuit Disruptions in Psychiatric Disorders. *Neuron*, *83*(5), 1019–1039. https://doi.org/10.1016/j.neuron.2014.08.031

Halekoh, U., & Hojsgaard, S. (2014). A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models—The R Package pbkrtest. *Journal of Statistical Software*, *59*(9), 1–30.

Hare, T. A., Schultz, W., Camerer, C. F., O'Doherty, J. P., & Rangel, A. (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proceedings of the National Academy of Sciences*, *108*(44), 18120–18125. https://doi.org/10.1073/pnas.1109322108

Harmer, C. J., Goodwin, G. M., & Cowen, P. J. (2009). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *British Journal of Psychiatry*, *195*(2), 102–108. https://doi.org/10.1192/bjp.bp.108.051193

Hauser, T. U., Moutoussis, M., NSPN Consortium, Dayan, P., & Dolan, R. J. (2017). Increased decision thresholds trigger extended information gathering across the compulsivity spectrum. *Translational Psychiatry*, *7*(1296).

Haxby, J. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *Neuroimage*, *62*(2), 852–855.

Hermans, D., Engelen, U., Grouwels, L., Joos, E., Lemmens, J., & Pieters, G. (2008). Cognitive confidence in obsessive-compulsive disorder: Distrusting perception, attention and memory. *Behaviour Research and Therapy*, *46*(1), 98–113. https://doi.org/10.1016/j.brat.2007.11.001

Holroyd, C. B., Ribas-Fernandes, J. J. F., Shahnazian, D., Silvetti, M., & Verguts, T. (2018). Human midcingulate cortex encodes distributed representations of task progress. *Proceedings of the National Academy of Sciences*, *115*(25), 6398–6403. https://doi.org/10.1073/pnas.1803650115

Hooks, M., Jones, G., Smith, A., Neill, D., & Justice, J. (1991). Response to novelty predicts the locomotor and nucleus accumbens dopamine response to cocaine. *Synapse*, *9*(2), 121–128.

Hooks, M. S., Jones, G. H., Smith, A. D., Neill, D. B., & Justice, J. B. (1991). Individual differences in locomotor activity and sensitization. *Pharmacology Biochemistry and Behavior*, *38*(2), 467–470. https://doi.org/10.1016/0091-3057(91)90308-O

Hooks, M. S., & Kalivas, P. W. (1995). The role of mesoaccumbens-pallidal circuitry in novelty-induced behavioral activation. *Neuroscience*, *64*(3), 587–597. https://doi.org/10.1016/0306-4522(94)00409-X

Horan, M., Daddaoua, N., & Gottlieb, J. (2019). Parietal neurons encode information sampling based on decision uncertainty. *Nature Neuroscience*, *22*(8), 1327–1335. https://doi.org/10.1038/s41593-019-0440-1

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

Hughes, R. N. (1965). Food deprivation and locomotor exploration in the white rat. *Animal Behaviour*, *13*, 30–32.

Hughes, R. N. (1968). Behaviour of male and female rats with free choice of two environments differing in novelty. *Animal Behaviour*, *16*(1), 92–96. https://doi.org/10.1016/0003-3472(68)90116-4

Hughes, R. N. (2007). Neotic preferences in laboratory rodents: Issues, assessment and substrates. *Neuroscience & Biobehavioral Reviews*, *31*(3), 441–464. https://doi.org/10.1016/j.neubiorev.2006.11.004

Hunt, L. T., Daw, N. D., Kaanders, P., MacIver, M. A., Mugan, U., Procyk, E., Redish, A. D., Russo, E., Scholl, J., Stachenfeld, K., Wilson, C. R. E., & Kolling, N. (2021). Formalizing planning and information search in naturalistic decision-making. *Nature Neuroscience*, *24*(8), 1051–1064. https://doi.org/10.1038/s41593-021-00866-w

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*, *15*(3), 470–476. https://doi.org/10.1038/nn.3017

Hunt, L. T., Malalasekera, W. M. N., de Berker, A. O., Miranda, B., Farmer, S. F., Behrens, T. E. J., & Kennerley, S. W. (2018). Triple dissociation of attention and decision computations across prefrontal cortex. *Nature Neuroscience*, *21*, 1471–1481.

