Goals and information processing in human decisions

Pradyumna Sepulveda Delgado

supervised by
Dr. Benedetto de Martino

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Cognitive Neuroscience

July 13, 2023
I, Pradyumna Sepulveda, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
ABSTRACT

We do not make decisions in the void. Every day, we act in awareness of our context, adjusting our objectives according to the situations we find. Operating effectively under multiple goals is fundamental for appropriate learning and decision-making, and deficiencies in this capacity can be at the core of mental disorders such as anxiety, depression or post-traumatic stress disorder. In this thesis, I present studies I conducted to investigate how goals impact different stages of the decision process, from simple perceptual choices to subjective value preferences.

Previous studies have described how animals assess alternatives and integrate evidence to make decisions. Most of the time, the focus of this work has been on simplified scenarios with single goals. In this thesis, my experiments tackle the issue of how people adjust information processing in tasks that demand more than one objective. Through various manipulations of the behavioural goals, such as decision framing, I show that (i) attention and evidence accumulation, (ii) brain representations, and (iii) decision confidence were all affected by context changes.

Using behavioural testing, computational models, and neuroimaging I show that goals have a crucial role in evidence integration and the allocation of visual attention. My findings indicate that brain patterns adapt to enhance goal-relevant information during learning and the valuation of alternatives. Finally, I report the presence of goal-dependent asymmetries in the generation of decision confidence, overweighting the evidence of the most-relevant option to fulfil the goal. In conclusion, I show how the entire process is highly flexible and serves the behavioural demands. These findings support the reinterpretation of some perspectives, such as reported biases and irrationalities in decisions, as attributes of adaptive processing towards goal fulfilment.
IMPACT STATEMENT

Being capable of resisting the temptation of eating dessert because you are on a diet is considered an example of “goal-directed” control. This type of decision allows people to regulate their behaviour, following higher-level objectives, beyond automatic habitual responses. These choices have been connected to the capacity of individuals to generate models of their world, beyond immediate actions. It is precisely this type of decision that may be affected by psychiatric conditions such as obsessive-compulsive disorder (OCD) or addiction. Given the lack of clarity on the origin of many psychiatric conditions, it is important to have a better characterisation of the process of goal-dependent choice to identify variables of interest or stages of processing that may have functional deficits. Therefore, understanding the mechanisms by which goals influence decisions is a key step in illuminating the divergent operation of the brain in these conditions. This thesis targets three stages of information processing in human decisions and how they are affected by goals: information sampling, brain representations and post-decisional metacognitive confidence.

Most of the neuroscience studies on decision-making have characterised choices in scenarios where the objective is constant along the tasks, e.g., choosing the preferred item in a value-based decision or indicating the option with higher luminosity in a perceptual choice. These studies have been extremely important to illuminate the underlying mechanisms of choice. Within this framework, some proposals have assumed that choices rely on factors entirely dependent on the alternatives, such as the hedonic reward value or perceptual evidence. However, this view does not consider that at a more fundamental level, it is the setting of the goals of the agents that inform how the available evidence in the alternatives is assessed and integrated for choice. In the experiments presented in this thesis, I point to dissociating these features of the alternatives from goals, supporting a perspective that places goals as the primary scenario in the decision. These findings can support a revision of the understanding of choices in neuroeconomics and perceptual neuroscience.

How decisions are made is a topic of high interest in various areas, beyond clinical and neuroscience circles. From a sociological, commercial, and even political perspective, understanding why people make certain choices, reassure their beliefs, or change their minds, is of capital importance. For example, social phenomena such as, how people are persuaded (or not) of getting a vaccine, or understanding the formation of financial bubbles, rely on understanding how individuals make choices, and the factors involved in this process. Good examples of this type of research in this thesis are the exploration of confirmation bias and how it is influenced by the way humans sample information, or how framing bias can be interpreted from a goal-directed perspective. Additionally, I capitalised on computational approaches to characterise the potential algorithms implemented by the brain when we make decisions. This interaction can be reciprocal since our findings can be relevant for researchers in machine learning and artificial intelligence looking for inspiration in the solutions that biological systems give to make choices in more complex scenarios.
ACKNOWLEDGEMENTS

Firstly I would like to thank Benedetto de Martino for his supervision during this PhD. I am deeply grateful for his patience all these years as a wise guide and a good friend. Being in the lab has been the best experience I could have asked for to grow as a researcher. Thank you for opening the door to a stranger.

Thank you to Aurelio Cortese, Steve Fleming, Pietro Ortoleva, Marius Usher and Hugo Spiers for their fundamental support in the projects I was lucky to be part of during my time at UCL.

Thank you to the members of the Decision Making group across the years, Mariana Zurita, Paula Kaanders, Amy Benson, Mihaela Nemes, Giuseppe Castegnetti, Yumeya Yamamori, Ned Davies, Annika Boldt, Tomas Folke, and Harry Costello (honorary member). I learned so much from you. My ICN memories are joyful, thanks to you.

I am also grateful to The National Research and Development Agency of Chile (ANID) for funding my PhD.

To all my friends in London, thank you for being my family in a foreign place. You made it feel like home.

To my parents and sister, thank you for encouraging me to ask questions.
STATEMENT OF PUBLICATIONS

The work I have been involved in during my PhD has resulted in the following peer-reviewed publications:


1. For a research manuscript that has already been published:
   (a) What is the title of the manuscript? Visual attention modulates the integration of goal-relevant evidence and not value
   (b) Please include a link to or doi for the work: https://doi.org/10.7554/eLife.60705
   (c) Where was the work published? eLife
   (d) Who published the work? eLife Sciences Publications, Ltd
   (e) When was the work published? November 17, 2020
   (f) List the manuscript’s authors in the order they appear on the publication: Pradyumna Sepulveda, Marius Usher, Ned Davies, Amy A Benson, Pietro Ortoleva, Benedetto De Martino
   (g) Was the work peer reviewed? Yes
   (h) Have you retained the copyright? Yes
   (i) Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)? If ‘Yes’, please give a link or doi: Yes, bioRxiv; https://doi.org/10.1101/2020.04.14.031971

2. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):
   (a) Pradyumna Sepulveda Conceptualization, Resources, Data curation, Software, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing - original draft, Project administration, Writing - review and editing
   (b) Marius Usher Supervision, Methodology, Writing - review and editing
   (c) Ned Davies Data curation, Formal analysis, Investigation, Writing - review and editing
   (d) Amy A Benson Data curation, Formal analysis, Investigation, Writing - review and editing
   (e) Pietro Ortoleva Conceptualization, Supervision, Methodology, Writing - original draft, Writing - review and editing, Developed model for optimal information acquisition
   (f) Benedetto De Martino Conceptualization, Supervision, Funding acquisition, Methodology, Writing - original draft, Project administration, Writing - review and editing

3. In which chapter(s) of your thesis can this material be found? Chapters 3 and 7

E-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/senior author unless this is not appropriate, e.g. if the paper was a single-author work):

Candidate: Pradyumna Sepulveda Date: 21 April 2023
Supervisor/Senior Author signature: Benedetto de Martino
Date: 21 April 2023
UCL RESEARCH PAPER DECLARATION FORM: REFERENCING THE DOCTORAL CANDIDATE’S OWN PUBLISHED WORK(S)

1. For a research manuscript that has already been published:
   (a) **What is the title of the manuscript?** Value signals guide abstraction during learning
   (b) **Please include a link to or doi for the work:** https://doi.org/10.7554/eLife.68943
   (c) **Where was the work published?** eLife
   (d) **Who published the work?** eLife Sciences Publications, Ltd
   (e) **When was the work published?** July 13, 2021
   (f) **List the manuscript's authors in the order they appear on the publication:** Aurelio Cortese, Asuka Yamamoto, Maryam Hashemzadeh, Pradyumna Sepulveda, Mitsuo Kawato, Benedetto De Martino
   (g) **Was the work peer reviewed?** Yes
   (h) **Have you retained the copyright?** Yes
   (i) **Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)?** If ‘Yes’, please give a link or doi: Yes, bioRxiv; https://doi.org/10.1101/2020.10.29.361469

2. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):
   (a) **Aurelio Cortese** Conceptualization, Resources, Data curation, Software, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing - original draft, Project administration, Writing - review and editing
   (b) **Asuka Yamamoto** Software, Formal analysis, Investigation, Methodology
   (c) **Maryam Hashemzadeh** Software, Formal analysis, Writing - review and editing
   (d) **Pradyumna Sepulveda** Conceptualization, Software, Investigation, Writing - review and editing
   (e) **Mitsuo Kawato** Conceptualization, Formal analysis, Supervision, Funding acquisition, Writing - original draft, Writing - review and editing
   (f) **Benedetto De Martino** Conceptualization, Formal analysis, Supervision, Validation, Writing - original draft, Project administration, Writing - review and editing

3. **In which chapter(s) of your thesis can this material be found?** Chapter 4

   e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/senior author unless this is not appropriate, e.g. if the paper was a single-author work):

   **Candidate:** Pradyumna Sepulveda **Date:** 21 April 2023
   **Supervisor/Senior Author signature:** Benedetto de Martino
   **Date:** 21 April 2023
1. For a research manuscript that has already been published:

(a) What is the title of the manuscript? Humans actively sample evidence to support prior beliefs.

(b) Please include a link to or doi for the work: https://doi.org/10.7554/eLife.71768

(c) Where was the work published? eLife

(d) Who published the work? eLife Sciences Publications, Ltd

(e) When was the work published? April 11, 2022

(f) List the manuscript's authors in the order they appear on the publication: Paula Kaanders, Pradyumna Sepulveda, Tomas Folke, Pietro Ortoleva, Benedetto De Martino

(g) Was the work peer reviewed? Yes

(h) Have you retained the copyright? Yes

(i) Was an earlier form of the manuscript uploaded to a preprint server (e.g. medRxiv)? If ‘Yes’, please give a link or doi Yes, bioRxiv; https://doi.org/10.1101/2021.06.29.450332

2. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4):

(a) Paula Kaanders
   Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft, Writing - review and editing

(b) Pradyumna Sepulveda
   Formal analysis, Methodology, Visualization, Writing - review and editing

(c) Tomas Folke
   Formal analysis, Software, Writing - review and editing

(d) Pietro Ortoleva
   Methodology, Writing - review and editing

(e) Benedetto De Martino
   Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing - review and editing

3. In which chapter(s) of your thesis can this material be found? Chapter 6

E-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/senior author unless this is not appropriate, e.g. if the paper was a single-author work):

Candidate: Pradyumna Sepulveda Date: 21 April 2023
Supervisor/Senior Author signature: Benedetto de Martino
Date: 21 April 2023
CONTENTS

Abstract 3
Impact Statement 4
Acknowledgements 5
Statement of Publications 6
UCL Research Paper Declaration Form 7

1 Introduction 15
  1.1 Perceptual decisions 16
  1.2 Value-based decisions 18
  1.3 Neural correlates of decisions 21
  1.4 Decision and confidence 22
  1.5 Goals and contexts 25

2 General Methods 29
  2.1 Summary 29
  2.2 Regression analysis 30
  2.3 Decision models: Signal Detection Theory 34
  2.4 Decision models: Sequential Sampling Models and Attention 37
  2.5 Probabilistic Graphical Models 41
  2.6 Reinforcement learning 46
  2.7 Functional Magnetic Resonance Imaging 48
    2.7.1 Univariate analysis 50
    2.7.2 Multivariate analysis 50
    2.7.3 Classifiers 52
    2.7.4 Representational similarity analysis 52
  2.8 Conclusion 54

III Information sampling and goals in human decision

3 Visual attention modulates the integration of goal-relevant evidence and not value 59
  3.1 Summary 59
  3.2 Introduction 59
  3.3 Methods 61
    3.3.1 Procedure 61
    3.3.2 Exclusion criteria 62
    3.3.3 Participants 63
    3.3.4 Eye-tracking 63
    3.3.5 Data analysis: behavioural data 64
    3.3.6 Data analysis: attentional model - GLAM 64
  3.4 Results 67
    3.4.1 The effect of attention on choice 67
    3.4.2 Fixations effects in choice 70
    3.4.3 Attentional model: GLAM 72
  3.5 Discussion 74

IV Brain representations and goals in human decisions

4 Goal-directed representations in learning: Value signal guides abstraction during learning 80
  4.1 Summary 80
CONTENTS

6.4.2 Fixed sampling condition 133
6.4.3 Attentional evidence accumulation modelling 135
6.5 Discussion 138
7 Decision goals and their impact in confidence 141
  7.1 Summary 141
  7.2 Introduction 141
  7.3 Methods 143
    7.3.1 Experiment Design 143
    7.3.2 Participants and Exclusion criteria 143
    7.3.3 Data analysis: behavioural data 143
    7.3.4 Modelling 144
  7.4 Results 145
    7.4.1 Which factors determine confidence? 145
    7.4.2 Attentional Model: GLAM 147
  7.5 Discussion 150
8 Goals modulate the evidence bias in human confidence 152
  8.1 Summary 152
  8.2 Introduction 152
  8.3 Methods 154
    8.3.1 Procedure 154
    8.3.2 Participants 157
    8.3.3 Data analysis: behavioural data 159
    8.3.4 Bayesian Models 159
    8.3.5 Data analysis: pupillometry experiment 163
  8.4 Results 164
    8.4.1 Evidence bias on confidence depends on frame 165
    8.4.2 Simulations of asymmetric variance can generate evidence bias 169
    8.4.3 Asymmetric variance model 171
    8.4.4 Equal variance model 172
    8.4.5 Heuristic model 172
    8.4.6 Asymmetric variance model fit to human data 174
    8.4.7 Attention as the cognitive mechanism 177
  8.5 Discussion 182
9 General Discussion 187
  9.1 Summary 187
  9.2 Decisions and the goal perspective 187
  9.3 Contexts and goals 191
  9.4 Psychiatric implications 193
  9.5 Conclusion 194
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Decision-making stages and the influence of goals</td>
<td>16</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Accumulator models</td>
<td>18</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Basic computations are required for value-based choice</td>
<td>20</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Neural activity in the LIP area during perceptual decisions</td>
<td>22</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Organisation of critical steps for value-based decision in the human brain cortex</td>
<td>23</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Confidence-related activity in rLPFC</td>
<td>25</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Context organization of neural activity in rats</td>
<td>26</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Relationships between variables and regressions</td>
<td>31</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Logistic regression</td>
<td>33</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Signal detection theory</td>
<td>35</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Two-dimensional interpretation of the 2AFC task</td>
<td>38</td>
</tr>
<tr>
<td>Figure 2.5</td>
<td>Drift diffusion models</td>
<td>39</td>
</tr>
<tr>
<td>Figure 2.6</td>
<td>Gaze-weighted linear accumulator model</td>
<td>41</td>
</tr>
<tr>
<td>Figure 2.7</td>
<td>Example Bayesian network</td>
<td>43</td>
</tr>
<tr>
<td>Figure 2.8</td>
<td>Bayesian graphical model for a perceptual decision experiment</td>
<td>44</td>
</tr>
<tr>
<td>Figure 2.9</td>
<td>Markov Chain Monte Carlo</td>
<td>45</td>
</tr>
<tr>
<td>Figure 2.10</td>
<td>Haemodynamic variations by blood flow</td>
<td>49</td>
</tr>
<tr>
<td>Figure 2.11</td>
<td>Univariate analysis in an auditory task</td>
<td>51</td>
</tr>
<tr>
<td>Figure 2.12</td>
<td>Multivariate pattern classification</td>
<td>53</td>
</tr>
<tr>
<td>Figure 2.13</td>
<td>Representational similarity analysis</td>
<td>55</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Task and behavioural results</td>
<td>68</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Attention and choice in Value and Perceptual Experiments</td>
<td>71</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Fixation effects on the chosen item</td>
<td>74</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Individual out-of-sample GLAM predictions for behavioural measures</td>
<td>74</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Learning task and behavioural results</td>
<td>89</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Mixture of reinforcement learning (RL) experts and value computation</td>
<td>90</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>Feature RL vs Abstract RL are related to learning speed and the use of abstraction increases with experience</td>
<td>92</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>Neural substrates of value construction during learning</td>
<td>95</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>Neural substrate of abstraction</td>
<td>96</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Task and behavioural analysis</td>
<td>111</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>Imagination and choice trials RSA in selected ROIs</td>
<td>113</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>Imagination searchlight RSA. Pet identity</td>
<td>114</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Frame representations in the occipital cortex</td>
<td>115</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>Imagination searchlight RSA. Goal-relevant value</td>
<td>116</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>Pattern similarity in imagination trials</td>
<td>117</td>
</tr>
<tr>
<td>Figure 5.7</td>
<td>Goal-directed pattern similarity in hippocampus during imagination trials</td>
<td>119</td>
</tr>
<tr>
<td>Figure 5.8</td>
<td>Choice trials brain patterns reinstatement</td>
<td>120</td>
</tr>
</tbody>
</table>
Figure 6.1  Task design and participant behaviour for the experiment 134
Figure 6.2  The effect of choice on sampling behaviour is mediated by confidence in experiment 135
Figure 6.3  Gaze impacted evidence accumulation (for the 2\textsuperscript{nd} choice) more strongly in the free than in the fixed sampling condition 137
Figure 7.1  Hierarchical linear regression model to predict confidence 146
Figure 7.2  Balance of evidence (\(\Delta e\)) simulated with GLAM reproduces \(\Sigma \text{Value}\) and \(\Sigma \text{Dots}\) effects over confidence 148
Figure 7.3  Pooled linear regressions to predict the balance of evidence (\(\Delta e\)) simulations 149
Figure 8.1  Tasks and hierarchical linear regressions predicting confidence 166
Figure 8.2  Bayesian observer model – asymmetric variance model (AVM) 170
Figure 8.3  Asymmetric Variance Model (AVM) simulations 172
Figure 8.4  Equal Variance Model (EVM) simulations 173
Figure 8.5  Heuristic Model (HM) simulations 174
Figure 8.6  AVM model fit 175
Figure 8.7  Bayesian graphical model indicating the formalization of the asymmetric variance model (AVM) 178
Figure 8.8  Model Comparison asymmetric variance model (AVM) and heuristic model (HM) 179
Figure 8.9  Pupil variations depend on the task goal 181
INTRODUCTION

Decisions are never made in the void. From seemingly trivial choices, e.g., putting ice or not into our drink, to complex choices, e.g., changing our career, real-world decisions always appear as steps, a fragment in a bigger picture. Decisions tend to be part of a flow of actions that point to satisfy internal needs and external demands, with the brain orchestrating many of these processes. Classical views in the study of behaviour, such as Bernard or Pavlov, define the brain as a control system that maintains a metastable balance between the internal world of the body and the external world through action [1]. Therefore, knowing the factors that help humans to make good decisions is fundamental to having a better understanding of this equilibrium. Decisions can have a local impact, like assuring individual survival by getting nutritious food [2], or have a wider influence, such as decisions on social issues, political opinions, and beliefs [3]. Difficulties to maintain these equilibriums are a key feature in many psychiatric and neurological conditions. Characterising the mechanisms behind decisions is pivotal to clarifying the origins and potential remedial measures for these disorders. For example, to form and execute plans, humans can engage in multistep behaviours such as preparing a cup of coffee or organizing a trip to Chile. When patients with lesions to the prefrontal cortex (PFC) are exposed to these complex tasks they often exhibit disordered action sequences that fail to achieve the specified goal [4, 5, 6]. Of all the factors controlling decisions, goals are primordial. Goals initiate the decision process, and they define the final state that agents hope to reach through their actions. Characterising the role of goals is not important only for humans and animals but can have critical consequences for the development of artificial intelligence systems. Indeed, that machines can respond in a human-like fashion may rely on their capacity to be aware of complex contextual cues and flexibly act to reach goals [7, 8]. In other words, the relationship between decisions and their goals is of capital importance.

Decisions constitute a complex pipeline of information, where inputs, such as external sensorial perception, and/or internal proprioception or emotions [9, 10] are processed and integrated into sequential stages to generate choice and action. Overall, the objective of the present work is to shed light on how goals impact three main stages of the decision process: (i) how evidence is attended to and accumulated, (ii) the generation of brain representations and, (iii) the sense of confidence (Figure 1.1).

In this introduction, I will present a general background on decision neuroscience to set the stage for more detailed analysis and discussions in the following chapters.
Figure 1.1: Decision-making stages and the influence of goals. The present thesis studies the influence of task goals on three main steps of the decision process. 1) In the preliminary stage, participants integrate information about the alternatives to make a choice, a process that is dynamically driven by attention. 2) During the deliberation process, the brain generates representations of the alternatives, a process that has been shown to rely on areas relevant to the decision process such as the hippocampus, and prefrontal cortex. 3) Finally, after a choice has been made, decision agents can generate an internal sense of how likely their selection was correct, i.e., an internal sense of uncertainty or confidence. This involves an assessment of the choice and the evidence employed to make a choice. In this thesis, I will show how these three stages adapt to fulfil the goals that the decision agents have to fulfil.

1.1 PERCEPTUAL DECISIONS

Even the most trivial decisions, like choosing coffee or tea, are extremely intricate as they involve an obscure and complex interaction of multiple factors. For this reason, the initial efforts to formalise a methodology to study decisions appear from the systematic analysis of perceptual choices in psychophysics [11]. In simple perceptual decisions, the experimental subjects use sensory stimuli, such as sound intensity or colour, to make their choices. This setup allows the experimenter to have a high degree of control over the evidence: it is easier and more reproducible to regulate experimentally the intensity of light or the volume of a sound than the internal strength of a memory or the subjective valuation of an item. The objective of psychophysics is to extract from choice behaviour (e.g., present/absent, more/less, left/right) properties of the sensory evidence. This implies that during the decision process, a representation of this evidence appears in the
brain and is compared with what is not evidence (or noise). Initial studies in monkeys performing a demanding visual discrimination task showed that neuronal activity in area MT/V5 of the extrastriate cortex can track decision variability, even at a single neuronal level [12]. The intuition that evidence and noise are tracked during decision inspired the application of signal detection theory (SDT, [13]) to perception (please check chapter 2 for more details on SDT). From SDT, two important concepts of the study of the decision in neuroscience emerged: the representation of raw evidence generates a decision variable, which is ‘used’ by the brain’s decision rule to make the choice (e.g., left or right, yes or no).

Experiments motivated by this framework suggest that neuronal activity, such as spikes, can be linked to the decision variable [12, 14, 15, 16]. Some of the potential operations using these brain representations could be subtraction, which allows a comparison between alternatives, or integration, which implies an evidence-processing timescale [11]. The latter operation is a crucial aspect of decisions, e.g., animals are slower to answer difficult choices and faster to respond to easy ones.

Sequential accumulator models have been used in psychophysics and decision neuroscience to capture this dynamic element of choice (Figure 1.2). This type of model can include two aspects related to timing, which can affect the response time: (1) a required elapsed time itself (e.g., a deadline), and (2) a level of evidence uncertainty. These two features are not mutually exclusive. The basic intuition behind these models is that integration of evidence could occur with neurons that “acquire” evidence in timescales higher than pure perceptual variation. In accumulator models, thresholds or boundaries represent the idea of a deadline that indicates the end of the integration of evidence. Adding noise into the accumulation process captures the uncertainty in the decision, which makes the integration more unstable and delays reaching the final threshold (for more details on accumulator models, please check chapter 2). Including this temporal dimension allows using a single mechanism for accounting for cases in which a fast choice, with a small number of samples, is favoured and in which it is better to take due time to deliver the correct choice, i.e. the speed-accuracy trade-off [17]. Further iterations of this model also consider choices driven by an “urgency signal”, which translates into collapsing decision boundaries in time, to explain behavioural signatures such as incorrect responses being slower than correct choices [18, 19].

Most of the perceptual decision research has been developed on the sense of vision [10], however, other sources of perceptual evidence have also been studied using a similar psychophysics approach. Perceptual decision processes for vibrotactile sensation in neurons of the somatosensory system [21, 22], smell and taste [23, 24, 25], hearing [26], multisensory scenarios [27], and even time passage [28] have been investigated. From the study of these different modalities, some distinctive features of perception have been found, such as the Weber-Fechner law, which states that the perceived change of a stimulus is proportional to the initial stimulus [29, 30, 31]. Overall, some of the general principles of decision-making computation resulting from the study of perceptual decisions are 1) the process is not defined by the perceptual changes in the environment as a reflex or tied to motor responses, but it is flexible in time, which allows the integration of evidence; 2) brain activity can be associated with a probabilistic representation or degree of belief in a choice, which favours the interpretation of the brain building a decision variable, i.e., a low-dimensional combination of the
Figure 1.2: Accumulator models. Drift diffusion model (DDM). A) Flowchart of evidence integration modelled by the DDM: the perceptual information is gathered by sensory channels and transformed to signals interpreted by the brain; a drift rate models how fast evidence is integrated which noisily accumulates over time until a boundary is reached. Superior or inferior boundaries represent the two alternatives in a binary choice. The decision process defines the features of the response: speed and accuracy. B) Noisy integration of evidence generates fast, medium, and slow responses (red lines). The integration of all the responses can be seen in the distribution of the reaction times (blue lines). C) Newer versions of the DDM include collapsing decision boundaries to capture an ‘urgency signal’, which could be interpreted as a reduction in the amount of evidence required to choose as the trial time passes. From [20], with Elsevier permission.

1.2 Value-based decisions

There is a whole world of decisions that do not depend exclusively on external perception but rely on internal preferences. Value-based decisions involve choosing among different options according to the subjective va-
The field of neuroeconomics has especially focused on this type of decision making choice models and protocols from economics, biological principles from neuroscience, behavioural data and theories about animal learning and choices from psychology, and computational models from computer science. The basis of this type of decision is the generation of internal value signals that allow the selection of the best options. Based on collaborative efforts from psychology and machine learning the “reward hypothesis” has been developed from the formalisation of reinforcement learning (RL). RL formalises operant learning paradigms in humans and other animals, characterising how the value of actions or options is learned. The RL model considers three main stages: 1) an agent receives sensory samples from the environment; 2) the agent takes actions that influence its future states; and (3) the agent receives a scalar reward signal generated by the environment and processed by the agent via a dedicated input channel. In scenarios of higher complexity, such as modelling subjective preferences in humans, methodologies from behavioural economics have been used to capture these values nurtured by years of experience, e.g., your favourite drink or go-to song for a karaoke night.

The computations involved in value-based decision-making have been divided into 5 general processes [32] (Figure 1.3). First, the representation stage considers that agents have to identify internal and external states and potential actions. Second, each one of the represented actions needs to be assessed in the valuation stage. Third, using the assigned values a comparison is performed to inform the action selection stage. Fourth, the outcome evaluation stage after choice assesses the desirability of the outcomes using the resulting feedback. Finally, the collected outcomes are used in a learning stage to update the three initial stages to generate better choices in the future. Note that this sequence assumes that learning is a continuous process, i.e., each new decision is informing the future choices.

The level of insight into these stages is quite dissimilar. Little is known about how the brain represents internal and external states (e.g., how the possible options are identified for each situation) since this is a complex process that requires the integration of various sources of information. The presence of valuation systems has been generally accepted, although their exact neural implementation and characteristics are an ongoing area of research. The Pavlovian valuation system assigns prominence to a restricted set of evolutionarily relevant actions, such as preparatory behaviours for food, or avoiding aversive stimuli, such as heat or electric shocks. For this reason, the behavioural repertoire of Pavlovian responses is rather limited and inflexible. In contrast, the habitual valuation system can cover a wider set of actions assuming that they are repeatedly experienced. In the habitual system, the associations of stimulus-response are learned, relying on past experiences through a training process involving trial and error. This makes habitual learning a slow process. For example, in this category includes a rat learning to press a lever for liquids in response to a sound, or a smoker’s desire for a cigarette after a meal. Finally, in the goal-directed valuation system, values are associated with actions. The values in this case depend on action-outcomes associations and the evaluation of those outcomes. Since each action can lead to multiple results, it is commonly assumed that an average value of all of them is representative of the action. Goal-directed behaviours are considered to be more flexible, not relying entirely on repeated training,
Basic computations are required for value-based choice. 1) Construction of representations involves identifying internal and external relevant states; 2) the value of the available options is assessed; 3) the previous evaluation is employed to choose an action; 4) after action, the brain measures the outcome in terms of its desirability; and 5) the outcomes of choice are used to learn and update all the stages to obtain future success. From [32], with Springer Nature permission.

However, they may require a higher investment of cognitive resources, given a more detailed assessment of the possibilities. In this category, we can find decisions such as choosing a new restaurant for a date. All these three systems can interact and be modulated by additional factors, such as risk, uncertainty and time discounting.

Regarding the action selection stage, the proposed avenues borrow the modelling from perceptual decisions. Diffusion and race models with termination thresholds have been used to describe this stage in value-based choice [10]. Further innovations to the standard accumulator models have been added to incorporate complementary dynamics during choices: e.g., how visual attention is deployed during deliberation. These models assume that gaze allocation boosts the accumulation of evidence [33, 34, 35]. For more details, please see chapter 2 on General Methods. Other features of the action selection, such as conflicts between valuation systems and the consideration of uncertainty estimates, among other issues, have been considered as part of the bigger ‘control assignment’ problem [32].

The assessment of the outcomes has been explored as well. Areas of the prefrontal cortex in humans or dorsal anterior cingulate cortex (ACC) in monkeys could be associated with this type of processing. Related to the learning stage, most of the studies focus on the habitual system. Learning has been formalised through the concept of the reward prediction error (PE) from reinforcement learning, which has been supported by repeated studies in animals and humans [36, 37, 38, 39, 40]. The main intuition is that after observing outcomes the agent can estimate the PE, which contrasts the preliminary forecast of reward (expected value) and the actual outcome experienced. This assessment changes the valuation (and choice) to reduce...
the error over time, allowing the animal to learn to assign proper values to their alternatives.

1.3 NEURAL CORRELATES OF DECISIONS

The characterisation of neurons or brain areas that correlate with the decision variables must fulfil the requirement of not tracking a purely sensory or motor response. Perceptual experiments in monkeys found that neurons in the lateral intraparietal area (LIP) could be candidates to encode this variable \[16, 41\] (Figure 1.4). After the stimuli onset, the firing rate of LIP neurons was found proportional to the strength of evidence and choice in a random dot motion task. Even more, the rise in average neuronal firing (build-up rate) was also proportional to motion strength, which was interpreted as connected to the integration of evidence in time. These findings suggest that the integral of the sensory signal could be encoded by the neural activity, which has been shown to start after \(~200\) ms of the motion onset, a long time for a purely visual signal. The neuronal firing reaches a stereotyped level of \(~85\) ms before the response, which has been associated with the decision variable reaching an accumulation bound. It has been proposed that alternative populations may be encoding the different options involved in the decision like in variations of a race architecture) which allows extending the mechanism beyond binary choices [42]. An important open question is how the activity of this neuronal circuit is adjusted to process the evidence relevant to each one of the multiple tasks that animals need to perform in natural environments [11].

In value-based decisions, the finding of dopaminergic firing in response to reward or reward prediction has been key to characterise the association between options/actions and value [37, 43]. Studies in the human brain have shown that areas such as the ventromedial prefrontal cortex (vmPFC) and orbitofrontal cortex (OFC) are associated with the computation of subjective preferences, even a “common currency”, across decision types [44, 45, 46, 47, 48]. Experiments in monkeys have confirmed the relevance of OFC regarding the computation of the subjective value of the options during the deliberation time [49]. In this study, OFC patterns representing each option appear alternatively in the neuronal population during the time leading to choice, with a predominant presence of the representation of the chosen option. Overall, a complex picture of value-based decisions has been proposed, including the integration of sensory and interoceptive circuits across areas encoding reward and affective values [50] (Figure 1.5).

While in perceptual decisions the increase in decision time can be directly connected to a longer sampling of sensory evidence, the situation in value-based choices is less straightforward. What is the evidence integrated into value decisions and from where is it coming? Recent perspectives have suggested that memory could be playing a crucial role in the sampling process, helping to rescue past events and experiences to construct value [51]. This line of research has shown the involvement of the hippocampus during the value-based decision, including scenarios where novel options need to be assessed [52] and a detrimental effect on the decision in patients with hippocampal lesions [53]. The interaction between regions involved in value computation and memory, i.e., vmPFC and hippocampus, has also been reported in decision paradigms [54]. Overall, these results suggest an important integration between these two cognitive modules to generate subjective preference decisions.
In perceptual and value-based decisions every choice is coupled with an internal level of uncertainty. Another advantage of the accumulator model is that it allows a mapping between the decision variable and the probability that the choice will be correct [11]. This means that in the brain, the internal assessment of the decision can be translated into a sense of certainty or confidence in the choice. In complex environments, decision confidence can play a pivotal role in guiding behaviour, assessing past choices to inform future decisions, or in changes of mind [35, 56, 57, 58]. Additionally, having a measure of confidence can be very important to communicate uncertainty in social decisions [59]. The study of decision confidence can be traced to
1.4 Decision and Confidence

Figure 1.5: Organisation of critical steps for value-based decision in the human brain cortex. Tier 1 operations compute and represent the presented stimulus or object at a perceptual level, e.g., visual features such as colour or shape across the primary visual cortex. The processing of “what” a stimulus is, involves multiple sensory modalities. In Tier 2 processes, the value of the stimulus is represented, which can include the reward or affective value of the objects. Finally, in Tier 3, choices and actions are characterised based on the value assessments. From [50], with Elsevier permission.

The early beginnings of experimental psychology (e.g. [60, 61, 62]). Most studies on metacognitive capacities have explored well-defined perceptual decisions, taking advantage of the systematic characterisation of choices in psychophysics [56, 63]; other studies have focused on confidence in memory capacity [64, 65]. Further studies have demonstrated that decision confidence follows similar behavioural signatures in value-based choices [66, 67]. Some of these signatures of confidence show that higher confidence is associated with higher choice accuracy. Confidence in easy decisions has been found higher in correct trials, while confidence was low in easy and incorrect trials. Faster response times have been associated with high confidence in decisions without time constraints [68].

Confidence studies are difficult to approach and are mostly confined to humans because it relies on subjective measures and verbal reports. However, studies in other animals using paradigms such as post-decision wagering [23, 69] have been very important to give further support to the research on confidence. In this experimental design, animals can opt out of uncertain choices in favour of a “sure bet”, an option that will deliver a smaller reward but with 100% probability. When animals choose the “sure bet” option, it
is associated with low confidence. Otherwise, when the animal is willing to wait for a higher magnitude reward, it is assumed as a high-confidence response. In these animal experiments, food pellets or water is used as a reward. Similarly to humans, these animal studies show some of the standard signatures of confidence, e.g., higher confidence is associated with better accuracy.

A big part of the literature on decision confidence is connected to measuring participants’ metacognitive skills. This refers to the capacity of the agents to correctly assess their own decisions, i.e., report high confidence when a correct choice has been made, or low confidence when the response was an error. Extremely useful metrics, such as meta-d’ [70, 71] have been proposed to characterise this introspective skill. In this work, I did not focus on this aspect of decision confidence, however, it is very important for its clinical and social implications [3, 72, 73].

Various models have been suggested to characterise the generation of decision confidence [74, 75]. The classical proposal by [76] formalises confidence as a measure extracted from the accumulator models at the endpoint of evidence integration. In this scenario, models with multiple accumulators such as the race models, allow us to estimate the difference between each accumulator, which can be used as a representation of confidence. The larger the difference between accumulators the higher the predicted confidence is. Other models, under the same framework of evidence accumulation, propose that confidence is generated from the integration of post-decision evidence, i.e., after choice, evidence is further integrated into confidence reports [68]. More recent proposals using Bayesian frameworks highlight that confidence can be described as a second-order inference process [74]. In this proposal, the evidence used to make the choice is fed to a distinct module to generate the confidence signal, akin to assessing the performance of another actor. Alternatively, this view defines as first-order models those in which decisions and confidence estimates are generated from the same internal state. In this work, I focused on first-order models, although allowing some departures from the standard formalisations (please see chapters 7 and 8 for more details).

Finally, the study of the neural substrate of confidence was initially based on metamemory studies in patients with Korsakoff’s syndrome. This disorder is characterised by anterograde amnesia as a result of alcohol abuse and nutritional deficiency, damaging areas such as the OFC and thalamus. It was found that the feeling of knowing, a metacognitive measure, was specifically affected in Korsakoff’s relative to amnesiac controls [77]. Further studies confirmed the role of frontal structures in metacognitive deficiencies [65], including regions such as vmPFC [78]. In psychophysics, the use of retrospective confidence reports and neuroimaging, lesion and neuroanatomical studies, have also supported the role of frontal areas in metacognitive processing. Areas including the dorsolateral prefrontal cortex (dlPFC) and rostrolateral prefrontal cortex (rlPFC) have been found relevant for confidence computation [79, 80, 81]. Value-based decision experiments have also reported the involvement of PFC in confidence processing, including coding of value and confidence in the vmPFC, and a metacognitive read-out in rlPFC [66] (Figure 1.6). More recent studies have reported that confidence in humans can indeed capture a probabilistic distribution of uncertainty, which can be found in neural populations in the PFC (plus insula and ACC) [82].
1.5 Goals and contexts

It is common in decision experiments to consider simple scenarios, with single and well-defined tasks, and rewards delivered directly by the environment. However, in real-life settings, rewards are obtained after agents identify the characteristics of the environment and perform the required actions to fulfil specific demands. The scenarios in natural decisions are varied and unstable, requestsing a varied repertoire of rules and actions from the agent. For example, if we are hungry, we may take a banana to eat, however, if the fruit is rotten, we would throw it away immediately. It is not obvious that animals have a specific channel for the tracking of rewards, which means that value has to be inferred by the agent from sensory information, i.e., extracting the proper context and goals to obtain rewards [83, 84].

[1] characterise goal-directed behaviour, in contrast to simple habitual behaviour, as having two main requirements: 1) the agent shows knowledge of the causal efficacy of its actions and their results given the current state or context, and 2) the agent selects and regulates its behaviour using goal representations. In other words, specific goals are driven by the agent’s external (e.g. being stranded in the desert) or internal contexts (e.g. being thirsty), motivating specific value representations (e.g. water becomes important), and guiding actions (e.g., choosing to search for water instead of resting). The crucial aspect of goal-directed behaviour (as previously mentioned in the value-based section) is that it is not a stereotypical or prescribed procedure, but it allows flexibility pointing towards the desired end-state.

Memory researchers have been interested in the study of context, given the relevance that it has to discriminate or generalise past experiences, informing the suitable course of action to fulfil the goals. Classical experiments on the topic consider contextual fear conditioning in mice [85]. In this case, animals learn that aversive scenarios (e.g. electric shocks) are associated with distinct cues (e.g., contexts created by light or sounds), which makes them display fear-specific behaviours such as freezing. Eventually, the animals

![Figure 1.6: Confidence-related activity in rIPFC. A) Brain activity in right rIPFC was correlated with decreases in subjective confidence (significant activity displayed, P<0.005 small-volume corrected). B) Signal extracted from an ROI in rIPFC (6-mm MNI space sphere) characterised a main effect of confidence but it was not affected by the difference in value (DV or difficulty of the choice) in a binary value-based choice (food items). From [66], with Springer Nature permission.](image)
react immediately in response to the contextual cues, even when the aversive stimulus is not presented. More recent studies have been conducted featuring complex, feature-rich spatial contexts [86, 87] (Figure 1.7). These studies have highlighted the relevance of the hippocampus in the organisation of experiences, following a hierarchical structure where context is the dominant source of variance. Even when these studies have been performed in scenarios where the spatial component is relevant, it has been proposed that hippocampal operations play a more general role, involved in an abstract relational mapping of objects and actions [88, 89]. These maps, including objects or episodic memories, are constructed and organised around contexts [90]. Furthermore, context can be defined not only by space but also by perceptual regularities, time or reward contingencies [91, 92, 93, 94]. [95] proposed three tenets for the formalization of contexts: 1) context should be constant along a certain experiential dimension (e.g., time), 2) context should be relatively complex and adaptable, 3) context should have behavioural relevance (overt or incidental).

![Figure 1.7 Context organization of neural activity in rats. A) Experimental design in a study conducted by [86]. In this task, rats discriminated between two items considering context and item locations to obtain a reward. Contexts were generated using different colours and textures in the experimental rooms. Objects were terracotta pots with distinct aromas and digging media. The reward was a small bit of cereal buried in one of the pots. The context generated variations in the object’s reward valence. Different object pairs (A-B and C-D) were tested in the same two contexts. Symbol + indicates the rewarded alternative. B) Hippocampal neurons show activity consistent with a hierarchical model, with item, position and valence nested within context representations. Adapted from [86], with Elsevier permission.]

When contexts are reencountered, the associated task demands can be retrieved in a process supported by the PFC, allowing proactive control [96]. [96] reported that contextual identification can be supported by hippocampal processing, such as pattern separation, which predicts a reinstatement associated with task demands in the dIPFC. The retrieved knowledge is more than the “what” (e.g., sensory information or semantic knowledge), it also incorporates “how” to perform the tasks (e.g., how information is processed, how attention is allocated and the mapping from evidence to action). The influence that context has on the strategies used to make task-relevant choices is a developing area of research with questions that have inspired recent work in decision-making. For example, the processing of perceptual evidence in context-dependent choices has shown prefrontal networks involved in the filtering of task-relevant evidence [97]. New conceptualisations of value-based decisions have reinterpreted value as the distance to goals [83] or the usefulness of the options for a specific goal [98, 99, 84]. The latter type of study has found that when including multiple contexts and goals, vmPFC, the region related to hedonic value in standard decision studies, actually encodes how task-appropriate the alternatives are. A recent perspective has highlighted the relevance of including cognitive control (i.e., how people
1.5 GOALS AND CONTEXTS

adjust their information processing to fulfill tasks, focusing on target features and ignoring distractors) in the exploration of value-based decisions [98]. This thesis follows this lead, exploring how goals affect various stages of the decision process.

In many of the following studies, I used a simple experimental manipulation to modify the contexts and goals: the framing effect. Decision rationality, which is based on the fulfillment of principles such as consistency or coherence, is broken by the context and the way information is presented [100]. For example, if participants receive £1000 and then are given the option to keep a sure gain of £250 versus gambling with a 70% probability of losing all, most responders choose the former option. On the other hand, when participants are given £1000 and then presented with the option of a sure loss of £750 versus a bet on a 70% chance of losing all, most of the choices favour the gamble. In this case, both choices are rationally identical, and from the experimenter’s perspective, the presented changes should be inconsequential. However, participant responses demonstrate changes in the decision process: while they show risk aversion in the gain context, they are risk-seeking in the loss context. This context effect on choices can be found at an individual level, reflecting variations in brain activity in areas such as the amygdala or the OFC [101]. In this thesis, I interpret that these frame manipulations elicit different goals, causing the reported behavioural variations.

In Chapter 2 of this thesis, I present an overview of the main experimental methods I used in the conducted studies. The experimental section of this thesis is divided into three parts. In the first part, my experiments showed that from the inception of the decision process, goals change the way information is integrated, serving the objective of the decision. In Chapter 3, I used eye-tracking and computational modelling in value and perceptual decisions to show how visual attention is oriented towards goal-relevant and not merely to the “rewarding” alternatives.

The second part of my work explores how brain processing (measured using functional Magnetic Resonance Imaging, fMRI) is affected by goals during decisions. In Chapter 4, I studied how learning efficient strategies to fulfill goals is guided by value signals. I show that the neural patterns reorganize during learning, representing features that are recognised as relevant to fulfill the task. Additionally, fronto-occipital networks in the brain may play a role in facilitating goal-relevant feature integration. In Chapter 5, I explore how the change of goals reconfigures item representations, overcoming context invariance mappings that only consider sensory features or hedonic rewards. I show that in the value-based assessment of options, brain representations are informed by the goal, and then by other perceptual features.

In the third section, I explore the impact of goals in post-decision processing. Together with every decision comes a sense of confidence, an internal estimate of our level of uncertainty, fundamental to guide learning or trigger changes of mind. In Chapter 6, I exemplify the relevance of confidence for the decision process, via the study of confirmation bias and changes of mind. I show that confidence seems to guide future sampling behaviour to fulfill the goal of confirming previous choices. In Chapter 7, I complement the study presented in Chapter 3, showing the impact that frames have on the generation of confidence, and how this is related to the integration of evidence. Chapter 8 is a study that deepens on the findings presented in Chapter 7, revealing that the confidence signal outweighs the evidence supporting the
goal-relevant evidence. This chapter proposes a model plus further studies that support the view that asymmetries in the integration of evidence could cause this effect on confidence.