Hunt, L. T., Rutledge, R. B., Malalasekera, W. M. N., Kennerley, S. W., & Dolan, R. J. (2016). Approach-Induced Biases in Human Information Sampling. *PLOS Biology*, *14*(11), e2000638. https://doi.org/10.1371/journal.pbio.2000638

Iigaya, K., Hauser, T. U., Kurth-Nelson, Z., O'Doherty, J. P., Dayan, P., & Dolan, R. J. (2020). The value of what's to come: Neural mechanisms coupling prediction error and the utility of anticipation. *Science Advances*, *6*(25), eaba3828. https://doi.org/10.1126/sciadv.aba3828

Iigaya, K., Story, G. W., Kurth-Nelson, Z., Dolan, R. J., & Dayan, P. (2016). The modulation of savouring by prediction error and its effects on choice. *ELife*, *5*, e13747. https://doi.org/10.7554/eLife.13747

Isoda, M., & Hikosaka, O. (2007). Switching from automatic to controlled action by monkey medial frontal cortex. *Nature Neuroscience*, *10*(2), 240–248. https://doi.org/10.1038/nn1830

Jacobsen, P., Freeman, D., & Salkovskis, P. (2012). Reasoning bias and belief conviction in obsessive-compulsive disorder and delusions: Jumping to conclusions across disorders?: Reasoning bias and belief conviction in OCD and delusions. *British Journal of Clinical Psychology*, *51*(1), 84–99. https://doi.org/10.1111/j.2044-8260.2011.02014.x

Jamieson, D. G., & Petrusic, W. M. (1977). Preference and the time to choose. *Organizational Behavior and Human Performance*, *19*(1), 56–67. https://doi.org/10.1016/0030-5073(77)90054-X

Jang, A. I., Sharma, R., & Drugowitsch, J. (2021). Optimal policy for attention-modulated decisions explains human fixation behavior. *ELife*, *10*, e63436. https://doi.org/10.7554/eLife.63436

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Jenkinson, M., & Chappell, M. (2018). *Introduction to Neuroimaging Analysis*. Oxford University Press.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156. https://doi.org/10.1016/S1361-8415(01)00036-6

Jepma, M., Verdonschot, R. G., van Steenbergen, H., Rombouts, S. A. R. B., & Nieuwenhuis, S. (2012). Neural mechanisms underlying the induction and relief of perceptual curiosity. *Frontiers in Behavioral Neuroscience*, *6*. https://doi.org/10.3389/fnbeh.2012.00005

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review*, *32*(2), 125–160. https://doi.org/10.1016/j.dr.2012.04.002

Johnson, A., & Redish, A. D. (2007). Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point. *Journal of Neuroscience*, *27*(45), 12176–12189. https://doi.org/10.1523/JNEUROSCI.3761-07.2007

Kaanders, P., Nili, H., O'Reilly, J. X., & Hunt, L. T. (2021). Medial frontal cortex activity predicts information sampling in economic choice. *Journal of Neuroscience*, *41*(40), 8403–8413. https://doi.org/10.1101/2020.11.24.395814

Kaanders, P., Sepulveda, P., Folke, T., Ortoleva, P., & De Martino, B. (2022). Humans actively sample evidence to support prior beliefs. *ELife*, *11*, e71768. https://doi.org/10.7554/eLife.71768

Kakade, S., & Dayan, P. (2002). Dopamine: Generalization and bonuses. *Neural Networks*, *15*, 549–559.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–685. https://doi.org/10.1038/nn1444

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T., & Camerer, C. F. (2009). The Wick in the Candle of Learning: Epistemic

Curiosity Activates Reward Circuitry and Enhances Memory. *Psychological Science*, *20*(8), 963–973. https://doi.org/10.1111/j.1467-9280.2009.02402.x

Karlsson, M. P., Tervo, D. G. R., & Karpova, A. Y. (2012). Network Resets in Medial Prefrontal Cortex Mark the Onset of Behavioral Uncertainty. *Science*, *338*(6103), 135–139. https://doi.org/10.1126/science.1226518

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *J Risk Uncertain*, *38*, 95–115.