Finally, in chapter 9 I present a general review of the studies, discussing the main implications of the findings. Overall, in my studies using behavioural testing, computational models and neuroimaging, I show that from pre- to post-decision operations, in the perceptual and value-based domain, the choice process is strongly influenced by goals, demonstrating the flexibility that the brain deploys every time we make a decision.
GENERAL METHODS

2.1 SUMMARY

In this chapter, I will present some of the methodological approaches used in the following studies. Firstly, I describe regression analysis, particularly focusing on linear and logistic regressions. These methods are useful to characterise the relevant factors predictive of selected variables, such as choice or confidence. Secondly, I present a series of models used to characterise the decision process. I begin presenting signal detection theory (SDT) and sequential sampling models, including drift-diffusion models. I also describe the basics of probabilistic graphical models, a versatile family of models with multiple applications, which I use to model the pre- and post-decision processes. I present the principles of reinforcement learning (RL), a powerful model to characterise how animals and humans acquire the knowledge to make value-based decisions. Finally, I introduce functional Magnetic Resonance Imaging (fMRI), the neuroimaging technique used in this thesis. fMRI is a non-invasive technique used to infer brain activity by detecting changes associated with blood flow. I present some of the univariate and multivariate methods used for the analysis of fMRI brain volumes, including the pre-processing of these datasets.
2.2 Regression Analysis

Regression analysis is a basic, though very powerful tool for data analysis. It allows us to examine the relationships between two or more variables of interest. It focuses on the association between a dependent or predicted variable, and one or more independent variables or predictors. For example, we may want to predict the amount of rainfall in London during January. For that purpose, we can use different measurements we may think relevant as predictors, e.g., the temperature during the day, the hours of sunshine, humidity in the air, or the number of buses crossing Waterloo Bridge. We can collect data about all those measurements for January of the last hundred years and use them to estimate a regression model. The results of this regression analysis can tell us if there is a relationship between the variables: e.g., humidity in the air could be very relevant to forecast the rain, while the traffic on Waterloo Bridge could be less informative. We can also have an idea of the magnitude of these effects: e.g., a drop of 5° in temperature may predict an increase in 10mm of rainfall. There are many types of regression analysis but one of the most extended is linear regression (Figure 2.1). This regression model summarises how the average values of the predicted (also called the outcome) variable, vary over subpopulations of the data defined by linear functions of the predictors. This can be expressed using the following equation for the case of a single predictor:

\[ Y = \beta_0 + X\beta_1 + \epsilon(1) \]

with Y a vector of the outcome of interest, X a vector with the respective values of the predictor, and \( \beta_0 \) and \( \beta_1 \) are the free parameters of the model. \( \beta_0 \) corresponds to the constant term of the prediction, also called bias, offset term, or the y-intersect of the linear equation. \( \beta_1 \) corresponds to the regression coefficient associated with predictor X, also called the slope of the linear equation. The value of \( \beta_1 \) is informative of the effect or contribution of that variable over the outcome: if \( \beta_1 > 0 \), an increase in X will predict a higher value of Y; if \( \beta_1 < 0 \), an increase in X will predict a lower value of Y; if \( \beta_1 = 0 \), it means that change in the value of factor X will not affect the value of Y. In other words, the regression parameter represents the change in the dependent variable per unit of change in the independent variable. For this reason, the unit of measurement of the variables is relevant for a correct interpretation. The term \( \epsilon \) corresponds to the deviation of Y from the predicted values by the linear model and it is called the error. Another way of interpreting the simple linear regression is as Y having a distribution with mean \( \mu(X) \) for any given value of X, with \( \epsilon \) capturing the shape of the distribution. The univariate case can be generalised to include multiple predictors:

\[ Y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \ldots + X_n\beta_n + \epsilon(2) \]

Where n is the total number of predictors in the model. Given that the effect of the parameters depends on the units of the specific measurements, some problems in the interpretation of these models can arise when a correlation between the predictors is found. Normalisation of the variables (e.g., using standard scores) can help to interpret and compare the effects of multiple factors.

The proper use of regression models relies on a series of assumptions:
2.2 Regression Analysis

Figure 2.1: Relationships between variables and regressions. A) If the properties of Y do not change with X, no association is found. It is possible that no regression is found, even when other types of associations are present (e.g., $E(Y|X)$ is constant, but the variance of $Y$ increases with $X$). Linear regression can be found if the mean of the predicted variable changes in a linear fashion with $X$. Nonlinear relationships between the variables can also be found, although transformation have be applied to adjust for the use of linear regressions. B) Example of linear regression with the predicted variable in the Y axis and predictor in the X axis. The mean value of the predicted variable (Precipitation) can be estimated by the regression line. From [103], with Springer Nature permission.

1) Validity: the data analysed should map to the research question. The outcome measure should reflect the phenomenon of interest and the model should include all the relevant factors. The regression model should generalise to the cases it will be applied (e.g., a model of exam scores is not necessarily informative of child intelligence). This step can be the most challenging for regression analysis.

2) Additivity and linearity: the model assumes its deterministic component is a linear function of the separate predictors $Y = X_1\beta_1 + X_2\beta_2 + \ldots + X_n\beta_n$. In the case linearity is violated, transformations of the data can be used (e.g., a predictor may be added to the model as $\log(x)$ instead of $x$). If additivity is violated, it is possible to add interactions between the predictors (e.g., $X_1X_2\beta_n$).

3) Independence of errors: the simple regression model assumes errors from the prediction line are independent.

4) Equal variance of errors: in the case of unequal variance of errors, some adjustments can be made to have an efficient estimation (e.g., using weighted least squares to find the regression line). However, unequal variance does not affect the most important aspects of the model (e.g., the form of the predictor $X\beta$).

5) Errors are normally distributed: according to [102] this is the least important of the regression assumptions for the estimation of the regression line, but it may be relevant for the prediction of new data.

The most common way of estimating the parameters $\beta$ is using least-squares estimators [103]. This involves the minimisation of the residual sum of squares term (SSE). In the univariate case, this is expressed as:
\[ SSE = \sum (y_i - [\beta_0 + X_i\beta_1])^2 = \sum (\hat{y}_i - y_i)^2 \] (3)

Note here the subscript \(i\) indicates each element in the vector of outcomes or predictors. Commonly, the term \(\beta_0 + X_i\beta_1\) is interpreted as the estimated (or fitted) values by the regression model (\(\hat{Y}\)).

Standard linear regression has been proposed to deal with continuous variables. In some scenarios, like binary choices, the variables of interest are discrete, and they do not fulfill some of the assumptions of the regression model, such as the normal distributions of the error. Logistic regression is an extension of the linear regression models, which considers binary outcomes as predicted variables [105]. The logistic prediction, in addition to indicating the value of the variable (e.g., it will rain or not today), can also estimate the probability of it (e.g., there is an 89% chance it will rain today). The output variable is coded as 0 and 1 to express the two categories. To bypass the normality assumption, the model predicts the probability of \(y_i = 1\) (Pr\(y_i = 1\) = \(p_i\)), instead of \(y_i\) directly. In this way, the model of probability can be expressed as:

\[ Pr (y_i = 1) = \logit^{-1}(\beta_0 + X_i\beta) \] (4)

The logit function (logit\(^{-1}\) (\(x\)) = \(\frac{e^x}{1 + e^x}\)) transforms the values to a range \((0,1)\) to adjust to the values of the probabilities. The term \(X_i\beta\) corresponds to the linear predictor, which is a continuous variable with range \((-\infty, \infty)\). This last expression can be rewritten as:

\[ Pr (y_i = 1) = \frac{e^{\beta_0 + X_i\beta}}{1 + e^{\beta_0 + X_i\beta}} \] (5)

Which is equivalent to:

\[ \log \left( \frac{Pr (y_i = 1|X_i)}{Pr (y_i = 0|X_i)} \right) = \beta_0 + X_i\beta \] (6)

In other words, the logistic regression can be interpreted as a regression to predict the probability of the predicted variable taking one of the two values, represented as an odds ratio. The estimation of the regression parameters does not have an analytical solution, unlike linear regression, therefore numerical optimisation is required [105]. One of those approaches is to use maximum likelihood estimates (MLE), which searches iteratively the pair of parameters \(\beta_0\) and \(\beta\) that maximise the likelihood of the observed data (or equivalently, that reduce the negative log-likelihood) (Figure 2.2).

An important consideration when performing regression analysis is how data will be organised in the model. In real-world situations, it is common that the data can be arranged in various groups: e.g., within a country, we have multiple regions with unique weather which will affect the precipitation measurements, the grades students get in a national math exam could depend on the high school they attend, or responses in a cognitive test depend on each individual participant. A standard approach to this issue is to just pool together all the possible groups and fit a single regression line that is representative of the whole dataset (e.g., join all the participants and trials into a single model). An alternative is to individualise each level, and fit as many models as groups (e.g., fit a linear regression for each participant independently). Both approaches are suboptimal since they lose relevant information available in the data structure: for the pooled model, there is a conflation of within and between group effects; for the unpooled model,
the analysis is underpowered treating each group as independent when commonalities may be present. Therefore, for the analysis of nested data, a useful approach is to use *hierarchical or multilevel regressions* [102]. The main modification relative to the standard regression approach is that in this case each group has its own regression parameter, also called mixed or random.
effects, which are constrained by an overall parameter. Parameters that do not vary at each group are denominated fixed effects. For example, if we want to estimate a hierarchical regression analysis predicting confidence in a cognitive task ($y$) using as a predictor the reaction time in each trial ($x$) we can formalise the model as:

$$y_i = \beta_0 + x_i\beta_{1,j[i]} + \epsilon_i, \text{ for } j = 1, \ldots, n$$

Where $j$ corresponds to participant number and $i$ represents each trial. Therefore, $j[i]$ expresses the participant $j$ that generated each trial $i$. In the next part of the model, the value of the slope coefficients for the participants ($\beta_{1,j}$) depends on a common probability distribution:

$$\beta_{1,j} \sim N(\mu, \sigma^2)$$

In this model both parts are fitted simultaneously, making that the parameters at the subject level are constrained by the overall distribution. In this way, we can estimate subject-level slopes, i.e., we have as many regression lines as participants, still considering a dependency between them as representative of a common ‘human’ sample. In the example we considered that the slope varied at a participant level, however, the intercept and any other extra predictor’s coefficients can be modelled hierarchically as well. In conclusion, hierarchical modelling is a powerful extension to the standard regression framework, taking advantage of the underlying structure in the data to generate informative models, making it suitable for the analysis of nested data in human experiments.

### 2.3 Decision Models: Signal Detection Theory

Signal and noise are two fundamental ingredients in any decision. If in a decision experiment the objective is to identify the presence or absence of a single stimulus among noise (null stimulus), that type of task is called detection. The original proposal of signal detection theory (SDT) was to improve the radar identification of enemy bombers during the Second World War. Eventually, the application of SDT was expanded to psychology, with the initial objective of characterising decisions in psychophysics experiments [13, 106].

In SDT the determination of each state has uncertainty on its own (e.g., an ‘it is raining’ state can consider a soft drizzle or a dramatic downpour). In the basic detection task, the presence of noise and signal entails uncertainty in the decisions. The SDT proposal considers that the signal is constant (e.g., an auditory tone with constant volume and frequency) and superimposed on noise (e.g., white noise). The objective of the task is to identify whether the presented stimulus contains the signal or if pure noise is presented. Therefore, SDT can be used to distinguish between two states (e.g., it is raining / it is not raining; there is a tumour in the scan / no tumour; there is a red light / no light). Pure noise stimulus can be represented as a probability distribution, meaning that the signal-plus-noise stimuli is also characterised as a distribution with a curve identical in shape to the one describing noise alone, but shifted depending on the evidence strength of the signal (Figure 2.3). One of the main assumptions in SDT is that these distributions (signal and noise) are Gaussian with equal variance. Once in the task, there are two possible types of trials, ‘signal’ and ‘noise’ trials, with two possible responses, ‘yes’ or ‘no’ to the presence of the signal. When the signal is indeed presented
and the response is ‘yes’, that trial is called a hit; if the response is ‘no’, it is called a miss. When it is a noise trial and the response is ‘yes’, it is termed a false alarm; if the response is ‘no’, it is a correct rejection. The basic SDT analysis can be performed using those four categories (or alternatively using hits and false alarms plus the total number of signal and noise trials). As presented in Figure 2.3, the four categories can be associated with specific sections of the signal and noise distributions.

The overlap between noise and signal distributions will depend on the evidence strength of the signal (Figure 2.3). The horizontal axis in SDT graphical representations displays the strength of evidence in each category, with greater strength on average for the signal distribution. In the previous example, evidence can be considered as the sound volume, with greater volume when a tone is present. The largest is the difference between the distributions, higher is the sensitivity or discriminability in target detection. Formally, the model characterises this as the value $d'$, the separation between the averages of each distribution.

Having this representation of two states, how can we determine if a new sample contains the target signal? A simple approach would be to set a value of the signal strength that will work as a criterion (c): if the evidence strength of the new sample is above c, then it is categorised as containing the signal, otherwise, it is noise. Under the assumption of equal variance, the distance $d'/2$ is important since placing the criterion at this point maximises the probability of correct identification of the signal, since here, signal and noise samples are equally likely to happen. This corresponds to the neutral or ‘unbiased’ criterion. However, the value of c can be lowered to favour a higher hit rate (a more liberal criterion); if c is raised, the correct rejection rate increases (a more conservative criterion) (Figure 2.3b). These biases generated by the placement of the criterion can be relevant in specific decisions, e.g., a false alarm can have terrible consequences in the control of a nuclear weapon, therefore a more conservative criterion may be used.

![Figure 2.3: Signal detection theory. A) In SDT two distributions are used to identify whether the observed stimulus contains a signal, or if it is only noise. Those two distributions are standardly assumed to have equal variance. A criterion can be set to report as a signal if the perceived stimulus is above that threshold and noise if the perception is under it. B) Changing the position of the criterion can generate a liberal (responding ‘signal’ in most of the cases but increasing the number of false alarms) or conservative response (responding ‘noise in most of the cases, but increasing the number of misses). From [107] with Springer Nature permission.](image)
Instead of the direct evidence strength, an alternative way of discriminating between the two states is using another decision variable: the likelihood ratio \([106]\). Since the probability distributions for noise and signal are known, it is possible to calculate the likelihood of a new sample for each one of them. The likelihood corresponds to the height at the location \(x\) by the distribution \(f(x)\). Therefore, the ratio between the likelihood of the new sample \(x\) is signal or noise is:

\[
LR(x) = \frac{f(x|\text{signal})}{f(x|\text{noise})}(9)
\]

For each point \(x\) a value of the likelihood ratio can be calculated: \(LR = 1\) when the new sample is located between the two distributions; \(LR > 1\) when it is located towards the signal distribution; \(LR < 1\) when it is located towards the noise distributions. Setting an \(LR = 1\), allows an unbiased criterion. In the normal distribution model, the height of the likelihood function depends on the distribution’s mean \((\mu)\) and standard deviation \((\sigma)\):

\[
f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)}(10)
\]

Since the LR is a ratio, an equivalent form using logarithms (log-likelihood ratio) is employed to express the division as a subtraction. Further equivalencies have shown that there is a relationship between LR and the multiplication of criterion (\(c\)) and sensitivity (\(d’\)) \((LR = e^{cd’})\).

To assess the accuracy of the observer (characterised by the signal and noise distributions) it is common to use the receiving operating characteristic (ROC) curve. This curve shows the false-alarm rates plotted against hit rates for all the possible values of the criterion, but with constant \(d’\). In the present work, ROC analysis was not relevant, however, it is a crucial part of the classic SDT approach \([106, 107]\).

While the basic SDT is constructed for single stimulus decisions, it can be extended to two alternative forced choice (2AFC) tasks. In this case, the observer does not report whether a single stimulus occurred or not, since the stimulus could be presented in any of two alternatives. 2AFC task requires a comparison between the two alternatives to identify which one contains the signal. An example of this type of task is a memory experiment where various words are presented in an initial phase. Then the participants observe pairs of words, and they report which one has been already shown and which one is a new word \([106]\). Using the previous example, a task where a listener is asked to report the side where an auditory tone is presented (i.e., right or left ear) would also be considered a task of this type. In 2AFC the notation of hits and false alarms needs to be repurposed to express the possibility the choice was correct or incorrect (i.e., the selected option did or did not contain the signal, respectively). It is assumed that the observer estimates the evidence for each option independently, meaning that the alternative can be treated as a separate dimension in the decision space. Since now the distributions correspond to surfaces over a 2D space, they can be represented as circles of equal likelihood in the plane (Figure 2.4). In the 2AFC task, as in the detection case, the characterisation of the distance between signal and noise distributions is denominated as \(d’\). Arbitrarily, the mean of the noise distribution can be set equal to zero. Therefore, the observations in a 2AFC task trial compare the possibility that option A (represented in the horizontal axis of a 2D plot, Figure 2.4) or B (represented in the vertical axis of the 2D plot) contain the signal. Option A for example could be the
option presented on the right side and Option B on the left side. When the option with the signal is A this can be represented in the 2D plot as a distribution centred around the coordinate (d',0); and (0,d') when the option with the signal is B. The decision axis can be defined as the line between the average of the two distributions (segmented line in Figure 2.4) and the decision boundary (criterion) as a perpendicular line. The 2AFC can be expressed as a unidimensional problem, along the decision axis (details can be found in [106]). An alternative way to express the decision variable is using the LR adapted for the two alternative scenarios:

\[ LR(x) = \frac{f(<A,B>)}{f(<B,A>)} \] (11)

In this case, the equation is adapted for the two alternatives considering the comparison of the two possibilities: the likelihood that option A contains the signal and option B is noise vs option B contains the signal and A is noise. In this thesis, log-likelihood ratio in a two-option scenario was used as the basis of the model presented in Chapter 8. However, in our model, we altered one of the main assumptions in SDT: equal variance of the probability distributions. In our model we repurposed ‘noise’ and ‘signal’ categories to fit the categories in our discrimination tasks in perceptual and value-based binary decisions. We modified the variance of the observer distributions for the two categories, finding that an increase in the variance of the goal-relevant distribution generates specific behavioural effects on simulated decision confidence. More details on the formalisation of our model can be found in the Methods section of Chapter 8.

In this section, I have not presented important work oriented to expand SDT to quantify the observer’s metacognitive capacities (meta-d’, [108, 74]. In these studies, participants report their confidence level after making their choices. These reports are created to construct an additional SDT for type-2 performance, i.e., that classifies whether participants could predict if they were correct or incorrect in their choices. The characterisation of individual or group metacognitive capacities was beyond the scope of the studies presented in this dissertation.

2.4 Decision Models: Sequential Sampling Models and Attention

Most of the insights we have on the decision process have been obtained from binary choice experiments. In this case, the usual and straightforward measures of performance are the probability of choosing one option or the other (which allows estimating the accuracy) and how long it takes to deliver the answer (which constitutes the response time, RT). Using SDT we can model the choice behaviour; however, the timing of decisions cannot be captured by this approach. Sequential sampling models are an expansion of the decision models to include this dynamic aspect. This allows us to model additional phenomena, such as the speed-accuracy trade-off [17] or the response time distributions [109]. Sequential sampling models can be of different varieties, but the basic building blocks include at least one accumulator, an assumed structure that gathers evidence in favour (or against) the decision alternatives, and a decision rule that indicates the moment the accumulation ends and the decision is made. A detailed account of different types of Sequential Sampling Models can be found elsewhere (e.g., [110, 20]). In the present work, I will focus on the drift-diffusion model
Figure 2.4: Two-dimensional interpretation of the 2AFC task. The space characterises the decision evidence (observation strengths) for the alternatives. Distributions are represented with concentric circles that correspond to contours of equal likelihood for the possibility the stimulus is present. The decision boundary can be defined as a perpendicular to the line that connects the mean of the two distributions. \(<A,B>\) indicates that option left contains the signal, and \(<B,A>\) that the right option contains the signal. Therefore, the observer responds left option for the region 'under' the boundary, and right option for the region 'above' the boundary. The segmented line indicates the decision axis.

(DDM) and its derivations that integrate the effect of visual attention in the accumulation process.

The standard DDM \([111, 112, 113]\) assumes that decisions are made by accumulating stochastic information over time. In this model, the accumulation of information begins at a starting point and continues until the total amount of information reaches a predefined boundary. In the DDM, the single accumulator is usually thought of as a particle that can move towards one of two boundaries (superior or inferior boundary in Figure 2.5a), one for each option (e.g., it could be a Yes/No question or right vs left alternative), which will trigger the choice. The response time in the model corresponds to the time required to reach the decision boundary. The model also can include a non-decision time parameter (\(t_0\)) which considers the constant encoding and response-execution time. The rate at which the accumulation approaches the boundary, i.e., the average amount of information accumulated per unit of time, is captured by the drift rate parameter (\(v\)). The accumulation is not constant but varies depending on a noise term, usually obtained from a normal distribution with standard deviation \(\sigma\), which is another parameter of the model. The model considers two boundaries, with the separation between them usually included as an additional parameter (\(a\)). The boundary separations can characterise specific decision-making styles. A high value of
the parameter $a$ is related to more conservative decisions, that require more evidence to be collected, with longer reaction times and fewer errors; a low boundary separation is associated with a riskier style with faster response times and a higher error rate. The starting point of accumulation ($z$) is another parameter that can capture decisions biased towards one of the options, e.g., if there is priming towards one of the alternatives.

**Figure 2.5:** Drift diffusion models (DDM). A) Visualization of the standard DDM parameters. B) The attentional DDM. The main difference of this model with respect to the standard DDM is that the drift rate is modulated by the alternative that is being fixated, generating a bias in the accumulation of evidence towards the attended alternative. In this case of binary choice, the options are presented in distinct locations so visual fixation displacements can inform attentional variation. The inset shows the example of a value-based experiment where two snacks are presented in a choice trial.

An important feature of the DDM is that for the measurement of the relative evidence in the binary case, the evidence accumulation for different options is anti-correlated, i.e., an increase in evidence towards one option means a decrease in evidence towards the other. This is the main difference
with race models, another type of sequential sampling model. In race models, each option is represented by an individual accumulator and a single boundary indicates when the decision is made [114].

DDM has been an important model for the characterisation of the decision process, extendedly used in perceptual, memory [110, 115] and value-based [116] studies, in healthy and clinical populations [117, 118, 119]. However, the observation of everyday decisions indicates the deliberation process is likely to be more sophisticated. Consider the case of a customer in a supermarket, deciding whether to buy chips or popcorn. Instead of approaching the shelf and immediately choosing their preferred snack, the customer would likely stop for a second and gaze repeatedly between the two items, shifting their visual fixations between options. This observation inspired an extension to the DDM to include visual attention, the attentional DDM (aDDM; [34, 33]). The standard DDM assumes fixations do not play a role in the accumulation process given that most of the decision experiments keep central fixation during the experiment. Therefore, the experimental design of the aDDM considers a presentation of the options that allow for variations in spatial attention with a displacement of visual fixations towards the available alternatives. The aDDM takes the period of distinct attentional allocation to introduce temporary modulations of the drift rate, which supports the accumulation of evidence towards the fixated item (Figure 2.5b). In other words, it generates a choice bias towards the items that are fixated on more, which is in line with behavioural findings [120, 34]. Therefore, eye-tracking information is required during the experiment to insert the attention component in the model.

The aDDM assumes that the brain computes a relative decision value that evolves over the decision time as a Markov Gaussian process until one of the boundaries is reached, in a similar way to the DDM. The main differences are in the expression used to estimate the variation in the accumulators. In every timestep, the drift rate is modulated by fixated and non-fixated items. The accumulation process can be characterised by two expressions:

\[ X_t = X_{t-1} + d(r_{\text{left}} - \theta r_{\text{right}}) + \epsilon_t \]  
\[ X_t = X_{t-1} - d(r_{\text{right}} - \theta r_{\text{left}}) + \epsilon_t \]

The first equation describes the accumulation when the item on the left side is fixated, and the second equation is when the right-side item is fixated. The variable X corresponds to the relative decision value which is updated every timestep (t). The drift rate parameter defined in standard DDM is separated into two parts in the aDDM: (1) a constant parameter d that controls the speed of integration (units ms\(^{-1}\)) and (2) a term for the weighted difference between the values (r) of attended and unattended items. Since the original presentation of aDDM was developed for value-based decision, r corresponds to the subjective preferences for each one of the items, which usually is quantified as the amount of money participants would be willing to pay for the items [33]. This process can be adapted to quantify evidence in perceptual experiments, e.g., relative angles of lines relative to a target [121]. The second term includes the crucial attentional parameter (\(\theta\)) which takes values between 0 and 1, multiplying the unattended item evidence. This means that the value of the unattended item is discounted favouring the accumulation of the attended item. A white Gaussian noise term (\(\epsilon\)) with variance \(\sigma^2\) is randomly sampled every timestep.

The aDDM has been an important model to formalise the influence of attention on the evidence accumulation process, even for multi-alternative
choice [34]. Recent work has reported the relevance of the frontal eye field area in the human brain on the gaze bias, which means that gaze could indeed have a causal role in choice [122]. In the present work, I used a model that has been inspired by the aDDM, the Gaze-weighted accumulator model (GLAM, [35, 123]). GLAM implementation points to facilitate the gaze-choice association at an individual level. GLAM is strictly a linear stochastic race model, which facilitates its application in scenarios with more than two alternatives (Figure 2.6). Another important difference is that while aDDM focuses on the effect of individual fixations and their sequence within a trial, GLAM consider a summary of the gaze bias effects in a trial. This means that for GLAM the full sequence of fixations is averaged on each trial, avoiding the complex problem of modelling and simulating the fixation process (which can be even more difficult in multi-alternative scenarios). These modifications allow for computations that are analytically tractable which facilitates more complex applications, such as Bayesian parameter estimation. An additional advantage of GLAM for our specific research question was that the estimation of decision confidence in the model is facilitated using race models. In race models of binary choice, the balance of evidence approach [76, 124, 23, 66, 125] proposes that the difference between the two accumulators at the decision time (i.e., when one of the alternatives has reached the boundary) can be used as a measure of confidence. A higher difference between the winning and loser accumulator indicates higher choice confidence. Further details on the GLAM model implementation can be found in the Methods section of Chapter 3.

![Figure 2.6](image.png)

**Figure 2.6**: Gaze-weighted linear accumulator model (GLAM). A) The fixation time for each one of the options is extracted within the trial. B) Using the fixation time, the average proportion of time each item is fixated is used to estimate the average absolute evidence for each option. In this case, the value depends on a gaze bias parameter $\gamma$. C) This is transformed into relative evidence, which compares each item with the other available alternatives. Relative evidence is used in the equation for the accumulation of evidence for each alternative. Note that the accumulation of evidence includes in its value $\gamma$, the attentional discount. For more details on the model, please check the Methods section, chapter 3 (Adapted from [35], with Springer Nature permission).

### 2.5 Probabilistic Graphical Models

Models for complex systems need to include uncertainty, an unavoidable aspect of real-world applications. Uncertainty can be defined by various factors, some of them described as 1) unobservable aspects of the world (e.g.,
a patient’s true disease may not be directly identifiable, only symptoms); 2) our measurements and observations are per se noisy (e.g., there is a degree of error in a blood sample or a PCR test). The construction of models must include not only what is possible but also what is probable [126]. Probabilistic Graphical Models (PGM) use graph-based representations as the basis for compactly encoding complex probability distributions over a high-dimensional space. This helps to characterise problems with hundreds or even thousands of relevant features (e.g., various attributes like muscle pain or cholesterol level in a medical diagnosis) in a single probabilistic model. There are many types of PGMs, but for this thesis, I will focus on Bayesian Networks.

The core of the Bayesian networks is the representation of a directed acyclic graph (unlike Markov networks where cycles are allowed). Here, the variables of the model are represented as nodes (ovals) and the relationships between variables are indicated by the edges of the network. For example, we can create a model to characterise the probability a student will get a recommendation letter from a professor, captured by a variable \( l \) with two levels \{yes, no\} (Figure 2.7). In this case, we can consider that the professor is quite distracted and does not remember well all her students, so she will decide whether to write the letter or not by checking the student’s grade, a variable \( g \) with possible values \{low, medium, high\}. On the other hand, the grade depends on two random variables: student’s intelligence \((i)\) and course difficulty \((d)\), which can be characterised by two levels \{low, high\}. The conditional probability distribution that specifies the values of the variable given the possible arrangement of joint assignment values, derived from the parent nodes (i.e., the variables that affect the probabilities of each node) is shown in Figure 2.7. In this way, the model allows characterising the events from the probabilities of each one of the variables.

In the example above, the various elements of the model were defined as discrete variables, but the model can also be constructed using continuous variables. The example below shows a graphical model using continuous variables to capture the generation of confidence in a perceptual decision task, an application of Bayesian models in cognitive science [127]. In the modelled experiment, participants complete a random dot motion task where they should choose if most of the dots are moving towards the right or left side of the screen. The experiments consider two stages where participants observe the dot motion: pre-decision and post-decision, after which they report their confidence. This model (Figure 2.8) seems more complex, but the principles are similar to the model with discrete variables. In [127] the variables that capture the evidence accumulated in the pre- and post-decision stages \( X_{\text{pre}} \) and \( X_{\text{post}} \) are described as Normal distributions \( \text{N}(\mu, \sigma) \) of mean \( \mu \) and standard deviation \( \sigma \) that are informed by the true direction of motion \((d, -1 \text{ or } 1 \text{ for left or right, respectively})\), the coherence of the dot motion \((\theta_{\text{pre}} \text{ or } \theta_{\text{post}})\) and \( k \), a sensitivity parameter. Note that in the standard notation, the \( \sim \) symbol indicates that the variable “is distributed as”. The variable \( k \) demonstrates another important advantage of building up models: it allows to describe latent variables, i.e., factors that exist in the world but that we cannot observe or are not even aware of. While it is possible to have a measure of the motion direction and dot coherence (they are controllable experimental features after all), we cannot directly measure the sensitivity parameter in participants. In Bayesian models, it is possible to propose a range of potential values for the parameters (e.g., for parameter \( k \) it was used a Normal distribution of \( \mu = 0 \) and \( \sigma = 10 \)). These assumed values
Figure 2.7: Example Bayesian network. The probability a student gets a recommendation letter from their professor depends on the student’s grade, which in turn depends on intelligence and difficulty. All the variables in the model are stochastic and can be characterised using probabilities for each discrete state. The arrows indicate the dependence between variables. The tables depict the probabilities for the combination of the model variables. Recreated from [126].

are also known as the priors of the model. The modelling also includes participant’s choices (a) using a Bernoulli distribution (bernoulli_logit(s), with logit-transformed chance-of-success parameter, s). The logit function implements a steep softmax relating \( X_{pre} \) to \( a \) and is applied for computational stability. Confidence is computed by a deterministic function that includes a complex log odds function (LO) that depends on the evidence accumulated and the choice (please refer to [127] for details). The mapping between model confidence and observed confidence allowed a small degree of imprecision (\( \sigma = 0.025 \)) in subjects’ ratings.

While the models on their own can be very informative (e.g., giving an idea of the relationship between the factors and outputs from simulations) the number of potential combinations between all the variables generates a gigantic space of possible models. For that reason, it is important to constrain the space using observations from real-life experiments. The process of looking for models that can generate predictions that resemble real-life data corresponds to the fitting process. In the model above, the fitting process employs observations from human participants performing the two-stage decision, e.g., motion directions \( d \), motion coherences \( \theta_{pre} \) and \( \theta_{post} \), subjects’ choices \( a \) and confidence ratings \( r \). There are multiple methods to search for adequate model parameters. In this work I focused on Bayesian inference. Note that Bayesian inference and Bayesian modelling are not the same: we are using Bayesian inference to estimate the parameters in a Bayesian graphical model. The approach taken is the one described by Probabilistic Programming (PP). A probabilistic program is run in both the forward and backward direction. It runs forward to compute the consequences of the assumptions it contains about the world (i.e., the model space it
Figure 2.8: Bayesian graphical model for a perceptual decision experiment. Adapted from [127], with Springer Nature permission.

represents), but it also runs backwards from the data to constrain the possible explanations. Bayesian inference points to obtain the posterior distributions \( p(\lambda|x) \), from Bayes theorem equation:

\[
p(\lambda|x) = \frac{p(x|\lambda)p(\lambda)}{p(x)} \tag{14}
\]

The posterior characterises the probability distribution of the set of model parameters \( \lambda \) given the observed data \( x \). The likelihood \( p(x|\lambda) \) corresponds to the probability distribution that the observed data was obtained from the model given a specific set of parameters. The prior \( p(\lambda) \) expresses the probability distributions of the parameters, independent of any observation. This is very important since here we can include previous knowledge into the model. For example, the parameter \( k \) in the model proposed by [127] was defined to take values following a Gaussian distribution with \( \mu = 0 \) and \( \sigma = 10 \), this corresponds to the prior for that parameter. Finally, the evidence \( p(x) \), is the probability distribution of the observations independent of the model and parameter values. In realistic models, it is difficult to estimate the \( p(x) \), therefore, it is a common approach to ignore it considering that the posterior distribution is proportional to the product the likelihood and the prior \( p(\lambda|x) \propto p(x|\lambda)p(\lambda) \). Since \( p(x) \) works as a normalisation factor, it is possible to have a reasonable approximation to the posterior parameter space without considering it. To estimate that approximation I used the Markov Chain Monte Carlo (MCMC) approach [128] to obtain samples from the posterior distributions. The Monte Carlo feature of this method means that according to the Law of Large numbers, the average of the samples tends towards the expected value of the distribution the more samples we have. On the other hand, the Markov chain aspect indicates that the sampling is not fully random, but it follows a strategy where the next sample depends on the value of the present sample. This allows us to look for high-density areas on the target distribution, i.e., it is a targeted sampling.

Multiple algorithms have been proposed to sample the parameter space in MCMC. One of the most used is the Metropolis-Hastings algorithm [129], which can be widely described by the following steps: 1) start at the current
position in the parameter space; 2) propose moving to a new position; 3) accept/reject the new position based on the position’s adherence to the data and prior distributions. 4.1) if it is accepted: move to the new position and return to step 1. 4.2) if it is not accepted: do not move to the new position and return to step 1; 5) after many iterations, return all accepted positions.

After a large number of repetitions, the sampling process is expected to reach a steady state (i.e., a stable prediction of the parameters). The sequence of repetitions is usually called a chain (Figure 2.9). Given that at the beginning of the chain (i.e., the initial repetitions of the algorithm) the initialisation of the parameters may not be representative of the steady state (e.g., random initialisation of the parameters), it is common to define a burn-in time, which considers samples that are discarded since they do not follow the target probability distribution. Some other methods can also be used to obtain more informative samples, such as thinning, where sequential samples can be discarded to avoid internal correlation given the path followed by the chain, generating (almost) independent samples. To avoid initialisation problems or local zones of high density, multiple chains are run. At the end of the sampling process, it is important to check that the results are consistent, showing convergence to a unique set of parameters. Standard diagnostic tests consider Gelman-Rubin \(|\bar{R} - 1| < 0.05\) or the effective sample size \(\text{ESS} > 100\). The Gelman-Rubin diagnostic \([130]\) checks the lack of convergence by comparing the variance between multiple chains to the variance within each chain. If convergence is achieved, the between-chain and within-chain variances should be identical. The ESS, roughly speaking, captures how many independent draws contain the same amount of information as the dependent samples obtained by the MCMC algorithm \([131]\).

**Figure 2.9:** Markov Chain Monte Carlo. (A) Sampling process for the estimation of parameters \((\lambda)\). Multiple chains sampling from different starting points eventually converge, delivering samples of the posterior distribution. Burn-in samples correspond to discarded samples before the chains converge. (B) Convergence of the Metropolis–Hastings algorithm for different numbers of samples. MCMC attempts to approximate the blue distribution with the orange distribution.
2.6 REINFORCEMENT LEARNING

One of the major questions in cognitive sciences is how animals learn to select actions to obtain rewards and avoid punishment. From the psychological perspective, this question was initially investigated through classical and instrumental conditioning paradigms, which have been very informative in describing how the relationship between stimuli and rewards is generated during learning. Pavlovian conditioning is an example of prediction learning: learning the association between features of the environment (e.g., a sweet aroma coming out of the oven) and a desirable outcome (e.g., the fulfilling taste of warm bread) [132]. Instrumental conditioning characterises learning to select actions that facilitate the occurrence of reward (or reduce the possibility of aversive events). In other words, the target of instrumental conditioning is to find actions that influence the environment for the subjective benefit of the decision-maker.

The development of reinforcement learning (RL) in more recent times has provided an important normative framework to continue analysing the animal conditioning process [133, 134, 135]. A fundamental insight that informs RL is that learning should occur when outcomes are surprising [136]. One of the most influential models is the Rescorla-Wagner model [137] which characterises Pavlovian conditioning, where learning is driven by events that violate expectations. This is formalised by the following equation, where the associative strength between the conditioned stimulus (CS, e.g., a sound) and the unconditioned stimulus (US, e.g., food) is characterised by $V$:

$$V_{new}(CS_i) = V_{old}(CS_i) + \eta \left( \lambda_{US} - \sum_i V_{old}(CS_i) \right) \quad (15)$$

The error learning rule indicates that learning is driven by the difference between the expectations ($\sum_i V_{old}(CS_i)$, with $i$ indicating all the CSs involved) and what is actually experienced. $\lambda_{US}$ parameter indicates the magnitude of the US, quantifying the maximal associative strength that US can support. $\eta$ corresponds to the learning rate which can depend on the salience properties of both the US and the CS being associated. Rescorla-Wagner model can explain a variety of behavioural phenomena like blocking or overshadowing [132]; however, its results are based on CS and US and do not extend to second-order conditioning. Unlike US, that have an affective effect per se (e.g., food or electric shocks in mice experiments), second-order conditioning allows to elicit conditioned responses with CSs that were never associated with US. This type of conditioning is very common in human experiments, where the motivation is given by monetary outcomes, which generates second-order associations. Additionally, Rescorla-Wagner model assumes that learning occurs in a single conditioning trial, ignoring any temporal considerations.

As an answer to these shortcomings [138] proposed the temporal difference (TD) learning rule. In TD learning the agent estimates the value of different states or situations in terms of the future rewards (or punishments) they predict. The formalisation comes from Markov chains, a dynamic process where different states $S$ follow one another according to a probability distribution $P(S_{t+1} | S_t)$. In this case, $S_t$ indicates the current state and $S_{t+1}$ the following state. For example, for a rat moving in a maze, the states would correspond to specific locations each one accessible (or not) from the contiguous positions, e.g., the transition probability between two positions separated by a wall would be 0. Additionally, each state could present rewards also given
from a probability distribution \( P(r|S_t) \). The value \( V \) of a state is defined as any reward present in that state plus potential rewards accessible in future states that can be achieved from there. The *temporal difference prediction error* is expressed by the equation:

\[
\delta_t = P(r|S_t) + \gamma \sum_{S_{t+1}} P(S_{t+1}|S_t) V(S_{t+1}) - V(S_t)
\] (16)

Here \( \gamma \) is a temporal discount factor that reduces the effect of rewards that are distant in time. As in the previous model, this equation expresses that the discrepancy between the observed and expected values of a state will drive the learning of \( V(S_t) \):

\[
V(S_t)_{\text{new}} = V(S_t)_{\text{old}} + \eta \delta_t
\] (17)

In this way, the error signals tend to sequentially improve the estimates of the function \( V(S_t) \). One problem of this learning rule is that it assumes that the dynamics of the environment are known (\( P(S_{t+1}|S_t) \) and \( P(r|S_t) \)) which is hardly true in real-world cases. Most of the time, the agent will have to explore the situation and the environment itself will supply this information stochastically and incrementally. In other words, the agent “samples” the probabilities distributions for reward and state transitions. Incorporating this assumption, the learning rule can be expressed as:

\[
V(S_t)_{\text{new}} = V(S_t)_{\text{old}} + \eta [r_t + \gamma V(S_{t+1})_{\text{old}} - V(S_t)_{\text{old}}]
\] (18)

Where \( r_t \) is the reward observed a time \( t \), when in state \( S_t \). Therefore, this model allows learning optimal predictive values of events in the environment, even if they are unknown. However, while these various learning processes allow for generating correct predictions of the states, we must keep in mind that the environment gives rewards for actions, not predictions [132]. The steps presented above consider fixed behavioural policies, however, one of the fundamental goals in prediction learning is to help to select the best action. The problem of *credit assignment*, or how to elucidate retrospectively what actions are relevant to reach an outcome (e.g., win or loss) is a crucial and difficult problem to improve the behavioural policy. In RL this problem is solved by defining action selection depending not only on the immediate outcomes but also on future value predictions, as presented in TD learning. In this scenario, the best action to choose is the one that leads to the state with the highest value.

One of the methods employed for action selection considers explicitly learning the predictive value (i.e., future expected rewards) of taking a particular action at a specific state, in other words, to learn the value of the state-action pair, denoted as \( Q(S,a) \). This process has been denominated Q-learning [139]. In this case, the learning rule is similar to the one for TD learning, only that it also includes the action:

\[
Q(S_t,a_t)_{\text{new}} = Q(S_t,a_t)_{\text{old}} + \eta \delta_t
\] (19)

Where the prediction errors are expressed as:

\[
\delta_t = r_t + \max_a \gamma Q(S_{t+1},a) - Q(S_t,a_t)
\] (20)

Here the max operator indicates that the temporal difference is calculated with respect to the (believed) best action at the next state \( S_{t+1} \). One of the important consequences of learning Q-values is that it does not require to
learn a separate policy. Furthering the interest in this model, findings of
dopaminergic neurons in non-human animals have suggested that they
could be communicating prediction error based on state-action values [37, 140].
The connection between dopaminergic systems and RL has been extensive
[141, 142, 143, 144, 135]. These circuits spark special interest given their
importance in conditions such as Parkinson’s disease or addiction. Therefore,
RL is a model that can be very useful to understand how the brain learns
values and make decisions in novel environments.

2.7 FUNCTIONAL MAGNETIC RESONANCE IMAGING

There is a variety of methods that allow the measurement of various aspects
of brain activity. In humans, methods like electroencephalography (EEG) (or
more recently magnetoencephalography, MEG) have been used extensively,
giving access to the electromagnetic perturbations generated by neuronal
activity. These tools have been very useful to track the swift changes in brain
processes, however, they usually lack a precise spatial resolution and cover-
age. This deficiency has been saved thanks to technological breakthroughs
propelled by functional Magnetic Resonance Imaging (fMRI), which allows
reaching even the millimetric scale without relying on invasive interventions.
Since the development of its first papers in the early 1990s [145, 146, 147, 148],
this technique has fostered an explosion of research in cognitive neuroscience.
The principles of operation of MRI allow the generation of images based
on the magnetic properties of hydrogen nuclei (protons). Given the varied
composition and structure of different types of tissue, such as grey or white
matter, specific magnetic signatures can be used to generate anatomical
images of the brain.

MRI can be used both to explore the structure of the brain and to track
dynamic changes in its operation. Neurons are no different from other cells
in the body and they consume energy through oxygen and glucose. The
circulatory system supplies the brain’s metabolic demands, which overall can
take approximately 20% of all the oxygen we breathe [149]. However, that rate
of consumption is not homogeneous, and the activity of neuronal groups can
affect locally the supply of oxygen-rich blood. This variation in blood flow
when the demands increase has been characterised by the haemodynamic
response function (HRF) (Figure 2.10). The oxygen supplied to the neurons
travels bonded to haemoglobin, an iron-containing transport protein. Iron
has magnetic properties, which in this case vary depending on whether it
is part of oxygenated or deoxygenated haemoglobin. fMRI can measure the
ratio of oxygenated vs deoxygenated blood, which is referred to as the blood
oxygenation level-dependent effect, or BOLD effect [150]. When a brain area
becomes active the blood flow rises, which generates an overall increase
in the BOLD signal (Figure 2.10). The change in blood flow is slower than
neuronal activation, which means the variation in the signal takes a couple
of seconds to reach its peak after neuronal activity. Therefore, unlike EEG or
MEG, fMRI accesses brain activity indirectly, measuring metabolic changes
that appear because of neuronal activations, thanks to the conjunction of an
increase in blood flow and a change in magnetic properties.