Kay, K., Chung, J. E., Sosa, M., Schor, J. S., Karlsson, M. P., Larkin, M. C., Liu, D. F., & Frank, L. M. (2020). Constant Sub-second Cycling between Representations of Possible Futures in the Hippocampus. *Cell*, *180*(3), 552-567.e25. https://doi.org/10.1016/j.cell.2020.01.014

Keltikangas-Järvinen, L., Räikkönen, K., Ekelund, J., & Peltonen, L. (2004). Nature and nurture in novelty seeking. *Molecular Psychiatry*, *9*(3), 308–311. https://doi.org/10.1038/sj.mp.4001433

Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLoS Computational Biology*, *10*(11), e1003915. https://doi.org/10.1371/journal.pcbi.1003915

Kidd, C., & Hayden, B. Y. (2015). The Psychology and Neuroscience of Curiosity. *Neuron*, *88*(3), 449–460. https://doi.org/10.1016/j.neuron.2015.09.010

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks Effect: Human Infants Allocate Attention to Visual Sequences That Are Neither Too Simple Nor Too Complex. *PLoS ONE*, *7*(5), e36399. https://doi.org/10.1371/journal.pone.0036399

Kim, K., Sano, M., De Freitas, J., Haber, N., & Yamins, D. (2020). Active World Model Learning with Progress Curiosity. *ArXiv:2007.07853 [Cs, Stat]*. http://arxiv.org/abs/2007.07853

Kluger, J. (2020, May 12). Accidental Poisonings Increased After President Trump's Disinfectant Comments. *Time*. https://time.com/5835244/accidental-poisonings-trump/

Kobayashi, K., & Hsu, M. (2019). Common neural code for reward and information value. *Proceedings of the National Academy of Sciences*, *116*(26), 13061–13066. https://doi.org/10.1073/pnas.1820145116

Kobayashi, K., Lee, S., Filipowicz, A. L. S., McGaughey, K. D., Kable, J. W., & Nassar, M. R. (2021). *Dynamic Representation of the Subjective Value of Information* [Preprint]. Neuroscience. https://doi.org/10.1101/2021.02.12.431038

Kobayashi, K., Ravaioli, S., Baranès, A., Woodford, M., & Gottlieb, J. (2019). Diverse motives for human curiosity. *Nature Human Behaviour*, *3*(6), 587–595. https://doi.org/10.1038/s41562-019-0589-3

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural Mechanisms of Foraging. *Science*, *336*(6077), 95–98. https://doi.org/10.1126/science.1216930

Kolling, N., Scholl, J., Chekroud, A., Trier, H. A., & Rushworth, M. F. S. (2018). Prospection, Perseverance, and Insight in Sequential Behavior. *Neuron*, *99*(5), 1069-1082.e7. https://doi.org/10.1016/j.neuron.2018.08.018

Kolling, N., Wittmann, M. K., Behrens, T. E. J., Boorman, E. D., Mars, R. B., & Rushworth, M. F. S. (2016). Value, search, persistence and model updating in

anterior cingulate cortex. *Nature Neuroscience*, *19*(10), 1280–1285. https://doi.org/10.1038/nn.4382

Kragel, P. A., Kano, M., Van Oudenhove, L., Ly, H. G., Dupont, P., Rubio, A., Delon-Martin, C., Bonaz, B. L., Manuck, S. B., Gianaros, P. J., Ceko, M., Reynolds Losin, E. A., Woo, C.-W., Nichols, T. E., & Wager, T. D. (2018). Generalizable representations of pain, cognitive control, and negative emotion in medial frontal cortex. *Nature Neuroscience*, *21*(2), 283–289. https://doi.org/10.1038/s41593-017-0051-7

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. https://doi.org/10.1038/nn.2635

Kreps, D. M., & Porteus, E. L. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica*, *46*(1), 185. https://doi.org/10.2307/1913656

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*(8), 401–412. https://doi.org/10.1016/j.tics.2013.06.007

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, *60*(6), 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253. https://doi.org/10.1017/S0140525X16001837

Lau, J. K. L., Ozono, H., Kuratomi, K., Komiya, A., & Murayama, K. (2020). Shared striatal activity in decisions to satisfy curiosity and hunger at the risk of electric shocks. *Nature Human Behaviour*, *4*(5), 531–543. https://doi.org/10.1038/s41562-020-0848-3

Le Moal, M., & Simon, H. (1991). Mesocorticolimbic dopaminergic network: Functional and regulatory roles. *Physiological Reviews*, *71*(1), 155–234.

Lerman, C., Hughes, C., Lemon, S., Main, D., Snyder, C., Durham, C., Narod, S., & Lynch, H. (1998). What you don't know can hurt you: Adverse psychologic effects in members of BRCA1-linked and BRCA2-linked families who decline genetic testing. *Journal of Clinical Oncology*, *16*(5), 1650–1654.