The fMRI scanning sessions generate sequences of brain volumes, usually
sampled with a couple of seconds of separation (repetition time, TR). An
experimental session can have hundreds of volumes. The details of the
specific methods for the acquisition of the images are beyond the reach of
this work and can be found elsewhere [151]. It is relevant to note that
2.7 Functional Magnetic Resonance Imaging

Figure 2.10: Haemodynamic variations by blood flow. A) Origin of the BOLD effect in the brain. Arterial blood is similar to tissue in terms of magnetic properties. Deoxygenated blood is paramagnetic which causes inhomogeneities in the magnetic field, and a decay in the MRI signal. Regions of the brain with higher activation (blood flow) are more magnetically uniform which increases the MRI signal. B) Canonical haemodynamic response function (HRF). The peak of the signal indicates an increase in oxygenated blood flow. After some seconds, the needs of the neuronal tissue are met, returning the blood flow to homeostatic levels. Time zero in the plot corresponds to the onset of stimulus or action, which means that the peak of activation has a couple of seconds of delay with respect to the actual brain activity. Panel A from [150] (CC license).

The images used in fMRI analysis are mainly of two types: anatomical images (also known as T1-weighted images) and dynamic functional images (known as T2*-weighted images). Additionally, fieldmaps images account for variations in the magnetic field inside the MRI machine at the moment of scanning and are used to correct these distortions (inhomogeneities in the field). After the acquisition of the brain volumes, the pre-processing of images is crucial to allow a consistent and reliable extraction of brain patterns [152, 153]. In general, the stages included in this process consider:

1. Slice-timing within brain volumes: the brain 3D volumes are acquired as stacks of 2D images, one slice at a time. Even when the time taken to acquire a full volume can be in the range of seconds, the internal displacement and temporal delay can affect the capacity to identify the desired effects. Slice-timing interpolates information in each slice to match the timing of a reference slice (e.g., first or mean TR slice) for each volume.

2. Motion correction across brain volumes: head motion can be a critical confound in the analysis. The most common strategy is to realign each volume to a reference (e.g., the mean, first or last brain volume). In addition, when the subject-level analysis of the signal is performed using GLM, the motion coordinates extracted from this stage are used as regressors in the model, to control for potential motion effects.

3. Spatial transformations: given the natural anatomical differences between human brains, the volumes are transformed from the subject’s native space to a common normalised space so proper comparison and integration of the subject’s data can be performed. A brain-space
commonly used is the Montreal Neurological Institute (MNI) brain template, which is the average of the anatomical images of many subjects [154]. Complementary steps like skull stripping and segmenting non-brain areas can be useful to improve the normalisation step.

4. Spatial smoothing: to reduce the noise and boost the signal in the spatial dimension, brain data points (the volumetric unit, voxels) are averaged with their neighbours. This step is beneficial for the validity of statistical tests reducing anatomical differences. However, this can be detrimental to spatial resolution on eventual brain activity, displacing activation peaks or extinguishing smaller local but significant activations.

Different software alternatives are available for pre-processing and analysis of brain volumes (e.g., AFNI, BrainVoyager, FSL). In the current work, I used the Statistical Parametric Mapping (SPM) package ([155]; www.fil.ion.ucl.ac.uk/spm).

2.7.1 Univariate analysis

Once the pre-processing has been completed the search for relevant brain activations can be performed using different analysis methodologies. In general lines, there are two main approaches: a univariate and multivariate analysis. In this section, I will focus on the univariate type, while multivariate is developed in the following section.

The univariate analysis relates psychological or physical dimensions with the activation of single brain voxels, assuming they are independent of their surroundings [156]). The most employed method for the analysis of task-based fMRI is SPM, which is based on the general linear model (GLM; [157, 153]). The GLM is expressed as:

\[ Y = X\beta + \epsilon \quad (21) \]

To formalise the model, relevant onsets or periods in the experimental time course are convolved with the canonical HRF, which quantifies the estimated BOLD signal in response to the conditions of interest. These estimates together with confounding factors such as head motion or respiratory signal are used as independent variables (parameters) in the model. These parameters are joined together in the X matrix (in equation 21) which corresponds to the design matrix, with \( \beta \) the parameter estimates and \( \epsilon \) being the error. The dependent variable in the GLM (\( Y \) in equation 21) corresponds to each one of the voxel time series. In other words, a GLM is proposed for each voxel independently, estimating the effects of the different parameters on the voxel’s activity. This allows the estimation of a parametric map that identifies areas in the brain affected by specific model parameters (e.g., occipital area voxels are positively modulated by an image onset in a visual experiment). This analysis is performed separately for each subject, and it is usually known as the first-level analysis. To make inferences at a group level (second-level analysis), statistical tests are performed over the first-level results. In SPM, one-sample, two-sample and paired t-tests together with ANOVA tests are readily available to compare across subjects’ groups.

2.7.2 Multivariate analysis

The univariate analysis is capable of determining the areas of the brain involved in the task by measuring repeatedly thousands of locations, but by
analysing each one of them separately. This approach restricts the findings to individual brain locations where the differences between task conditions are sufficiently large to allow discrimination, which many times is not easy to achieve. However, the range of human neuroimaging can be expanded by considering full spatial patterns measured simultaneously at many locations, an approach used in the multivariate analysis [158, 159]. Some advantages of this approach have been listed [160]:

1. The weak information available in individual locations can be accumulated across many spatial locations. In this way, two areas that do not contain information on their own could be informative when considered jointly.

2. In univariate analysis, variance in the brain signal in the spatial dimension is discarded through pre-processing steps as spatial smoothing. Methods that use the pattern of activity in multiple areas simultaneously can take advantage of this spatial resolution.

3. The standard univariate approach requires multiple repetitions of events to measure the average differences between conditions and maximise statistical sensitivity. Multimodal analyses, given the higher sensitivity, could facilitate more direct estimates of participants’ perceptual or cognitive states.

In other words, the classical univariate approach allows for uncovering the areas of the brain processing specific stimulus classes, however, it does not reveal how those regions represent the stimuli [161, 162]. Various methodologies have been employed for the analysis of brain patterns, here we focus on machine learning classifiers [163, 162] and representational similarity analysis [164, 165].
2.7.3 Classifiers

Each one of the conditions in fMRI experiments is associated with specific configurations of brain activity. These multivariate samples of brain activity can be extracted directly from the unsmoothed MRI volumes or they can be the result of GLM analysis, encompassing the whole brain or selected regions of interest (ROI). Each one of these samples can then be assigned labels that associate it with the condition or category active the moment they were acquired. For example, the selected categories can be cat or dog, and the multiple presentations of images in those two conditions will generate many samples of how the brain reacts to those stimuli (Figure 2.12). The objective of the classifier is to find a way to separate the categories in the brain pattern space, estimating a decision boundary [166]. The classification algorithms that require a labelled organisation of the data are denominated supervised learning (as opposed to unsupervised learning where no tags on data are required). The estimation of decision boundary is done using various methods, for example, linear discriminant analysis (LDA) or support vector machines (SVM) [167]. These methods estimate a mathematical model that contains the weights of each voxel to linearly project to a single decision axis. The estimation of these weights is performed using only a part of the available labelled samples as a training dataset. LDA estimates projection weights that maximise within-class to between-class variance. On the other hand, SVM estimates the weights that define a maximum margin hyperplane, i.e., it projects the samples to a space where the gap between categories is maximised. In that way, using that new voxel space the decision boundary can be defined to generate the best classification.

The remaining part of the samples is used as a test dataset to check whether the estimated model (i.e., weights and decision boundary) can generalise the classification to new brain patterns. The proportion of correctly predicted labels in the test set is called classification accuracy. The process of partitioning the data in training and test sets in different configurations, many times, is denominated cross-validation. All these extra considerations are extremely relevant to avoid overfitting, i.e., the model fits its training data exactly, which means it will most likely perform poorly on unseen data.

2.7.4 Representational similarity analysis

The decoding analysis points to revealing multivariate explicit information in neural representations, relying on predefined categories. Representational models go one step further in describing the representational geometry in specific areas of the brain, i.e., defining features or conditions represented in the area, how strongly each one of them is represented (signal-to-noise ratio) and characterising the similarity between the different conditions [168]. Representational similarity analysis (RSA) describes the relationship between the experimental conditions estimating how different are their brain patterns from each other, in other words, how distant they are in representational space. The RSA approach can be very useful when a one-to-one mapping cannot be implemented, e.g., when comparing the brain activity between humans and monkeys, the distinct nature of the signal and anatomy hinders a direct comparison [161]. However, if the brain patterns are characterised from the representational perspective, it is possible to set a common ground for comparison. In other words, RSA does not characterise the patterns in
Figure 2.12: Multivariate pattern classification. Brain patterns are acquired from the presentation of different images. In this toy example, 8 voxels are shown to represent the brain patterns. Two categories of items are relevant for the classification: cat (A) and dog (B). Each stimulus will have a specific voxel response, which can be represented in a voxel space (for simplicity, only two voxels are plotted in the 2-D coordinate system). Multiple repetitions of the stimulus in each category generate various samples of brain patterns that populate the voxel space. These samples can be used for supervised binary classification. C) Part of the labelled samples (cat in red and dog in green) is used as training data, which allows fitting the parameters necessary to define the decision rule. In this example, a linear decision boundary (for example, obtained from an SVM) is used to separate the two categories. To assess the precision of the estimated rule, brain patterns that have not been used previously in training are categorised. Predicted and true labels are checked to estimate the accuracy of the classification. The proportion of correct classification can be tested against chance performance (in this case 0.5 since it is binary classification). Usually, the inclusion of more voxels tends to improve the accuracy of the classification. The weights of the (linear) classifier can be tracked to their voxels, indicating which specific brain voxels were more relevant (i.e., containing more information) to decide the category. Adapted from [162], with permission Elsevier.

a brain area per se, but the signatures of similarity between the different representations found in the brain area (Figure 2.13).

The first step in RSA is estimating the activity patterns. For this is necessary to define a vector of activity amplitudes across response channels (i.e., in the case of fMRI corresponds to voxels, but it could be neurons or sites in cell recording) within a region of interest (ROI). Note that patterns could also be extracted from weights or parameters in computational models, or even features of behaviour such as reaction time or errors [164]. For the analysis, the activity pattern typically considered is a spatial pattern,
Although it may also be a spatiotemporal pattern. The pattern associated with each experimental condition is considered a representation. In the case of fMRI, the activity estimate for each voxel and condition is usually obtained from univariate analysis with the design matrix used to model each voxel’s response based on the event sequence (e.g., each visual stimuli presented in the experiment) and a linear model of the hemodynamic response.

The second stage considers measuring the activity pattern dissimilarity. For this purpose, a representational dissimilarity matrix (RDM) for the ROI is built. The matrix is constructed by pairing the representations of the experimental conditions and calculating the dissimilarity between them. A measure of activity-pattern dissimilarity that normalises for both the mean level of activity and the variability of activity is the correlation distance, i.e., 1 minus the linear correlation between patterns. The diagonal entries in the RDM are 0 since they reflect the comparisons between identical conditions. The off-diagonal values indicate the dissimilarity between the activity patterns associated with two different stimuli/conditions. Therefore, dissimilarities can be interpreted as distances in the multidimensional response space. The size of the RDM matrix depends on the number of conditions represented.

Standard neuroscientific studies are mostly focused on the relationship between stimulus properties and brain activity levels in single cells or brain regions, also called the first-order isomorphism between stimuli and their representations. The second-order isomorphism is related to the comparison among the similarity structures, e.g., comparing the representations of a series of images by the inferior temporal cortex vs the early visual cortex in humans (Figure 2.13b, [161]). For this type of analysis, the comparison across RDMs can be performed. As mentioned above, brain RDMs can be compared with representational structures obtained from computational models or other sources. For example, RDMs obtained from the visual cortex could be compared with simulated retinotopic maps based on banks of Gabor filters for various spatial frequencies and orientations, or conceptual models that contain expected similarity between the objects presented in the images. An alternative distance measure, such as the Euclidean distance, can be used for comparing dissimilarity matrices.

Further steps in the analysis pipeline consider statistical tests using randomisation measures to verify that the similarities are significant. Visualisations of the representational structures can be presented as coloured matrices, dendrograms, or multidimensional scaling figures [165]. Overall, the RSA approach can be very useful to get insights into the way the brain represents information across various regions in a wide variety of neuroimaging studies.

2.8 Conclusion

In this chapter, I have presented a broad recap of the main experimental and analysis techniques used in this thesis. Initially, I presented techniques mainly employed for the analysis of human behaviour, including the decision processes and visual attention deployment using eye tracking. Drift diffusion models, signal detection theory and reinforcement learning approaches will have a central role in the characterization of the decision process. The brain imaging experiments in this work employ the fMRI technique, therefore, I supported further details on the fundamentals of the method and the steps required for the pre-processing and analysis of this data. The experimental section of this thesis has been separated into three parts: 1) information
sampling, 2) brain representations, and 3) confidence computation, regarding the human decision process.

**Figure 2.13:** Representational similarity analysis. (A) Computation of the RDM for a single ROI from the brain activity patterns generated for each sensory stimulus. For each pair of activity patterns, a dissimilarity is computed which corresponds to a single element in the RDM. RDMs serve as a signature of representation that can be compared between multiple ROIs in the brain, models, individuals and even species. The size of the matrix is defined by the number of conditions (N images in this example). The matrix can be represented as a graph, considering the conditions as nodes and the distance between them as the dissimilarity value, i.e. the closer the nodes the more similar the representations in that ROI. (B) RSA allows us to see the representation geometry in various areas of the brain and check its variation. RDMs were obtained in a visual human fMRI experiment with multiple images presented [161]. The images of various objects, plants, animals, and human faces were displayed, and activity patterns were extracted from the inferior temporal (IT) and early visual cortex. IT can be seen to represent hierarchically animate and inanimate objects, with further distinctions between faces and body parts in human and non-human images. The structure of representation in the early visual cortex can be seen as distinct and more diffuse across the presented images. Panel A from [165] (CC license) and panel B from [161], with permission from Elsevier.
Part I

INFORMATION SAMPLING AND GOALS IN HUMAN DECISION
Whenever we decide, e.g., picking the dessert we will get from the supermarket, it is likely we will explore the options, looking at them to evaluate and make our final resolution. Decisions in real life are a dynamic process where the sampling of evidence informs the future choice. Indeed, it has been reported that participants tend to allocate attention towards the alternative they will eventually choose \[33\]. This observation has been commonly interpreted as attention being allocated to items with higher hedonic reward in value-based choices or targeting the presence of perceptual evidence. However, these studies assume choices where the participants complete a single task, conflating value/perceptual evidence and the goal of the experiment. In the study presented in this part, I show using behavioural testing, eye-tracking and computational modelling that attention modulates the integration of goal-relevant evidence. Even more, the accumulation of evidence that leads to choice also reflects the influence of goals. These findings support a more general and flexible deliberation process at the service of the goal, and not simply relying on fixed features of the stimuli.
3

VISUAL ATTENTION MODULATES THE INTEGRATION OF GOAL-RELEVANT EVIDENCE AND NOT VALUE

3.1 SUMMARY

When choosing between options, such as food items presented in plain view, people tend to choose the option they spend longer looking at. The prevailing interpretation is that visual attention increases value. However, in previous studies, ‘value’ was coupled to a behavioural goal, since subjects had to choose the item they preferred. This makes it impossible to discern if visual attention influences value, or, instead, if attention modulates the information most relevant to the goal of the decision-maker. Here, we present the results of two independent studies—a perceptual and a value-based task—that allow us to decouple value from goal-relevant information using specific task framing. Combining psychophysics with computational modelling, we show that, contrary to the current interpretation, attention does not boost value, but instead it modulates goal-relevant information. This work provides a novel and more general mechanism by which attention interacts with choice.

3.2 INTRODUCTION

Attention is thought to play a central role, in prioritising and enhancing which information is accessed during the decision-making process. How attention interacts with value-based choice has been investigated in psychology and neuroscience [33, 34, 169, 170, 171, 57, 121, 172, 173, 35] and this question is at the core of the theory of rational inattention in economics [174, 175, 176, 177].

Robust empirical evidence has shown that people tend to look for longer at the options with higher values [178, 173, 171] and that they tend to choose the option they pay more visual attention to [33, 34, 57, 169, 35]. The most common interpretation is that attention is allocated to items based on their value and that looking or attending to an option boosts its value, either by amplifying it [33, 34, 179] or by shifting it upwards by a constant amount [169]. This intuition has been elegantly formalised using models of sequential sampling. Particularly, the attentional drift-diffusion model (aDDM), which considers that visual attention boosts the drift rate of the stochastic accumulation processes [33]. More recently, this same model has been also used to study the role of attention in the accumulation of perceptual information [121]. These lines of investigation have been extremely fruitful, as they have provided an elegant algorithmic description of the interplay between attention and choice.

As a consequence of this development, a predominant assumption in the field of neuroeconomics has become that attention modulates the values of the alternatives [179]. However, this view overlooks the fact that in the
majority of these studies, value is coupled to the agents’ behavioural goal, that is, participants had to choose the item they found more rewarding. Some recent studies have called into question this assumption and have hinted towards a flexible role of attention in sampling goal-relevant options [180, 172]. Even further, recent studies have shown that the ‘value networks’ in the brain could be tracking not purely reward value, but goal-congruent information [98, 181]. Considering all this, our study aimed to understand in more detail the role of goals on visual attention during both value-based and perceptual decisions: we aim to test the hypothesis that attention acts flexibly upon the accumulation of goal-relevant information and to examine the effects on the mechanism of preference formation.

Our experimental design decoupled reward value from choice using a simple task-framing manipulation. In the eye-tracking part of our value-based experiment, participants were asked to choose between different pairs of snacks. We used two frame manipulations: like and dislike. In the like frame, they had to indicate which snack they would like to consume at the end of the experiment; this was consistent with the standard tasks used in value-based decision studies. Whereas in the dislike frame, subjects had to indicate the snack that they would prefer not to eat. In the latter frame, the value was distinct from the behavioural goal of which item to select. In fact, in the dislike frame participants needed to consider the ‘anti-value’ of the item to choose the one to reject.

To anticipate our results, in the like frame condition we replicated the typical gaze-boosting effect: participants looked for longer at the item they were about to choose – the item they deemed most valuable. In the dislike frame, however, participants looked for longer at the item that they then chose to eliminate, that is, the least valuable item. This means that agents paid more attention to the option they selected in the task, not to the option to which they deemed more valuable or wanted to consume. This suggests that attention did not boost value but rather it was used to gather task-relevant information.

To understand the mechanism via which attention interacted with value in both frames, we used a dynamic accumulation model, which allowed us to account for the preference formation process and its dependency on task variables (values of the options). To test the generality of our findings, we also conducted a new perceptual decision-making experiment with a new set of participants. In this perceptual task, participants were asked to choose between two circles filled with dots. In some blocks, they had to indicate the circle with more dots – most frame; in others, the circle with fewer dots – fewest frame. In this second study, we replicated all the effects of the first, value-based one, corroborating the hypothesis of a domain-general role for attention in modulating goal-relevant information driving choice.

This work questions the dominant view in neuroeconomics about the relationship between attention and value, showing that attention does not boost value per se but instead modulates goal-relevant information.
3.3 Methods

3.3.1 Procedure

3.3.1.1 Value experiment

At the beginning of this experiment, participants were asked to report on a scale from £0 to £3 the maximum they would be willing to pay for each of the 60 snack food items. They were informed that this bid would allow them to purchase a snack at the end of the experiment, using the BDM [182], which gave them incentives to report their true valuation. Participants were asked to fast for 4 hr before the experiment, expecting they would be hungry and willing to spend money to buy a snack.

After the bid process, participants completed the choice task: in each trial, they were asked to choose between two snack items, displayed on-screen in equidistant boxes to the left and right of the centre of the screen (Figure 3.1A). After each binary choice, participants also rated their subjective level of confidence in their choice (see chapter 7 for detailed results on this item). Pairs were selected using the value ratings given in the bidding task: using a median split, each item was categorised as high- or low-value for the agent; these were then combined to produce 15 high-value, 15 low-value, and 30 mixed pairs, for a total of 60 pairs tailored to the participant’s preferences. Each pair was presented twice, inverting the position to have a counterbalanced item presentation.

The key aspect of our experimental setting is that all participants executed the choice process under two framing conditions: (1) a like frame, in which participants were asked to select the item that they liked the most, that is, the snack that they would prefer to eat at the end of the experiment and (2) a dislike frame in which participants were asked to select the item that they liked the least, knowing that this is tantamount to choosing the other item for consumption at the end of the experiment. See Figure 3.1A for a diagram of the task.

After four practice trials, participants performed a total of 6 blocks of 40 trials (240 trials in total). Like and dislike frames were presented in alternate blocks and the order was counterbalanced across participants (120 trials per frame). An icon in the top-left corner of the screen (‘thumbs up’ for like and ‘stop sign’ for dislike) reminded participants of the choice they were required to make; this was also announced by the investigator at the beginning of every block. The last pair in a block would not be first in the subsequent block.

Participants’ eye movements were recorded throughout the choice task and the presentation of food items was gaze-contingent: participants could only see one item at a time depending on which box they looked at; following [57], this was done to reduce the risk that participant, while gazing one item, would still look at the other item in their visual periphery.

Once all tasks were completed, one trial was randomly selected from the choice task. The BDM bid value of the preferred item (the chosen one in the like frame and the unchosen one in the dislike frame) was compared with a randomly generated number between £0 and £3. If the bid was higher than the BDM generated value, an amount equivalent to the BDM value was subtracted from their £20 payment and the participant received the food item. If the bid was lower than the generated value, participants were paid £20 for their time and did not receive any snacks. In either case, participants were required to stay in the testing room for an extra hour and were unable
to eat any food during this time other than the food bought in the auction. Participants were made aware of the whole procedure before the experiment began.

3.3.1.2 Perceptual experiment

The Perceptual Experiment had a design similar to the one implemented in the Value Experiment, except that alternatives were visual stimuli instead of food items. In this task, participants had to choose between two circles filled with dots (for a schematic diagram see Figure 3.1), again in two frames. In the most frame, they had to pick the one with more dots; and the one with fewer dots in the fewest frame. The total number of dots presented in the circles could have three numerosity levels (=50, 80 and 110 dots). For each pair in those three levels, the dot difference between the circles varied in 10 percentage levels (ranging from 2% to 20% with 2% steps). To increase the difficulty of the task, in addition to the target dots (blue-green coloured), distractor dots (orange coloured) were also shown. The number of distractor dots was 80% of that of target dots (40, 64, and 88 for the three numerosity levels, respectively). Pairs were presented twice and counterbalanced for item presentation. After 40 practice trials (20 initial trials with feedback, last 20 without), participants completed 3 blocks of 40 trials in the most frame and the same number in the fewest frame; they faced blocks with alternating frames, with a presentation order counterbalanced across participants. On the top left side of the screen, a message indicating Most or Fewest reminded participants of the current frame. Participants reported their confidence level in making the correct choice at the end of each trial (see chapter 7 for the confidence analysis). As in the previous experiment, the presentation of each circle was gaze contingent. Eye tracking information was recorded for each trial. Participants received £7.5 for 1 hr in this study.

Both tasks were programmed using Experiment Builder version 2.1.140 (SR Research).

3.3.2 Exclusion criteria

3.3.2.1 Value experiment

We excluded individuals that met any of the following criteria:

1. Participants used less than 25% of the BDM value scale.
2. Participants gave exactly the same BDM value for more than 50% of the items.
3. Participants used less than 25% of the choice confidence scales.
4. Participants gave exactly the same confidence rating for more than 50% of their choices.
5. Participants did not comply with the requirements of the experiment (i.e., participants that consistently choose the preferred item in dislike frame or their average blink time is over 15% of the duration of the trials).

3.3.2.2 Perceptual experiment

Since for the Perceptual Experiment the assessment of the value scale is irrelevant, we excluded participants according to criteria 3, 4, and 5.
3.3.3 Participants

3.3.3.1 Value experiment

Forty volunteers gave their informed consent to take part in this research. Of these, 31 passed the exclusion criteria and were included in the analysis (16 females, 17 males, aged 20–54, mean age of 28.8). One participant was excluded for using less than 25% of the bidding scale (criteria 1). A second participant was excluded according to criteria 2 as they frequently gave the same bid value. A further four participants were excluded under criteria 4. Three participants were excluded due to criteria 5. In the latter case, one participant’s eye-tracking data showed the highest number of blink events and made choices without fixating on any of the items; the other two did not comply with the frame manipulation. To ensure familiarity with the snack items, all the participants in the study had lived in the UK for 1 year or more (average 17 years).

3.3.3.2 Perceptual experiment

Forty volunteers were recruited for the second experiment. Thirty-two participants (22 females, 10 males, aged 19–50, mean age of 26.03) were included in the behavioural and regression analyses. Three participants were excluded due to the repetition of the confidence rating (criteria 4). Five participants were removed for criteria 5: four of them had performance close to chance level or did not follow the frame modification, and one participant presented difficulties with eye-tracking. Due to instability in parameter estimation (problem of MCMC convergence), four additional participants were removed from the GLAM modelling analysis.

All participants signed a consent form, and both studies were done following the approval given by the University College London, Division of Psychology and Language Sciences ethics committee.

3.3.4 Eye-tracking

3.3.4.1 Value and perceptual experiments

An Eyelink 1000 eye-tracker (SR Research) was used to collect the visual data. Left eye movements were sampled at 500 Hz. Participants rested their chin over a head-mount in front of the screen. The display resolution was 1024 × 768 pixels. To standardise the environmental setting and the level of detectability, the lighting was monitored in the room using a dimmer lamp and light intensity was maintained at 4 ± 0.5 lx at the position of the head-mount when the screen was black.

Eye-tracking data were analysed initially using Data Viewer (SR Research), from which reports were extracted containing details of eye movements. We defined two interest areas (IA) for left and right alternatives: two squares of 350 × 350 pixels in the Value Experiment and two circles of 170 pixels of radius for the Perceptual Experiment. The data extracted from the eye-tracker were taken between the appearance of the elements on the screen (snack items or circle with dots in experiments 1 and 2, respectively) and the choice button press (confidence report period was not considered for eye data analysis).

The time participants spent fixating on each IA was defined as the dwelling time (DT). From it, we derived a difference in dwelling time (ΔDT).
for each trial by subtracting DT of the right IA minus the DT of the left IA. Starting and ending IA of each saccade were recorded. This information was used to determine the number of times participants alternated their gaze between IAs, that is, ‘gaze shifts’. The total number of gaze shifts between IAs was extracted for each trial, producing the gaze shift frequency (GSF) variable.

3.3.5 Data analysis: behavioural data

during like/dislike and most/fewest frames were compared using statistical tests available in SciPy. Scikit-learn toolbox in Python was used to perform logistic regressions on choice data. Fixation time series analysis was performed following [180] methodology. We segmented the time series of all the trials in samples of 10 ms. We fixed all the trial time series to the beginning of the trial when the participant could start exploring the gaze-contingent alternatives. We considered an analysis window of 2000 ms after the presentation of stimuli for all the trials. Please note that not all the trials had the same duration and no temporal normalisation was performed in this analysis. For each time sample, we obtained the gaze position and the difference in evidence (i.e. ∆Value or ∆Dots) for all trials across participants and then Pearson correlation was calculated. Permutation testing was used to assess the difference between the time series in like/dislike and most/fewest frames. Instantaneous fixations (across trials and frames) were shuffled 200 times to create a null distribution of the difference in correlation coefficients between frames. False discovery rate (FDR) was used to correct for multiple tests the p-values obtained from the permutation test (α ≤ 0.01). All the hierarchical analyses were performed using lme4 package [183] for R integrated in a Jupyter notebook using the rpy2 package (https://rpy2.readthedocs.io/en/latest/).

For choice models, we predicted the log odds ratio of selecting the item appearing at the right.

3.3.6 Data analysis: attentional model - GLAM

To get further insight into potential variations in the evidence accumulation process due to the change in frames we used the Gaze-weighted Linear Accumulator Model (GLAM) developed by [35]. GLAM is part of the family of linear stochastic race models in which different alternatives (i.e. left or right) accumulate evidence (E_i) until a decision threshold is reached by one of them, determining the chosen alternative. The accumulator for an individual option was described by the following expression:

\[ E_i(t) = E_i(t-1) + v R_i + \epsilon_t \quad \text{with} \quad \epsilon_t \sim N(0, \sigma) \quad \text{and} \quad E_i(t=0) = 0(1) \]

With a drift term (v) controlling the speed of relative evidence (R_i) integration and i.i.d. noise terms with a normal distribution (zero-centred and standard deviation σ). R_i was a term that expressed the amount of evidence that was accumulated for item i at each time point t. This term was calculated as follows. We denoted by g_i, the relative gaze term, calculated as the proportion of time that participants observed item i:

\[ g_i = \frac{DT_i}{DT_1 + DT_2} \]
with DT as the dwelling time for item i during an individual trial. Let \( r_i \) denoted the value for item i reported during the initial stage of the experiment. We defined the average absolute evidence for each item (\( A_i \)) during a trial:

\[
A_i = g_i r_i + (1 - g_i) \gamma r_i \]

This formulation considered a multiplicative effect of the attentional component over the item value, capturing different rates of integration when the participant is observing item i or not (unbiased and biased states, respectively). The parameter \( \gamma \) was the gaze bias parameter: it controlled the weight that the attentional component had in determining absolute evidence. [35] interpret \( \gamma \) as follows: when \( \gamma = 1 \), bias and unbiased states have no difference (i.e. the same \( r \) is added to the average absolute evidence regardless the item is attended or not); when \( \gamma < 1 \), the absolute evidence is discounted for the biased condition; when \( \gamma < 0 \), there is a leak of evidence when the item is not fixated. Following [35], in our analysis, we allowed \( \gamma \) to take negative values, but our results did not change if \( \gamma \) is restricted to \([0, 1]\).

Finally, the relative evidence of item i, \( R_i^\ast \), was given by:

\[
R_i^\ast = A_i - \max_j(A_j) = A_i - A_j \rightarrow R_{\text{right}}^\ast = -R_{\text{left}}^\ast
\]

Since our experiment considered a binary choice, \( R_i^\ast \) from the original formulation of the model [35], proposed for more than two alternatives, was reduced to subtract the average absolute evidence of the other item. Therefore, for the binary case, the \( R_i^\ast \) for one item was the additive inverse of the other, for example, if the left item had the lower value, we had \( R_{\text{left}}^\ast < 0 \) and \( R_{\text{right}}^\ast > 0 \). Additionally, in their proposal for GLAM, [35] noted that \( R_i^\ast \) range depends on the values that the participant reported, for example, evidence accumulation may appear smaller if the participant valued all the items similarly, since \( R_i^\ast \) may be lower in magnitude. This may not represent the actual evidence accumulation process since participants could be sensitive to marginal differences in relative evidence. To account for both issues, a logistic transformation was applied over \( R_i^\ast \) using a scaling parameter \( \tau \):

\[
R_i = \frac{1}{1 + e^{-\tau R_i}}
\]

In this case, \( R_i \) would always be positive and the magnitude of the difference between \( R_{\text{left}} \) and \( R_{\text{right}} \) was controlled by \( \tau \). For example, higher \( \tau \) would imply a bigger difference in relative evidence (and hence accumulation rate) between the left and right items. In the case that \( \tau = 0 \) the participant will not present any sensitivity to differences in relative evidence.

Given that \( R_i \) represents an average of the relative evidence across the entire trial, the drift rate in \( E_i \) was assumed to be constant, which enabled the use of an analytical solution for the passage of time density. Unlike aDDM [33], GLAM does not deal with the dynamics of the attentional allocation process in choice. Details of these expressions are available at [35].

In summary, we had four free parameters in the GLAM: \( \nu \) (drift term), \( \gamma \) (gaze bias), \( \tau \) (evidence scaling), and \( \sigma \) (normally distributed noise standard deviation).

The model fit for GLAM was implemented at a participant level in a Bayesian framework using PyMC3 [184]. Uniform priors were used for all the parameters.
\[ v \sim \text{Uniform}(1^{-10}, 0.01) \]
\[ \gamma \sim \text{Uniform}(-1, 1) \]
\[ \sigma \sim \text{Uniform}(1^{-10}, 5) \]
\[ \tau \sim \text{Uniform}(0, 5) \]

3.3.6.1 Value experiment

We fitted the model for each participant and for like and dislike frames, separately. To model participants’ behaviour in the like frame, we used the RTs and choices as input for GLAM, alongside BDM bid values and relative gaze for left and right alternatives in each trial. The original GLAM formulation (as presented above) assumed that evidence was accumulated in line with the preference value of a particular item (i.e., 'how much I like this item'). When information about visual attention was included in the model, the multiplicative model in GLAM assumed that attention boosts the evidence accumulation already defined by value. We proposed that evidence accumulation is a flexible process in which attention is attracted to items based on the match between their value and task-goal (accept or reject) and not based on value alone, as most of the previous studies have assumed. Since in the dislike frame the item with the lower value becomes relevant to fulfil the task, we considered the opposite value of the items \( r_{i, \text{dislike}} = 3 - r_{i, \text{like}} \), e.g., item with value 3, the maximum value, becomes value 0) as input for GLAM fit. For both conditions, model fit was performed only on even-numbered trials using Markov-Chain-Monte-Carlo sampling, using implementation for No-U-Turn-Sampler (NUTS), four chains were sampled, 1000 tuning samples were used, 2000 posterior samples to estimate the model parameters. The convergence was diagnosed using the Gelman-Rubin statistic (\(|\bar{R} - 1| < 0.05\)) and by corroborating that the effective sample size (ESS) was high (ESS > 100) for the four parameters \( \nu, \gamma, \sigma, \) and \( \tau \). Considering all the individual models, we found divergences in less than 3% of the estimated parameters. Model comparison was performed using Watanabe-Akaike Information Criterion (WAIC) scores available in PyMC3, calculated for each individual participant fit.

Pointing to check if the model replicates the behavioural effects observed in the data [185], we simulated choice and response time (RT) using participant’s odd trials, each one repeated 50 times. For each trial, value and relative gaze for left and right items were used together with the individual estimated parameters. Random choice and RT (within a range of the minimum and maximum RT observed for each participant) were set for 5% of the simulations, replicating the contaminating process included in the model as described by [35].

3.3.6.2 Perceptual experiment

In the Perceptual Experiment, we repeated the same GLAM analysis done in the Value Experiment. Due to instabilities in the parameters’ fit for some participants, we excluded four extra participants. Twenty-eight participants were included in this analysis. We removed outlier trials, that is, trials with RT higher than 3 standard deviations (within participant) or higher than 20
3.4 Results

In our first experiment, hungry participants (n = 31) made binary choices between snacks in one of two task-frames, like and dislike. In the like frame, participants had to report the item they would prefer to eat; in the dislike frame, they chose the item they wanted to avoid eating (Figure 3.1A). After each choice, participants reported their confidence in having made a good choice [66, 57]. At the beginning of the experiment, participants reported the subjective value of individual items using a standard incentive-compatible Becker-DeGroot-Marschak mechanism (BDM; see Materials and methods).

Our second experiment was done to test whether the results observed in value-based decisions could be generalised to perceptual decisions. A different group of participants (n = 32) made binary choices between two circles containing a variable number of dots (Figure 3.1D). In the most frame, participants reported the circle containing the higher number of dots; in the fewest frame, the one with the lower. As in the Value Experiment, at the end of each trial participants reported their confidence in their choice.

3.4.1 The effect of attention on choice

3.4.1.1 Value experiment

Our results confirmed that participants understood the task and chose higher value items in the like frame and lower value items in the dislike frame (Figure 3.1B,C). This effect was modulated by confidence (Figure 3.1B) similarly to previous studies [66, 57, 58]. For a direct comparison of the differences between the goal manipulations in the two tasks (Value and Perceptual) see [186].

We then tested how attention interacts with choice by examining the eye-tracking variables. Our frame manipulation, which orthogonalised choice and valuation, allowed us to distinguish between two competing hypotheses. The first hypothesis, currently dominant in the field, is that visual attention is always attracted to high values items and that it facilitates their choice. The alternative hypothesis is that attention is attracted to items whose value matches the goal of the task. These two hypotheses make starkly different experimental predictions in our task. According to the first one, the gaze will mostly be allocated to the more valuable item independently of the frame.
Figure 3.1: Task and behavioural results. Value-based decision task (A). Participants choose between two food items presented in an eye-contingent way. Before the choice stage, participants reported the amount of money they were willing to bid to eat that snack. In the like frame (top) participants select the item they want to consume at the end of the experiment. In the dislike frame (bottom) participants choose the opposite, the item they would prefer to avoid. After each choice participants reported their level of confidence. (B) After a median split for choice confidence, a logistic regression was calculated for the probability of choosing the right-hand item depending on the difference in value (Value\textsubscript{Right} – Value\textsubscript{Left}) for like (top) and dislike (bottom) framing conditions. The logistic curve calculated from the high confidence trials is steeper, indicating an increase in accuracy. (C) The slope of logistic regressions predicting choice for each participant, depending on the frame. The shift in the sign of the slope indicates that participants are correctly modifying their choices depending on the frame. Perceptual decision task (D) Participants have to choose between two circles containing dots, also presented eye-contingently. In the most frame (top), participants select the circle with more white dots. In the fewest frame (bottom), they choose the circle with the lower number of white dots. Distractor dots (orange) are included in both frames to increase the difficulty of the task. Confidence is reported at the end of each choice. We obtained a similar pattern of results to the one observed in the Value Experiments in terms of probability of choice (E) and the flip in the slope of the choice logistic model between the most and fewest frames (F). Reprinted from [186], eLife Sciences Publications.

The second hypothesis instead predicts that in the like frame participants will look more at the more valuable item, while this pattern would reverse in the dislike frame, with attention mostly allocated to the least valuable item. In other words, according to this second hypothesis, visual attention should predict choice (and the match between value and goal) and not value, independently of the frame manipulation.

Our data strongly supported the second hypothesis because we found participants preferentially gaze (Figure 3.2A) at the higher value option during like ($t(30) = 7.56, p<0.001$) and the lower value option during dislike.
frame ($t(30) = -4.99, p<0.001$). From a hierarchical logistic regression analysis predicting choice (Figure 3.2B), the difference between the time participants spent observing the right over left item ($\Delta DT$) was a positive predictor of choice both in *like* ($z = 6.448, p<0.001$) and *dislike* ($z = 6.750, p<0.001$) frames. This means that participants looked for longer at the item that better fits the frame and not at the item with the highest value. Notably, the magnitude of this effect was slightly lower in the *dislike* case ($t(30) = 2.31, p<0.05$). In Figure 3.2B are also plotted the predictors of the other variables on choice from the best fitting model.

**Figure 3.2:** Attention and choice in Value and Perceptual Experiments. (A) Gaze allocation time depends on the frame: while visual fixations in the *like* frame go preferentially to the item with a higher value (top), during the *dislike* frame participants look for longer at the item with a lower value (bottom). Dots in the bar plot indicate participants’ average gaze time across trials for high and low value items. Time is expressed as the percentage of trial time spent looking at the item. Similar results were found for gaze distribution in the Perceptual Experiment (C) participants gaze at the circle with a higher number of dots in the most frame and the circle with a lower number of dots in the fewest frame. Hierarchical logistic modelling of choice (probability of choosing the right item) in Value (B) and Perceptual (D) Experiments, shows that participants looked for longer ($\Delta DT$) at the item they chose in both frames. All predictors are z-scored at the participant level. In both regression plots, bars depict the fixed-effects and dots the mixed-effects of the regression. Error bars show the 95\% confidence interval for the fixed effect. In Value Experiment: $\Delta Value$: difference in value between the two items ($Value_{Right} - Value_{Left}$); $RT$: reaction time; $\Sigma Value$: summed value of both items; $\Delta DT$: difference in dwell time ($DT_{Right} - DT_{Left}$); Conf: confidence. In Perceptual Experiment: $\Delta Dots$: difference in dots between the two circles ($Dots_{Right} - Dots_{Left}$); $\Sigma Dots$: summed number of dots between both circles. $***p<0.001$, $**p<0.01$, $*p<0.05$. Reprinted from [186], eLife Sciences Publications.
3.4.1.2 Perceptual experiment

We then analysed the effect of attention on choice in the perceptual case to test the generality of our findings. As in the Value Experiment, our data confirmed that participants did not have issues in choosing the circle with more dots in the most frame and the one with the least amount of dots in the fewest frame (Figure 3.1D,F). Furthermore, as in the Value Experiment and many other previous findings [66, 57], confidence modulated the accuracy of their decisions (Figure 3.1E). Critically for our main hypothesis, we found that participants’ gaze was preferentially allocated to the relevant option in each frame (Figure 3.2C): they spent more time observing the circle with more dots during most frame (t(31)=13.85, p<0.001) and the one with less dots during fewest frame (t(31)=-10.88, p<0.001). ∆DT was a positive predictor of choice (Figure 3.2D) in most (z = 10.249, p<0.001), and fewest (z = 10.449, p<0.001) frames. Contrary to the results in the Value Experiment in which the effect of ∆DT on choice was slightly more marked in the like condition (Figure 3.2B), in the Perceptual Study, the effect of ∆DT was the opposite: ∆DT had a higher effect in the fewest frame (∆DT_{Most-Few} : t(31)=-2.17, p<0.05)(Figure 3.2D). However, and most importantly, in both studies ∆DT was a robust positive predictor of choice in both frame manipulations. To summarise, these results show that in the context of a simple perceptual task, visual attention also has a specific effect in modulating information processing in a goal-directed manner: subjects spend more time fixating on the option they will select, not necessarily the option with the highest number of dots.

In both Value and Perceptual Experiments, the most parsimonious models were reported in the manuscript and in Figure 3.2B and D. For a full model comparison and more details on the choice models please check [186].

3.4.2 Fixations effects in choice

An important prediction of attentional accumulation models is that the chosen item is generally fixated upon last (unless that item is much worse than the other alternative), with the magnitude of this effect related to the difference in value between the alternatives. This feature of the decision has been consistently replicated in various previous studies [33, 34, 187]. Therefore, we tested how the last fixation was modulated by the frame manipulation.

3.4.2.1 Value experiment

In the Value Experiment in both frames, we replicated the last fixation effect and its modulation by the value difference between the last fixated option and the other one (Figure 3.3A). In the like frame, the probability of choosing the last item fixated upon increases when the value of the last item is higher, as is shown by the positive sign of the slope of the logistic curve (mean $\beta_{\text{Like}} = 0.922$). Crucially, during the dislike frame the opposite effect was found: the probability of choosing the last seen option increases when the value of the non-chosen item is higher, seen from the negative slope of the curve (mean $\beta_{\text{Dislike}} = -0.951$; $\Delta \beta_{\text{Like-Dislike}}$: t(30)=7.963, p<0.001).

3.4.2.2 Perceptual experiment

We observed the same pattern of results that in the Value Experiment (Figure 3.3B). In the most frame, it was more probable that the last fixation was on
Figure 3.3: Fixation effects on the chosen item. Last fixation effects: (A) in the Value Experiment, a logistic regression was calculated for the probability the last fixation is on the chosen item (P(LastFix = Chosen)) depending on the difference in value of the item last fixated upon and the alternative item. As reported in previous studies, in *like* frame, we find it is more probable that the item last fixated upon will be chosen when the value of that item is relatively higher. In line with the hypothesis that goal-relevant evidence, and not value, is being integrated to make the decision, during the *dislike* frame the effect shows the opposite pattern: P(LastFix = Chosen) is higher when the value of the item last fixated on is lower, that is, the item fixated on is more relevant given the frame. (B) A similar analysis in the Perceptual Experiment mirrors the results in the Value Experiment with a flip in the effect between *most* and *fewest* frames. Lines represent the model predictions and dots are the data binned across all participants. ∆Value and ∆Dots measures are z-scored at the participant level. Gaze preference in time: (C) Pearson correlation between gaze position and difference in value (∆Value) was calculated for each time point during the first 2 s of the trials. In the Value Experiment, after an initial phase of random exploration, fixations are positively correlated with the high value item in *like* frame, while this effect is the opposite for *dislike* frame, that is, fixations are directed to the low value item. (D) In the Perceptual Experiment, a similar pattern of goal-relevant fixations emerges. Lines in both figures correspond to the time point correlation considering all trials and participants. The shaded area corresponds to the standard error. The black line indicates time points with statistically significant differences between frames, resulting from a permutation test (p-value<0.01 for at least 6 time-bins, 60 ms). Correction for multiple comparisons was performed using FDR, α ≤ 0.01. Reprinted from [186], eLifeSciencesPublications.

the chosen item when the fixated circle had a higher number of dots (mean $\beta_{\text{Most}} = 1.581$). In the *fewest* frame, the effect flipped: it was more likely that the last circle seen was chosen when it had fewer dots (mean $\beta_{\text{Few}} = -0.944$; $\Delta\beta_{\text{Most-Few}}$: $t(31)=3.727$, p<0.001).
The previous set of analyses shows that the last fixation is modulated by the difference in evidence according to the goal that the participant is set to achieve. However, since the last fixation is in general followed by the participant response, one could suspect that the goal-dependent modulation of attention (i.e., $\Delta DT$) we identified in our choice regression analysis (Figure 3.2) is entirely driven by the final fixation. This would be problematic since one would have similar results to the one presented in Figure 3.2 even if participants’ pattern of attention is not modulated by the goal (i.e., attention is directed in both frames to the most valuable item) or even if the pattern of fixation, before the last fixation, is random. To control for this possibility, we performed a series of further analyses:

First of all, we repeated the analysis presented in the previous section (hierarchical choice regression – Figure 3.2), removing the last two fixations when calculating the $\Delta DT$. Note that we removed the last two fixations and not just the last one to avoid statistical artefacts (i.e. since the final fixation is mostly directed towards the chosen item there would be an increased probability that the second to the last fixation is on the unchosen item). We found that once removed the last two fixations the pattern of results is unchanged (please see [186] for details).