Liebowitz, M. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, *22*, 141–173.

Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2011). The Decision Value Computations in the vmPFC and Striatum Use a Relative Value Code That is Guided by Visual Attention. *Journal of Neuroscience*, *31*(37), 13214–13223. https://doi.org/10.1523/JNEUROSCI.1246-11.2011

Litman, J., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition & Emotion*, *19*(4), 559–582. https://doi.org/10.1080/02699930441000427

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98.

Logothetis, N., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), 150–157.

Luyckx, F., Nili, H., Spitzer, B., & Summerfield, C. (2019). Neural structure mapping in human probabilistic reward learning. *ELife*, *8*, e42816.

Mansouri, F. A., Matsumoto, K., & Tanaka, K. (2006). Prefrontal Cell Activities Related to Monkeys' Success and Failure in Adapting to Rule Changes in a Wisconsin Card Sorting Test Analog. *The Journal of Neuroscience*, *26*(10), 2745–2756. https://doi.org/10.1523/JNEUROSCI.5238-05.2006

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*(7474), 78–84. https://doi.org/10.1038/nature12742

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., … Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), 654–660. https://doi.org/10.1038/s41586-022-04492-9

Marin, R., Biedrzycki, R., & Firinciogullari, S. (1991). Reliability and validity of the apathy evaluation scale. *Psychiatry Research*, *38*, 143–162.

Mars, R. B., Jbabdi, S., Sallet, J., O'Reilly, J. X., Croxson, P. L., Olivier, E., Noonan, M. P., Bergmann, C., Mitchell, A. S., Baxter, M. G., Behrens, T. E. J., Johansen-Berg, H., Tomassini, V., Miller, K. L., & Rushworth, M. F. S. (2011). Diffusion-Weighted Imaging Tractography-Based Parcellation of the Human Parietal Cortex and Comparison with Human and Macaque Resting-State Functional Connectivity. *Journal of Neuroscience*, *31*(11), 4087–4100. https://doi.org/10.1523/JNEUROSCI.5102-10.2011

Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, *21*(11), 1609–1617. https://doi.org/10.1038/s41593-018-0232-z

Milner, A., Beech, H., & Walker, V. (1971). Decision processes and obsessional behaviour. *British Journal of Social & Clinical Psychology*, *10*(1), 88–89.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making*, *5*(6), 437–449.

Misslin, Ren., Haberkorn, E., & Ropartz, P. (1981). Responses to Novelty and Changes in Behavior Across a 3-Week Postoperative Period in Hippocampal-Lesioned Mice I. *Physiology & Behavior*, *27*, 413–418.

Molenberghs, P., Mesulam, M. M., Peeters, R., & Vandenberghe, R. R. C. (2007). Remapping Attentional Priorities: Differential Contribution of Superior Parietal Lobule and Intraparietal Sulcus. *Cerebral Cortex*, *17*(11), 2703–2712. https://doi.org/10.1093/cercor/bhl179

Moutoussis, M., Bentall, R., El-Deredy, W., & Dayan, P. (2011). Bayesian modelling of Jumping-to-Conclusions bias in delusional patients. *Cognitive Neuropsychiatry*, *16*(5), 422–447.

Müller, N. G., & Ebeling, D. (2008). Attention-modulated activity in visual cortex—More than a simple 'spotlight'. *NeuroImage*, *40*(2), 818–827. https://doi.org/10.1016/j.neuroimage.2007.11.060

Murray, J. D., Bernacchia, A., Freedman, D. J., Romo, R., Wallis, J. D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., & Wang, X.-J. (2014). A

hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience*, *17*(12), 1661–1663. https://doi.org/10.1038/nn.3862

Nair, A., Pong, V., Dalal, M., Bahl, S., Lin, S., & Levine, S. (2018). Visual Reinforcement Learning with Imagined Goals. *ArXiv:1807.04742 [Cs, Stat]*. http://arxiv.org/abs/1807.04742

Navarro, D. J., Newell, B. R., & Schulze, C. (2016). Learning and choosing in an uncertain world: An investigation of the explore–exploit dilemma in static and dynamic environments. *Cognitive Psychology*, *85*, 43–77. https://doi.org/10.1016/j.cogpsych.2016.01.001

Nemeroff, C., & Kalivas, P. (1988). *The Mesocorticolimbic Dopamine System*. New York Academy of Sciences.