Second, we specifically investigated the middle fixations. Previous studies [33, 34, 121] have reported that middle fixations duration increases when the difference in value ratings (or perceptual evidence) of the fixated minus unfixed item increases. We replicated this result for our like and most frames but critically the effect was reversed in dislike and fewest frames (i.e. middle fixations durations decreased when the relative value of the fixated item was higher). The results suggesting that the goal-relevant modulation of attention affects also the middle fixations is presented in [186].

Finally, we investigated in more detail how the relation between attentional allocation and difference in value or perceptual evidence changed over time in the context of the goal manipulation. We calculated the Pearson correlation between fixation position ($0$: left, $1$:right) and the difference in evidence (i.e. $\Delta Value$ or $\Delta Dots$, in both cases right – left item) at different time points (Figure 3.3C). We observed that after an initial phase in which there was no clear gaze preference for any of the items (note that given the gaze-contingent design participants must explore both alternatives), fixations were correlated with the frame-relevant item: during like frame, fixations positions were positively correlated with $\Delta Value$, that is the fixations were directed towards the item with higher value; during dislike frame the behaviour was the opposite: fixations were negatively correlated with $\Delta Value$, indicating a preference for the option with a lower value. Note that these results are in line with the ones reported by [180]. We see a very similar pattern of results in the Perceptual Experiment too (Figure 3.3D).

3.4.3 Attentional model: GLAM

To gain further insights into the dynamic of the information accumulation process, we modelled the data from both experiments by adapting a Gaze-weighted Linear Accumulator Model (GLAM) recently developed by [35]. The GLAM belongs to the family of race models and approximates the aDDM model [33, 34] in which the dynamic aspect is discarded, favouring a more efficient estimation of the parameters. This model was chosen since, unlike the aDDM, it allowed us to test the prediction of the confidence measures as the balance of evidence [124, 23, 66]. Crucially, in both experiments,
we used goal-relevant evidence (not the value or the number of dots) to fit the models in the dislike and fewest frames (for further details see the Materials and methods Attentional Model: Glam section). For further details on the modelling results on Confidence, please refer to Chapter 7.

### 3.4.3.1 Parameter fit and simulations

#### 3.4.3.1.1 Value experiment

The simulations estimated with the parameters fitted for like and dislike frames (even-trials) reproduced the behaviour observed in the data not used to fit the model (odd-trials). In both like and dislike frames, the model replicated the observed decrease of RT when $|\Delta \text{Value}|$ is high, that is, the increase in speed of response in easier trials (bigger value difference). The RT simulated by the models significantly correlated with the RT values observed in participants’ odd-numbered trials (Like: $r(29)=0.90$, $p<0.001$; Dislike: $r(29)=0.89$, $p<0.001$) (Figure 3.4A). In the like frame, the model also correctly predicted a higher probability of choosing the right item when $\Delta \text{Value}$ is higher. In the dislike frame, the model captured the change in the task goal and predicted that the selection of the right item will occur when $-\Delta \text{Value}$ is higher, that is when the value of the left item is higher. Overall, in both frames, the observed and predicted probabilities of choosing the most valuable item were significantly correlated (Like: $r(29)=0.80$, $p<0.001$; Dislike: $r(29)=0.79$, $p<0.001$) (Figure 3.4B).

In both frames, the models also predicted choice depending on the difference in gaze ($\Delta \text{Gaze} = g_{\text{right}} - g_{\text{left}}$), that is, the probability of choosing the right item increases when the time spent observing that item is higher. However, in this case, we cannot say if gaze allocation itself is predicting choice if we do not account for the effect of $|\Delta \text{Value}|$. To account for the relationship between choice and gaze, we used a measure devised by [35], ‘gaze influence’. Gaze influence is calculated by taking the actual choice (1 or 0 for right or left choice, respectively) and subtracting the probability of choosing the right item given by a logistic regression for $\Delta \text{Value}$ calculated from actual behaviour. The averaged ‘residual’ choice probability indicates the existence of a positive or negative gaze advantage. Then, we compared the gaze influence predicted by GLAM with the empirical one observed for each participant. As in [35], most of the participants had a positive gaze influence and it was properly predicted by the model in both frames (Like: $r(29)=0.68$, $p<0.001$; Dislike: $r(29)=0.63$, $p<0.001$) (Figure 3.4C).

#### 3.4.3.1.2 Perceptual experiment

As in the Value Experiment, we fitted the GLAM to the data and we conducted model simulations. Again, these simulations showed that we could recover most of the behavioural patterns observed in participants. We replicated the relationship between RT and $|\Delta \text{Dots}|$ (Most: $r(26)=0.97$, $p<0.001$; Fewest: $r(26)=0.98$, $p<0.001$) (Figure 3.4D). As in the value-based experiment, the model also predicted a higher probability of choosing the right-hand item when $\Delta \text{Dots}$ is higher in the most frame and when $-\Delta \text{Dots}$ is higher in the fewest frame. However, in the Perceptual Experiment, the simulated choices only in the fewest frame were significantly correlated with the observed data, although we observed a non-significant trend in the most frame (Most: $r(26)=0.69$, $p<0.001$; Fewest: $r(26)=0.37$, $p=0.051$) (Figure 3.4E). In both frames, we observed that the model predicted that choice was linked to
Figure 3.4: Individual out-of-sample GLAM predictions for behavioural measures in Value (A–C) and Perceptual Experiments (D–F). In value-based decisions, (A) the model predicts individuals’ mean RT; (B) the probability of choosing the item with higher value in like frame, and the item with lower value in dislike frame; and (C) the influence of gaze on choice probability. In the Perceptual Experiment, (D) the model also predicts RT and (F) gaze influence. (E) The model significantly predicts the probability of choosing the best alternative in the fewest frame only (in the most frame a trend was found). The results corresponding to the models fitted with like/most frame data are presented in blue, and with dislike/fewest frame data in red. Dots depict the average of observed and predicted measures for each participant. Lines depict the slope of the correlation between observations and predictions. Mean 95% confidence intervals are represented by the shadowed region in blue or red, with full colour representing Value Experiment and striped colour Perceptual Experiment. All model predictions are simulated using parameters estimated from individual fits for even-numbered trials. Reprinted from [186], eLife Sciences Publications.

DGaze and, as in the Value Experiment, we show that the gaze influence predicted by the model is indeed observed in the data (Most: r(26)=0.65, p<0.001; Fewest: r(26)=0.47, p<0.05) (Figure 3.4F). Results of the models fitted without accounting for the change in goal-relevant evidence provided a poor fit of the data. Additionally, we were able to mirror the results obtained with GLAM using aDDM [33, 121]. For dislike and fewest frames, the best model was the one fitted using goal-relevant evidence Please see [186] for further details.

3.5 DISCUSSION

In this study, we investigated how framing affects the way in which information is acquired and integrated during value-based and perceptual choices. Here, using psychophysics together with computational and economic models we have been able to discern between two contrasting hypotheses. The first one, currently the dominant one in the field of neuroeconomics, proposes that attention modulates (either by biasing or boosting) a value integration that starts at the beginning of the deliberation process. Subsequently, at the time of the decision, the participant would give the appropriate response (in
our task accepting the option with the highest value or rejecting the one with the lowest one) using the value estimate constructed during this deliberation phase. The second hypothesis suggests that, from the very start of the deliberation process, the task-frame (goal) influences the type of information that is integrated. In this second scenario, attention is not automatically attracted to high value items to facilitate their accumulation but has a more general role in prioritising the type of information that is useful for achieving the current behavioural goal. Importantly, these two hypotheses make very distinct predictions about the pattern of attention and suggest very different cognitive architectures underpinning the decision process.

Our results favour the second hypothesis: specifically, we show that, in both perceptual and value-based tasks, attention is allocated depending on the behavioural goal of the task. Although our study does not directly contradict previous findings [33, 34, 169, 179], it adds nuance to the view that this is a process specifically tied to value integration (defined as a hedonic or reward attribute). Our findings speak in favour of a more general role played by attention in prioritising the information needed to fulfil a behavioural goal in both value and perceptual choices [188, 180, 172]. Importantly, the seeking of goal-relevant information is observed during the trial, opposing the assumption that attentional sampling is random except for the last fixation [33, 34](see [173, 171] for additional support for this idea). Pavlovian influences have been proposed to play a key role in the context of accept/reject framing manipulation [101, 189, 190, 191]. In this case, Pavlovian responses associated with predictions of reward (e.g., water or food in deprived animals) usually involve a vigorous active approach and engagement with the actions leading to the outcome [180]. This could translate into a deployment of attention towards preferred options [180], similar to a mouse looking for the button to press and obtain food. On the other hand, in reject frames, when the animal is exposed to negative outcomes (e.g., electric shocks or other punishment), animals express in general a behavioural inhibition, avoiding the aversive stimulus or inhibiting behaviour [192, 193]. However, the "reject" frame of our experiment did not fit this description, since participants engaged more with the item they disliked more, i.e., they paid attention to the stimulus that was relatively more aversive before choosing it. Additionally, the fact that we found almost identical results in a follow-up perceptual study in which the choice was not framed in terms of ‘accept’ or ‘reject’ but using a different kind of instruction (i.e. ‘choose the option with fewer or more dots’) suggests that attention acts on a more fundamental mechanism of information processing that goes beyond simple Pavlovian influences.

The idea that the goal of the task plays a central role in shaping value-based decisions should not be surprising. Indeed, value-based decision is often called goal-directed choice. Nevertheless, there has been a surprisingly little amount of experimental work in which the behavioural goal has been directly manipulated as the key experimental variable for studying the relation between attention and value. Notable exceptions are two recent studies from [98] and [180]. In the first study ([98]), participants were shown a set of four items and asked, in half of the trials, to determine the best item and, in the second half, the worst item. In line with our findings, they found that behaviour and neural activity in the ‘value network’, including vmPFC and striatum, was determined by goal-congruency and did not simply reflect the expected reward. In the second study, [180] implemented a design similar to our value-based experiment in which participants were required to
indicate the item to keep and the one to discard. They found, similarly to our findings in the value-based experiment, that the overall pattern of attention was mostly allocated according to the task goal. However, in the first few hundred milliseconds, these authors found that attention was directed more prominently to the most valuable item in both conditions. We did not replicate this last finding in our experiment (see Figure 3.3C and D, showing that fixations were randomly allocated during the early moments of the trial). One possible reason for this discrepancy is that the experiment by Kovach and colleagues [180] presented both items on the screen at the beginning of the task – unlike in our task, in which the item was presented in a gaze-contingent way (to avoid processing in the visual periphery). This setting might have triggered an initial and transitory bottom-up attention grab from the most valuable (and often most salient) item before the accumulation process started.

The most far-reaching conclusion of our work is that context and behavioural demand have a powerful effect on how information is accumulated and processed. Notably, our data show that this is a general effect that spans both more complex value-based choices and simpler perceptual choices. We conclude that, given the limited computational resources of the brain, humans have developed a mechanism that prioritises the processing or recollection of the information that is most relevant for the behavioural response that is required. This has profound implications when we think about the widespread effect of contextual information on decision-making that has been at the core of the research in psychology, behavioural economics and more recently neuroeconomics [194, 195, 196, 101, 197]. Most of these contextual or framing effects have been labelled as ‘biases’ because, once one strips away the context, the actual available options should remain identical. However, this perspective may not be putting enough emphasis on the fact that the decision-maker has to construct low-dimensional (and therefore imperfect) representations of the decision problem. As we have shown here, from the very beginning of the deliberation process, the context — even when it is simple (like/dislike, most/fewest) or irrelevant from the experimenter’s perspective — affects which information is processed, recalled, or attended to, with effects that spread into post-decision processing such as confidence estimation (please see Chapter 7 for details on the confidence effects). This, as a consequence, will produce profoundly dissimilar representations according to the behavioural goal set by the context. With this shift of perspective, it may well be the case that many of the so-called ‘biases’ will be shown in a new light, given that participants are dealing with very different choices once the behavioural goal changes. This viewpoint might provide a more encouraging picture of the human mind, by suggesting that evolution has equipped us well to deal with ever-changing environments in the face of limited computational resources.
Part II

BRAIN REPRESENTATIONS AND GOALS IN HUMAN DECISIONS
So far, I have presented how the sampling stage in decisions is modulated by participants’ goals. But the dynamic attentional component is just an external aspect of the entire process of deliberation. For any decision, the accumulation of evidence should be reflected in internal processes in the brain, appearing as specific brain activity patterns. In the following section, I focus on how the brain generates representations of the task to be completed and the available alternatives during goal-dependent decisions. In the following neuroimaging studies, I concentrate on two important aspects of the decision process: abstraction learning and preference evaluation. In Chapter 4, I studied how participants generate abstractions from multiple sensory features, adjusting them to the goal of the task. I showed how brain regions computing value signals, e.g., the ventromedial prefrontal cortex (vmPFC), guide the selection of relevant features, hinting at a goal-dependent role during learning. In the second study (Chapter 5), I presented the evaluation of single items and their brain representations in a value-based context. Item representations in brain regions were affected by the task, not only by the features of the items. Specific areas represented context information and hippocampal activity seems to represent specific features for each context. These results support that at the level of subjective preferences, goals impact brain representations to generate task-appropriate mappings.
GOAL-DIRECTED REPRESENTATIONS IN LEARNING: VALUE SIGNAL GUIDES ABSTRACTION DURING LEARNING

4.1 SUMMARY

The human brain excels at constructing and using abstractions, such as rules, or concepts. This is a fundamental step to efficiently identify the relevant features and strategies to fulfil the goals. In the present study, we used neuroimaging to show a mechanism of abstraction built upon the valuation of sensory features. Human volunteers learned novel association rules based on simple visual features. Reinforcement-learning models revealed that high-value abstract representations increasingly guided participant behaviour along the learning process, resulting in better choices and higher subjective confidence. We also found that the brain area computing value signals – the ventromedial prefrontal cortex – prioritised and selected latent task elements during abstraction, both locally and through its connection to the visual cortex. These findings provide a novel interpretation of value as a goal-dependent, key factor in forging abstract representations.

4.2 INTRODUCTION

Successful interaction with our environment is based on the capacity to correctly perform our goals. For example, if our goal is to drive from A to B, to achieve it involves many important steps such as navigating the traffic (e.g., knowing that a red light indicates to stop) or operating the automobile (e.g., pressing the gas and not the brake when the green light appears). Therefore, fulfilling a goal involves a process of learning where we highlight relevant aspects and silence irrelevant features.

The unique ability of the human mind to organise information beyond the immediate sensory reality is the basis of the process of abstraction. Abstractions are everywhere, from high art and government plans to online shopping and city maps. Goal implementation relies on these abstractions, which are ultimately represented in the human brain. The need for abstractions is well exemplified in reinforcement learning (RL). Classic RL algorithms rapidly collapse when an agent has to deal with complex and/or multidimensional problems [198, 199, 200]. However, when the agent can ‘abstract’ the current state to a lower dimensional manifold, representing only relevant features, responses become more flexible and efficient [201, 202, 135].

From a psychological or neuroeconomic point of view, the process of learning is guided by task goals that determine what is valuable [203, 204, 205]. For example, if I want to reach my destination, the brand of air freshener in my car is not as important to fulfil the goal as the level of gasoline [99].
Hence, we hypothesised that valuation processes are directly related to the construction of abstractions.

Value representations have been linked to neural activity in the ventromedial prefrontal cortex (vmPFC) in the context of economic choices \[205, 44\]. Recently, the role of the vmPFC has been further expanded including its role in the computation of confidence \[66, 206, 67, 207\]. However, in these cases, the characterisation of value has mostly been focused on the hedonic and rewarding aspects instead of its broader functional role. In the field of memory, vmPFC plays a crucial role in the formation of schemas or conceptual knowledge \[88, 208, 209, 210, 211\], as well as generalisations \[212\]. The vmPFC also combines goal-relevant information from elsewhere in the brain \[213\]. Considering its connectivity pattern \[214\], the vmPFC is well suited to serve a pivotal function in the circuit that involves the hippocampal formation (HPC) and the orbitofrontal cortex (OFC), dedicated to extracting latent task information and regularities important for navigating behavioural goals \[135, 215, 216, 217, 218\]. Thus, this study aims to show how abstraction emerges during learning and to investigate how the brain, and specifically the vmPFC, uses valuation upon low-level sensory features to generate abstract representations.

We designed a task in which human participants repeatedly learned novel association rules to perform their goal, while fMRI recorded their brain activity. Reinforcement learning (RL) modelling was used to track participants’ valuation processes and to distinguish at a behavioural and neural level between two potential learning strategies: the full representation of sensory features or the dimensionality reduction through abstractions. Participants’ confidence in having performed the task well was positively correlated with their ability to abstract. We found that vmPFC and its connection to the visual cortex constructed abstract representations through a goal-dependent valuation process that is implemented as top-down control of sensory cortices.

### 4.3 METHODS

#### 4.3.1 Participants

Forty-six participants with normal or corrected-to-normal vision were recruited for the main experiment (learning task). The sample size was chosen according to prior work and recommendations on model-based fMRI studies \[219\]. Based on pilot data and the available scanning time in one session (60 min), we set the following conditions of exclusion: failure to learn the association in three blocks or more (i.e., reaching a block limit of 80 trials without having learned the association), or failure to complete at least 11 blocks in the allocated time. Eleven participants were removed based on the above-predetermined conditions, 2 of which did not go past the initial practice stages. Additionally, two more participants were removed due to head motion artifacts. Thus, 33 participants \(22.4 \pm 0.3\) y.o.; eight females) were included in the main analyses. All results presented up are from the 33 participants who completed the learning task.

All experiments and data analyses were conducted at the Advanced Telecommunications Research Institute International (ATR) in Japan. The study was approved by the Institutional Review Board of ATR with ethics protocol numbers 18–122, 19–122, 20–122. All participants gave written informed consent.
### 4.3.2 Learning task

The task consisted of learning the fruit preference of Pacman-like characters. These characters had three features, each with two levels (colour: green vs red, stripe orientation: horizontal vs vertical, mouth direction: left vs right). On each trial, a character composed of a unique combination of the three features was presented. The experimental session was divided into blocks, for each of which a specific rule directed the association between features and preferred fruits. There were always two relevant features and one irrelevant feature, but these changed randomly at the beginning of each block. Blocks could thus be of three types: CO (colour-orientation), CD (colour-direction), and OD (orientation-direction). Furthermore, to avoid a simple logical deduction of the rule after one trial, we introduced the following pairings: The four possible combinations of two relevant features with two levels were paired with the two fruits in both a symmetric or asymmetric fashion - 2x2 or 3x1. The appearance of the two fruits was also randomly changed at the beginning of each new block (see Figure 4.1B,e.g., green-vertical: fruit 1, green-horizontal: fruit 2, red-vertical: fruit 1, red-horizontal: fruit two or green-vertical: fruit 2, green-horizontal: fruit 2, red-vertical: fruit 1, red-horizontal: fruit 2).

Each trial started with a black screen for 1 s, followed by the character presented for 2 s. Then, while the character continued to be at the centre of the screen, the two fruit options appeared below, to the right and left sides. Participants had 2 s to indicate the preferred fruit by pressing a button (right for the fruit to the right, left for vice versa). Upon registering a participant’s choice, a coloured square appeared around the chosen fruit: green if the choice was correct, red otherwise. The square remained for 1 s, following which the trial ended with a variable ITI - bringing the trial to a fixed 8 s duration.

Participants were simply instructed that they had to learn the correct rule for each block, and the rule itself (relevant features + association type) was hidden. Each block contained up to 80 trials, but a block could end earlier if participants learned the target rule. Learning was measured as a set of correct trials (between 8 and 12, determined randomly in each block). Participants were instructed that each correct choice added one point, while incorrect choices did not alter the balance. They were further told that points obtained would be weighted by the speed of learning on that block. That is, the faster the learning, the greater the net worth of the points. The end of a block was explicitly signalled by presenting the reward obtained on the screen. Monetary reward was computed at the end of each block according to the formula:

\[
R = A \times (\frac{\sum pts}{\sum tr}) - (\sum tr - mcs) \times a(t)
\]

where \(R\) is the reward obtained in that block, \(A\) the maximum available reward (150 JPY), \(\sum pts\) the sum of correct responses, \(\sum tr\) the number of trials, \(mcs\) the maximum length of a correct strike (12 trials), and \(a\) is a scaling factor (\(a = 1.5\)). This formula ensures time-dependent decay of the reward that approximately follows a quadratic fit. In case participants completed a block in less than 12 trials, the reward was capped at 150 JPY.

The maximum terminal monetary reward over the whole session was set at 3,000 JPY. On average, participants earned 1251 ± 46 JPY (blocks in which participants failed to learn the association within the 80-trial limit were not rewarded). For each experimental session, there was a sequence of 20
blocks that was pre-generated pseudo-randomly, and on average, participants completed $13.6 \pm 0.3$ blocks.

For sessions done in the MR scanner, participants were instructed to use their dominant hand to press buttons on a dual-button response pad to register their choices. Concordance between responses and buttons was indicated on the display, and importantly, randomly changed across trials to avoid motor preparation confounds (i.e. associating a given preference choice with a specific button press). The task was coded with PsychoPy v1.82.01 [220].