Neubert, F.-X., Mars, R. B., Sallet, J., & Rushworth, M. F. S. (2015). Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proceedings of the National Academy of Sciences*, *112*(20), E2695–E2704. https://doi.org/10.1073/pnas.1410767112

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Prentice-Hall.

Ng, A. Y., Harada, D., & Russell, S. (1999). Policy invariance under reward transformations: Theory and application to reward shaping. *Proceedings of the 16th International Conference on Machine Learning*.

Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, *2*(2), 175–220.

Nicki, R. M. (1970). The reinforcing effect of uncertainty reduction on a human operant. *Canadian Journal of Psychology*, *24*(6), 389–400.

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154. https://doi.org/10.1016/j.jmp.2008.12.005

Olley, A., Malhi, G., & Sachdev, P. (2007). Memory and executive functioning in obsessive–compulsive disorder: A selective review. *Journal of Affective Disorders*, *104*(1–3), 15–23. https://doi.org/10.1016/j.jad.2007.02.023

Olshausen, B., & Field, D. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607–609.

Op de Beeck, H. P. (2010). Against hyperacuity in brain reading: Spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage*, *49*(3), 1943–1948. https://doi.org/10.1016/j.neuroimage.2009.02.047

O'Reilly, J. X., Schuffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, *110*(38), E3660–E3669. https://doi.org/10.1073/pnas.1305373110

Padoa-Schioppa, C. (2011). Neurobiology of Economic Choice: A Good-Based Model. *Annual Review of Neuroscience*, *34*(1), 333–359. https://doi.org/10.1146/annurev-neuro-061010-113648

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), 223–226. https://doi.org/10.1038/nature04676

Paterson, A. (1999). Dopamine D4 Receptor Gene Novelty or Nonsense? *Neuropsychopharmacology*, *21*(1), 3–16. https://doi.org/10.1016/S0893-133X(98)00115-8

Patton, J., Stanford, M., & Barratt, E. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology*, *51*, 768–774.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, *16*(2), 366–387. https://doi.org/10.1016/0030-5073(76)90022-2

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pélissier, M.-C., & O'Connor, K. P. (2002). Deductive and inductive reasoning in obsessive-compulsive disorder. *British Journal of Clinical Psychology*, *41*(1), 15–27. https://doi.org/10.1348/014466502163769

Philiastides, M. G., Biele, G., & Heekeren, H. R. (2010). A mechanistic account of value computation in the human brain. *Proceedings of the National Academy of Sciences*, *107*(20), 9430–9435. https://doi.org/10.1073/pnas.1001732107

Piazza, P., Deminiere, J., Le Moal, M., & Simon, H. (1989). Factors That Predict Individual Vulnerability to Amphetamine Self-Administration. *Science*, *245*, 1511–1513.

Pierce, R. C., Crawford, C. A., Nonneman, A. J., Mattingly, B. A., & Bardo, M. T. (1990). Effect of forebrain dopamine depletion on novelty-induced place

preference behavior in rats. *Pharmacology Biochemistry and Behavior*, *36*(2), 321–325. https://doi.org/10.1016/0091-3057(90)90411-A

Pisauro, M. A., Fouragnan, E., Retzler, C., & Philiastides, M. G. (2017). Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nature Communications*, *8*(1), 15808. https://doi.org/10.1038/ncomms15808

Poldrack, R. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*, *2*, 67–70.

Poyurovsky, M., & Koran, L. M. (2005). Obsessive–compulsive disorder (OCD) with schizotypy vs. schizophrenia with OCD: Diagnostic dilemmas and therapeutic implications. *Journal of Psychiatric Research*, *39*(4), 399–408. https://doi.org/10.1016/j.jpsychires.2004.09.004

Raileanu, R., & Rocktäschel, T. (2020). RIDE: Rewarding Impact-Driven Exploration for Procedurally-Generated Environments. *ArXiv:2002.12292 [Cs]*. http://arxiv.org/abs/2002.12292

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Reed, P., Mitchell, C., & Nokes, T. (1996). Intrinsic reinforcing properties of putatively neutral stimuli in an instrumental two-lever discrimination task. *Animal Learning & Behavior*, *24*(1), 38–45. https://doi.org/10.3758/BF03198952

Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 585–590. https://doi.org/10.1038/nature12160

Rigotti, M., Ben Dayan Rubin, D., Morrison, S. E., Salzman, C. D., & Fusi, S. (2010). Attractor concretion as a mechanism for the formation of context representations. *NeuroImage*, *52*(3), 833–847. https://doi.org/10.1016/j.neuroimage.2010.01.047

Rigotti, M., & Fusi, S. (2016). Estimating the dimensionality of neural responses with fMRI Repetition Suppression. *ArXiv:1605.03952 [q-Bio, Stat]*. http://arxiv.org/abs/1605.03952

Robbins, T. W., Gillan, C. M., Smith, D. G., de Wit, S., & Ersche, K. D. (2012). Neurocognitive endophenotypes of impulsivity and compulsivity: Towards dimensional psychiatry. *Trends in Cognitive Sciences*, *16*(1), 81–91. https://doi.org/10.1016/j.tics.2011.11.009

Rodriguez, C. A., Turner, B. M., Van Zandt, T., & McClure, S. M. (2015). The neural basis of value accumulation in intertemporal choice. *European Journal of Neuroscience*, *42*(5), 2179–2189. https://doi.org/10.1111/ejn.12997

Rodriguez Cabrero, J. A. M., Zhu, J.-Q., & Ludvig, E. A. (2019). Costly curiosity: People pay a price to resolve an uncertain gamble early. *Behavioural Processes*, *160*, 20–25. https://doi.org/10.1016/j.beproc.2018.12.015

Ruitenberg, M. F. L., Koppelmans, V., De Dios, Y. E., Gadd, N. E., Wood, S. J., Reuter-Lorenz, P. A., Kofman, I., Bloomberg, J. J., Mulavara, A. P., & Seidler, R. D. (2018). Neural correlates of multi-day learning and savings in

sensorimotor adaptation. *Scientific Reports*, *8*(1), 14286. https://doi.org/10.1038/s41598-018-32689-4

Ruscio, A. M., Stein, D. J., Chiu, W. T., & Kessler, R. C. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry*, *15*(1), 53–63. https://doi.org/10.1038/mp.2008.94

Rushworth, M. F. S., Paus, T., & Sipila, P. K. (2001). Attention Systems and the Organization of the Human Parietal Cortex. *Journal of Neuroscience*, *21*(14), 5262–5271.

Sachdev, P. S., & Malhi, G. S. (2005). Obsessive–Compulsive Behaviour: A Disorder of Decision-Making. *Australian and New Zealand Journal of Psychiatry*, *39*, 757–763.

Sallet, J., Mars, R. B., Noonan, M. P., Neubert, F.-X., Jbabdi, S., O'Reilly, J. X., Filippini, N., Thomas, A. G., & Rushworth, M. F. (2013). The Organization of Dorsal Frontal Cortex in Humans and Macaques. *Journal of Neuroscience*, *33*(30), 12255–12274. https://doi.org/10.1523/JNEUROSCI.5108-12.2013

Saunders, J., Aasland, O., Babor, T., De La Fuente, J., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, *88*, 791–804.

Schuck, N. W., Gaschler, R., Wenke, D., Heinzle, J., Frensch, P. A., Haynes, J.-D., & Reverberi, C. (2015). Medial Prefrontal Cortex Predicts Internally Driven Strategy Shifts. *Neuron*, *86*(1), 331–340. https://doi.org/10.1016/j.neuron.2015.03.015

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*, 1593–1598.

Schwartenbeck, P., Passecker, J., Hauser, T. U., FitzGerald, T. H., Kronbichler, M., & Friston, K. J. (2019). Computational mechanisms of curiosity and goal-directed exploration. *ELife*, *8*, e41703. https://doi.org/10.7554/eLife.41703

Sekar, R., Rybkin, O., Daniilidis, K., Abbeel, P., Hafner, D., & Pathak, D. (2020). Planning to Explore via Self-Supervised World Models. *ArXiv:2005.05960 [Cs, Stat]*. http://arxiv.org/abs/2005.05960

Sengupta, A., Yakupov, R., Speck, O., Pollmann, S., & Hanke, M. (2017). The effect of acquisition resolution on orientation decoding from V1 BOLD fMRI at 7 T. *NeuroImage*, *148*, 64–76. https://doi.org/10.1016/j.neuroimage.2016.12.040

Sepulveda, P., Usher, M., Davies, N., Benson, A. A., Ortoleva, P., & De Martino, B. (2020). Visual attention modulates the integration of goal-relevant evidence and not value. *ELife*, *9*, e60705. https://doi.org/10.7554/eLife.60705

Shadlen, M. N., & Shohamy, D. (2016). Decision Making and Sequential Sampling from Memory. *Neuron*, *90*(5), 927–939. https://doi.org/10.1016/j.neuron.2016.04.036

Shahnazian, D., & Holroyd, C. B. (2018). Distributed representations of action sequences in anterior cingulate cortex: A recurrent neural network approach. *Psychonomic Bulletin & Review*, *25*(1), 302–321. https://doi.org/10.3758/s13423-017-1280-1

Shepard, R. (1987). Toward a Universal Law of Generalization for Psychological Science. *Science*, *237*(4820), 1317–1323.

Shepard, R., & Chipman, S. (1970). Second-Order Isomorphism of Internal Representations: Shapes of States. *Cognitive Psychology*, *1*, 1–17.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322.

Singh, S., Lewis, R. L., Barto, A. G., & Sorg, J. (2010). Intrinsically Motivated Reinforcement Learning: An Evolutionary Perspective. *IEEE Transactions on Autonomous Mental Development*, *2*(2), 70–82. https://doi.org/10.1109/TAMD.2010.2051031

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. https://doi.org/10.1002/hbm.10062

Smith, S. M., & Krajbich, I. (2019). Gaze Amplifies Value in Decision Making. *Psychological Science*, *30*(1), 116–128. https://doi.org/10.1177/0956797618810521

Smith, S., & Nichols, T. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Spielberger, C., Gorsuch, R., Lushene, R., Vagg, P., & Jacobs, G. (1983). Manual for the State-Trait Anxiety Inventory. *Consulting Psychologists Press*.

Stolba, A., & Wood-Gush, D. G. M. (1980). Arousal and exploration in growing pigs in different environments. *Applied Animal Ethology*, *6*(4), 382–383. https://doi.org/10.1016/0304-3762(80)90140-6

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006. https://doi.org/10.1038/nn.2163

Taghzouti, K., Simon, H., Louilot, A., Herman, J. P., & Le Moal, M. (1985). Behavioral study after local injection of 6-hydroxydopamine into the nucleus accumbens in the rat. *Brain Research*, *344*(1), 9–20. https://doi.org/10.1016/0006-8993(85)91184-9

Tang, E., Mattar, M. G., Giusti, C., Lydon-Staley, D. M., Thompson-Schill, S. L., & Bassett, D. S. (2019). Effective learning is accompanied by high-dimensional and efficient representations of neural activity. *Nature Neuroscience*, *22*(6), 1000–1009. https://doi.org/10.1038/s41593-019-0400-9

Tanner, J., & Itti, L. (2019). A top-down saliency model with goal relevance. *Journal of Vision*, *19*(1), 11. https://doi.org/10.1167/19.1.11

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision*, *11*(5), 5–5. https://doi.org/10.1167/11.5.5

Tolin, D. F., Abramowitz, J. S., Brigidi, B. D., Amir, N., Street, G. P., & Foa, E. B. (2001). Memory and memory confidence in obsessive–compulsive disorder. *Behaviour Research and Therapy*, *39*(8), 913–927. https://doi.org/10.1016/S0005-7967(00)00064-4

Tolin, D. F., Abramowitz, J. S., Brigidi, B. D., & Foa, E. B. (2003). Intolerance of uncertainty in obsessive-compulsive disorder. *Journal of Anxiety Disorders*, *17*(2), 233–242. https://doi.org/10.1016/S0887-6185(02)00182-2

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208.

Tomov, M. S., Truong, V. Q., Hundia, R. A., & Gershman, S. J. (2020). Dissociable neural correlates of uncertainty underlie different exploration strategies. *Nature Communications*, *11*(1), 2371. https://doi.org/10.1038/s41467-020-15766-z

Tremblay, S., Doucet, G., Pieper, F., Sachs, A., & Martinez-Trujillo, J. (2015). Single-Trial Decoding of Visual Attention from Local Field Potentials in the Primate Lateral Prefrontal Cortex Is Frequency-Dependent. *Journal of Neuroscience*, *35*(24), 9038–9049. https://doi.org/10.1523/JNEUROSCI.1041-15.2015

Trudel, N., Scholl, J., Klein-Flügge, M. C., Fouragnan, E., Tankelevitch, L., Wittmann, M. K., & Rushworth, M. F. S. (2021). Polarity of uncertainty representation during exploration and exploitation in ventromedial prefrontal cortex. *Nature Human Behaviour*, *8*(1), 83–98. https://doi.org/10.1038/s41562-020-0929-3

Tsividis, P. A., Pouncy, T., Xu, J. L., Tenenbaum, J. B., & Gershman, S. J. (2017). Human learning in Atari. *AAAI Spring Symposium Series, Science of Intelligence: Computational Principles of Natural and Artificial Intelligence, Technical Report SS-17-07*.

Twomey, K. E., & Westermann, G. (2018). Curiosity-based learning in infants: A neurocomputational approach. *Developmental Science*, *21*(4), e12629. https://doi.org/10.1111/desc.12629

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592.

van Lieshout, L. L. F., Vandenbroucke, A. R. E., Müller, N. C. J., Cools, R., & de Lange, F. P. (2018). Induction and Relief of Curiosity Elicit Parietal and Frontal Activity. *The Journal of Neuroscience*, *38*(10), 2579–2588. https://doi.org/10.1523/JNEUROSCI.2816-17.2018

Venkatraman, V., Rosati, A. G., Taren, A. A., & Huettel, S. A. (2009). Resolving Response, Decision, and Strategic Control: Evidence for a Functional Topography in Dorsomedial Prefrontal Cortex. *Journal of Neuroscience*, *29*(42), 13158–13164. https://doi.org/10.1523/JNEUROSCI.2708-09.2009

Volans, P. (1976). Styles of decision-making and probability appraisal in selected obsessional and phobic patients. *British Journal of Social & Clinical Psychology*, *15*(3), 305–317.

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, *137*, 188–200. https://doi.org/10.1016/j.neuroimage.2015.12.012

Wang, J., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., & Botvinick, M. (2018). Prefrontal Cortex as a Meta-Reinforcement Learning System. *Nature Neuroscience*, *21*(6), 860.

Wang, M. Z., & Hayden, B. Y. (2019). Monkeys are Curious about Counterfactual Outcomes. *BioRxiv*.

Wang, M. Z., & Hayden, B. Y. (2020). Curiosity is associated with enhanced tonic firing in dorsal anterior cingulate cortex. *BioRxiv*. https://doi.org/10.1101/2020.05.25.115139

White, J. K., Bromberg-Martin, E. S., Heilbronner, S. R., Zhang, K., Pai, J., Haber, S. N., & Monosov, I. E. (2019). A neural network for information seeking. *Nature Communications*, *10*(1), 5168. https://doi.org/10.1038/s41467-019-13135-z

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*(6), 2074–2081. https://doi.org/10.1037/a0038199

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*, 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060

Wise, T., & Dolan, R. J. (2020). Associations between aversive learning processes and transdiagnostic psychiatric symptoms in a general population sample. *Nature Communications*, *11*(1), 4179. https://doi.org/10.1038/s41467-020-17977-w

Wittmann, B. C., Daw, N. D., Seymour, B., & Dolan, R. J. (2008). Striatal Activity Underlies Novelty-Based Choice in Humans. *Neuron*, *58*(6), 967–973. https://doi.org/10.1016/j.neuron.2008.04.027

Wood-Gush, D. G. M., & Vestergaard, K. (1991). The seeking of novelty and its relation to play. *Animal Behaviour*, *42*(4), 599–606. https://doi.org/10.1016/S0003-3472(05)80243-X

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–1747. https://doi.org/10.1016/j.neuroimage.2003.12.023

Yantis, S. (2008). The Neural Basis of Selective Attention: Cortical Sources and Targets of Attentional Modulation. *Current Directions in Psychological Science*, *17*(2), 86–90. https://doi.org/10.1111/j.1467-8721.2008.00554.x

Zhang, J., Hughes, L. E., & Rowe, J. B. (2012). Selection and inhibition mechanisms for human voluntary action decisions. *NeuroImage*, *63*(1), 392–402. https://doi.org/10.1016/j.neuroimage.2012.06.058

Zheng, Z., Oh, J., Hessel, M., Xu, Z., Kroiss, M., van Hasselt, H., Silver, D., & Singh, S. (2020). What Can Learned Intrinsic Rewards Capture? *ArXiv:1912.05500 [Cs]*. http://arxiv.org/abs/1912.05500

Zung, W. (1965). A self-rating depression scale. *Archives of General Psychiatry*, *12*, 63–70.