4.3.3 Computational modelling part 1: mixture-of-experts RL model

We built on a standard RL model [200] and prior work in machine learning and robotics to derive the mixture-of-experts architecture [221, 222, 223]. In this work, the mixture-of-experts architecture is composed of several ‘expert’ RL modules, each tracking a different representational space, and each with its own value function. In each trial, the action selected by the mixture-of-experts RL model is given by the weighted sum of the actions computed by the experts. The weight reflects the responsibility of each expert, which is computed from the SoftMax of the squared prediction error. In this section, we define the general mixture-of-expert RL model and in the next section we define each specific expert, based on the task-state representations being used.

Formally, the mixture-of-expert RL model global action is defined as:

$$A_t = \sum_{j=1}^{N} \lambda_j^t a_j^t$$ (2)

where $N$ is the number of experts, $\lambda$ the responsibility signal, and $a$ the action selected by the $j$th-model. Thus, $\lambda$ is defined as:

$$\lambda_j^t = \exp\left(\frac{-RPE_j^t - 1}{\nu}\right) / \left\{ \sum_{k=1}^{N} \exp\left(\frac{-RPE_k^t - 1}{\nu}\right) \right\}$$ (3)

where $N$ is the same as above, $\nu$ is the RPE (reward prediction error) variance. Expert uncertainty $RPE_j$ is defined as:

$$RPE_j^t = \gamma RPE_j^{t-1} + (1 - \gamma)(RPE_j^t)^2$$ (4)

where $\gamma$ is the forgetting factor that controls the strength of the impact of prior experience on the current uncertainty estimate. The most recent RPE is computed as:

$$RPE_j^t = O - Q^j(S_t, A_t)$$ (5)

where $O$ is the outcome (reward: 1, no reward: 0), $Q$ is the value function, $S$ the state for the current expert, and $A$ is the global action computed with Equation (2). The update to the value function can therefore be computed as:

$$\Delta Q^j(S_t, A_t) = \lambda_j^t aRPE_j^t$$ (6)

where $\lambda$ is the responsibility signal computed with Equation (3), $a$ is the learning rate (assumed equal for all experts), and $RPE$ is computed with Equation (5). Finally, for each expert, the action $a$ at trial $t$ is taken as the argmax of the value function, as follows:
\[ a^t_i = \text{argmax}[Q^t(S_t, a)](7) \]

where \( Q \) is the value function, \( S \) the state at current trial, and \( a \) the two possible actions.

Hyperparameters estimated through likelihood maximisation were the learning rate \( \alpha \), the forgetting factor \( \gamma \), and the RPE variance \( \nu \).

4.3.4 Computational modelling part 2: Feature RL and Abstract RLs

Each (expert) RL algorithm is built on a standard RL model \([200]\) to derive variants that differ in the number and type of states visited. Here, a state is defined as a combination of features. Feature RL has \( 2^3 = 8 \) states, where each state was given by the combination of all three features (e.g. colour, stripe orientation, mouth direction: green, vertical, left). Abstract RL is designed with \( 2^2 = 4 \) states, where each state was given by the combination of two features.

Note that the number of states does not change for different blocks, only features used to determine them. These learning models create individual estimates of how participant action-selection depended on features they attended and their past reward history. Both RL models share the same underlying structure and are formally described as:

\[ Q(s, a) \leftarrow Q(s, a) + \alpha(r - Q(s, a))(8) \]

where \( Q(s, a) \) in Equation (8) is the value function of selecting either fruit-option \( a \) for Pacman-state \( s \). The value of the action selected in the current trial is updated based on the difference between the expected value and the actual outcome (reward or no reward). This difference corresponds to the reward prediction error (RPE). The degree to which this update affects the expected value depends on the learning rate \( \alpha \). For larger \( \alpha \), more recent outcomes will have a strong impact. On the contrary, for small \( \alpha \) recent outcomes will have little effect on expectations. Only the value function of the selected action, which is state-contingent in Equation (8), is updated. Expected values of the two actions are combined to compute the probability \( p \) of predicting each outcome using a SoftMax (logistic) choice rule:

\[ P_{s, A} = 1/ (1 + \exp(\beta(Q(s, a_1) - Q(s, a_2))))(9) \]

The greediness hyperparameter \( \beta \) controls how much the difference between the two expected values for \( a_1 \) and \( a_2 \) influences choices. Hyperparameters estimated through likelihood maximisation were the learning rate \( \alpha \), and the greediness (inverse temperature) \( \beta \).

4.3.5 Procedures for model fitting, simulations

Hierarchical Bayesian Inference (HBI) was used to fit the models to participant behavioural data, enabling precise estimates of hyperparameters at the block level for each participant \([224]\). Hyperparameters were selected by maximising the likelihood of estimated actions, given the true actions. For the mixture-of-experts architecture, we fit the model on all participants block-by-block to estimate hyperparameters at the single-block and single-participant level. For the subsequent direct comparison between Feature RL and Abstract RL models, we used HBI for concurrent model fitting and comparison at the
single-block and single-participant basis. The model comparison provided the likelihood that each RL algorithm best explained participants’ choice data. That is, it was a proxy to whether learning followed a Feature RL or Abstract RL strategy. Because the fitting was done block-by-block, with a hierarchical approach considering all participants, blocks were first sorted according to their lengths, from longer to shorter, at the participant level. This ensured that each block of a given participant was more similar to the blocks of all other participants, thus avoiding unwanted effects in the fitting due to block length. The HBI procedure was then implemented on all participant data, proceeding block-by-block.

We also simulated model action-selection behaviours to benchmark their similarity to human behaviour, and in the case of Feature RL vs Abstract RL comparisons, to additionally compare their formal learning efficiency. In the case of the mixture-of-experts RL architecture, we simply used estimated hyperparameters to simulate 45 artificial agents, each completing 100 blocks. The simulation allowed us to compute, for each expert RL module, the mean responsibility signal, and the mean expected value across all states for the chosen action. Additionally, we also computed the learning speed (time to learn the rule of a block) and compared it with the learning speed of human participants.

In the case of the simple Feature RL and Abstract RL models, we added noise to the state information in order to have a more realistic behaviour (from the perspective of human participants). In the empirical data, the action (fruit selection) in the first trial of a new block was always chosen at random because participants did not have access to the appropriate representations (states). In later trials, participants may have followed specific strategies. For model simulations, we simply assumed that states were corrupted by a decaying noise function:

\[ n_t = n_0(1/t^{1/\text{rte}})(10) \]

where \( n_t \) is the noise level at trial \( t \), \( n_0 \) the initial noise level (randomly drawn from a uniform distribution within the interval \([0.3, 0.7]\)), and \( \text{rte} \) was the decay rate, which was set to 3. This meant that in early trials in a block, there was a higher chance of basing the action on the wrong representation (at random), rather than following the appropriate value function. Actions in later trials had a decreasing probability of being chosen at random. This approach is a combination of the classic \( \epsilon \)-greedy policy and the standard SoftMax action-selection policy in RL. Hyperparameter values were sampled from obtained participant data maximum likelihood fits. We simulated 45 artificial agents solving 20 blocks each. The procedure was repeated 100 times for each block with new random seeds. We used two metrics to compare the efficiency of the two models: learning speed (same as above, the time to learn the rule of a block), as well as the fraction of failed blocks (blocks in which the rule was not learned with the 80-trials limit).

### 4.3.6 fMRI scans: acquisition and protocol

All scanning sessions employed a 3T MRI scanner (Siemens, Prisma) with a 64-channel head coil in the ATR Brain Activation Imaging Centre. Gradient T2*-weighted EPI (echoplanar) functional images with blood-oxygen-level-dependent (BOLD)-sensitive contrast and multi-band acceleration factor six were acquired [225, 226]. Imaging parameters: 72 contiguous slices (TR = 1 s, TE = 30 ms, flip angle = 60 deg, voxel size = 2×2×2 mm³, 0 mm slice
gap) oriented parallel to the AC-PC plane were acquired, covering the entire brain. T1-weighted images (MP-RAGE; 256 slices, TR = 2 s, TE = 26 ms, flip angle = 80 deg, voxel size = 1 × 1 × 1 mm³, 0 mm slice gap) were also acquired at the end of the first session.

4.3.7 fMRI scans: standard and parametric general linear models

BOLD-signal image analysis was performed with SPM12 [http://www.fil.ion.ucl.ac.uk/spm/], running on MATLAB v9.1.0.96 (2016b) and v9.5.0.94 (2018b). fMRI data for the initial 10 s of each block were discarded due to possible unsaturated T1 effects. Raw functional images underwent realignment to the first image of each session. Structural images were re-registered to mean EPI images and segmented into grey and white matter. Segmentation parameters were then used to normalise (MNI) and bias-correct the functional images. Normalised images were smoothed using a Gaussian kernel of 7 mm full-width at half-maximum.

GLM1: regressors of interest included ‘High value’, ‘Low value’ (trials were labelled as such based on the median split of the trial-by-trial expected value for the chosen option computed from the best fitting algorithm - Feature RL or Abstract RL), ‘Feature RL’, ‘Abstract RL’ (trials were labelled as such based on the best fitting algorithm at the block level). For all, we generated boxcar regressors at the beginning of the visual stimulus (character) presentation, with a duration of 1 s. Contrast of [1 -1] or [−1 1] were applied to the regressors ‘High value’ - ‘Low value’, and ‘Feature RL’ - ‘Abstract RL’. Specific regressors of no interest included the time in the experiment: ‘early’ (all trials falling within the first half of the experiment), and ‘late’ (all trials falling in the second half of the experiment). The early/late split was done according to the total number of trials: taking as ‘early’, trials from the first block onward, adding blocks until the trial sum exceeded the total trials number divided by two.

GLM2 (psychophysiological interaction, PPI): the seed was defined as a sphere (radius = 6 mm) centred around the individual peak voxel from the ‘High value’ > ‘Low value’ group-level contrast, within the vmPFC (peak coordinates [2 50 -10], radius 25 mm). The ROI mask was defined individually to account for possible differences among participants. Two participants were excluded from this analysis because they did not show a significant cluster of voxels in the bounding sphere (even at very lenient thresholds). The GLM for the PPI included three regressors (the PPI, the mean BOLD signal of the seed region, and the psychological interaction), as well as nuisance regressors described below.

For all GLM analyses, additional regressors of no interest included a parametric regressor for reaction time, regressors for each trial event (fixation, fruit options presentation, choice, button press [left, right], ITI), block regressors, the six head motion realignment parameters, framewise displacement (FD) computed as the sum of the absolute values of the derivatives of the six realignment parameters, the TR-by-TR mean signal in white matter, and the TR-by-TR mean signal in cerebrospinal fluid.

Second-level group contrasts from all models were calculated as one-sample t-tests against zero for each first-level linear contrast. Statistical maps were z-transformed, and then reported at a threshold level of P(fpr) < 0.001 (Z > 3.09, false positive control meaning cluster forming threshold), unless otherwise specified. Statistical maps were projected onto a canonical MNI template.
4.3 Methods

with MRICroGL [https://www.nitrc.org/projects/mricrogl/] or a glassbrain MNI template with Nilearn 0.7.1 [https://nilearn.github.io/index.html].

4.3.8 fMRI scans: pre-processing for decoding

As above, the fMRI data for the initial 10 s of each run were discarded due to possible unsaturated T1 effects. BOLD signals in native space were pre-processed in MATLAB v7.13 (R2011b) (MathWorks) with the mrVista software package for MATLAB [http://vistalab.stanford.edu/software/]. All functional images underwent 3D motion correction. No spatial or temporal smoothing was applied. Rigid-body transformations were performed to align functional images to the structural image for each participant. One region of interest (ROI), the HPC, was anatomically defined through cortical reconstruction and volumetric segmentation using the Freesurfer software [http://surfer.nmr.mgh.harvard.edu/]. Furthermore, VC subregions V1, V2, and V3 were also automatically defined based on a probabilistic map atlas [227]. The vmPFC ROI was defined as the significant voxels from the GLM ‘High value’ > ‘Low value’ contrast at $p(fpr) < 0.0001$, within the OFC. All subsequent analyses were performed using MATLAB v9.5.0.94 (r2018b). Once ROIs were individually identified, time-courses of BOLD signal intensities were extracted from each voxel in each ROI and shifted by 6 s to account for the hemodynamic delay (assumed fixed). A linear trend was removed from time-courses, which were further z-score-normalised for each voxel in each block to minimise baseline differences across blocks. Data samples for computing individual feature identity decoders were created by averaging BOLD signal intensities of each voxel over two volumes, corresponding to the 2 s from stimulus (character) onset to fruit options onset.

4.3.9 Decoding: multivoxel pattern analysis (MVPA)

All ROI-based MVPA analyses followed the same procedure. We used sparse logistic regression (SLR) [228], to automatically select the most relevant voxels for the classification problem. This allowed us to construct several binary classifiers (e.g. feature id.: colour - red vs green, stripes orientation - horizontal vs vertical, mouth direction - right vs left).

Cross-validation was used for each MVP analysis to evaluate the predictive power of the trained (fitted) model. In the primary analysis (reported in Figure 4.5C), cross-validation was done with a leave-one-run-out scheme, whereby each run was iteratively held out as a test set, and all other runs were used for training the algorithm. The final accuracy was taken as the averaged accuracy across the runs. This approach is generally used because there may be subtle differences across runs: holding out one run as a test ensures higher generalizability of the results while avoiding within-run information leaks. Yet, because of the nature of our task and experiment, the leave-one-run-out cross-validation leads to other confounds due to the varying number of training trials across classes (e.g. colour red vs green) or conditions (Feature RL vs Abstract RL blocks). To control for this idiosyncratic feature of our design, we performed a second cross-validation scheme. Here, we first merged the data from all blocks for each condition, and then computed the lowest bound of trial number from the minority class across conditions (e.g. if Feature RL had 128 trials labelled as ‘green’, and 109 as ‘red’, while Abstract RL had 94 trials labelled as ‘green’, and 99 labelled as ‘red’; then the minority class lowest bound was 94). In each fold (N folds = 20), a number of trials equivalent to 80% of the minority class lowest
bound were assigned to the training set from each class, and the remaining trials to the test set. The training samples were randomly chosen in each fold. Furthermore, for all MVP analysis, SLR classification was optimised by using an iterative approach [229] In each fold of the cross-validation, the feature-selection process was repeated 10 times. In each iteration, selected features (voxels) were removed from pattern vectors, and only features with unassigned weights were used for the next iteration. At the end of the cross-validation, test accuracies were averaged for each iteration across folds, to evaluate accuracy at each iteration. The number of iterations yielding the highest classification accuracy was then used for the final computation. Results report the cross-validated average of the best-yielding iteration.

4.4 RESULTS

Our experiment generated a scenario where the goal of the task was clear but the relevant features to achieve it had to be identified. In particular, the learning task we designed could be solved according to two strategies, based on the sampled task-space dimensionality. A simple, slower strategy akin to pattern recognition, and a more sophisticated one that required abstraction to use the underlying structure. Participants (N = 33) learned the fruit preference of Pacman-like characters formed by the combination of three visual features (colour, mouth direction, and stripe orientation, Figure 4.1 A–B). The preference was governed by a combination of two features, selected randomly by our computer program for each block (Figure 4.1 A–B). Learning the block rules essentially required participants to uncover hidden associations between features and fruits. Although participants were instructed that one feature was irrelevant, they did not know which. A block ended when a sequence of 8–12 (randomly set by our computer program) correct choices was detected or upon reaching its upper limit (80 trials). Variable stopping criteria were used to prevent participants from learning that a fixed sequence was predictive of block termination. During each trial, participants could see the outcome after selecting a fruit. A green box appeared around the chosen fruit if the choice was correct (red otherwise). Additionally, participants were instructed that the faster they learned a block rule, the higher the reward. At the end of the session, a final monetary reward was delivered, commensurate with participant performance (see Materials and methods). Participants failing to learn the association in three blocks or more (i.e. reaching a block limit of 80 trials without having learned the association), and/or failing to complete more than 10 blocks in the allocated time, were excluded (see Materials and methods). All main results reported in this chapter are from the included sample of N = 33 participants.

4.4.1 Behavioural accounts of learning

First, we verified that participants completed and learned the task properly. Within blocks, performance was higher than chance as early as the second trial (Figure 4.1 C, one-sample t-test against the mean of 0.5, trial 2: \( t_{32} = 4.13, P_{(FDR)} < 10^{-3} \), trial 3: \( t_{32} = 2.47, P_{(FDR)} = 0.014 \), all trials \( t > 3 \): \( P_{(FDR)} < 10^{-3} \). Considering the whole experimental session, learning speed (i.e., how quickly participants completed a given block) increased significantly across blocks (Figure 4.1 D, N = 11 time points, Pearson’s \( r = 0.80, p = 0.003 \)). These results not only confirmed that participants learned the task rule in each block, but also that they learned to use more efficient strategies. Notably, in
**Figure 4.1**: Learning task and behavioural results. (A) Task: participants learned the fruit preferences of Pacman-like characters, which changed on each block. (B) Associations could form in three ways: colour – stripe orientation, colour – mouth direction, and stripe orientation – mouth direction. The left-out feature was irrelevant. Examples of the two types of fruit associations. The four combinations arising from two features with two levels were divided into symmetric (2x2) and asymmetric (3x1) cases. f1-3: features 1 to 3; fruit: rule refers to the fruit as being the association rule. Both block types were included to prevent participants from learning rules by simple deduction. If all blocks had symmetric association rules and participants knew this, they could simply learn one feature-fruit association (e.g. green-vertical), and from there deduce all other combinations. Both the relevant features and the association types varied on a block-by-block basis. (C), Trial-by-trial ratio-correct improved as a measure of within-block learning. Dots represent the mean across participants, while error bars indicate the SEM, and the shaded area represents the 95% CI (N = 33). Participant-level ratio correct was computed for each trial across all completed blocks. (D), Learning speed was positively correlated with time, among participants. Learning speed was computed as the inverse of the max-normalised number of trials taken to complete a block. Thin grey lines represent the least square fits of individual participants, while the black line represents the group-average fit. The correlation was computed with group-averaged data points (N = 11). Average data points are plotted as coloured circles, the error bars are the SEM. (E), Confidence judgements were positively correlated with learning speed, among participants. Each dot represents data from one participant, and the thick line indicates the regression fit (N = 31 [2 missing data]). The experiment was conducted once (n = 33 biologically independent samples), **p<0.01. Reprinted from [23], eLife Sciences Publications.

this task, the only way to solve blocks faster was by using the correct subset of dimensions (the abstract representation). When at the end of a session, participants were asked about their degree of confidence in having performed the task well, their self-reports correlated with their learning speed (N = 31 [2 missing data], robust regression slope = 0.024, t_{29} = 3.27, p = 0.003, Figure 4.1E), but not with the overall number of trials, or the product of the proportion of successes (learning speed: Pearson’s r = 0.53, p = 0.002, total trials: r = -0.13, p = 0.47, test for difference in r: z = 2.71, p = 0.007; product of the proportion of successes: r = -0.06, p = 0.75, test for difference in r: z = 2.43, p = 0.015). We confirmed that block type (defined by relevant features, e.g., colour-orientation) or association type (e.g. symmetric 2x2) did not systematically affect learning speed, by pooling block-wise learning speed from all participants for each block or association type. None of the pairwise
tests survived multiple comparison correction (FDR). Excluded participants (see Materials and methods) had overall lower performance, although some had comparable ratios correct (Wilcoxon rank sum test, $z = 2.76$, $p = 0.006$).

### 4.4.2 Discovery of abstract representations

![Figure 4.2: Mixture of reinforcement learning (RL) experts and value computation.](image)

(A) Outline of the representational spaces of each RL algorithm comprising the mixture-of-experts architecture. (B) Illustration of the model architecture. See Methods for a formal description of the model. All experts had the same number of hyperparameters: the learning rate $\alpha$ (how much the latest outcome affected agent beliefs), the forgetting factor $\gamma$ (how much prior RPEs influenced current decisions), and the RPE variance $v$, modulating the sharpness with which the mixture-of-expert RL model should favour the best performing algorithm in the current trial. (C) The approach used for data analysis and model simulation. The model was first fitted to participant data with Hierarchical Bayesian Inference [224]. Estimated hyperparameters were used to compute value functions of participant data, as well as to generate new, artificial choice data and to compute simulated value functions. (D) Averaged expected value across all states for the chosen action in each RL expert, as well as responsibility signal for each model. Left: simulated data, right: participant empirical data. Dots represent individual agents (left) or participants (right). Bars indicate the mean and error bars depict the SEM. Statistical comparisons were performed with two-sided Wilcoxon signed rank tests. ***$p<0.001$. AbRL: Abstract RL, FeRL: Feature RL, AbRL$_{w1}$: wrong-1 Abstract RL, AbRL$_{w2}$: wrong-2 Abstract RL. (E) RPE variance was negatively correlated with learning speed (outliers removed, $N = 29$). Dots represent individual participant data. The thick line shows the linear regression fit. The experiment was conducted once ($n = 33$ biologically independent samples), * $p<0.05$. Reprinted from [230], eLife Sciences Publications.

Was participants’ learning behaviour guided by the selection of accurate representations? To answer this question, we built upon a classic RL algorithm (Q-learning) [231] in which state-action value functions (beliefs) used to predict future rewards, are updated according to the task state of a given trial and the action outcome. In this study, task states were defined by the number of feature combinations that the agent may track; hence, we
devised algorithms that differed in their state-space dimensionality. The first algorithm, called Feature RL, explicitly tracked all combinations of three features, \(2^3 = 8\) states (Figure 4.2A, top left). This algorithm is anchored at a low feature level because each combination of the three features results in a unique fingerprint – one simply learns direct pairings between visual patterns and fruits (actions). Conversely, the second algorithm, called Abstract RL, used a more compact or abstract state representation in which only two features are tracked. These compressed representations reduce the explored state-space by half, \(2^2 = 4\) states (Figure 4.2A, top right). In this task, as many as three Abstract RL in parallel were possible, one for each combination of two features.

The above four RL algorithms (Feature RL + three Abstract RL) were combined in a mixture-of-experts architecture \([232, 222, 223]\) (Figure 4.2B and Materials and methods). The key intuition behind this approach was that at the beginning of a new block, the agent did not know which abstract representation was correct (i.e., which features were relevant). Thus, the agent needed to learn which representations were most predictive of reward, to exploit the best representation for action selection. While all experts competed for action selection, their learning uncertainty (RPE: reward prediction error) determined their strength in doing so \([221, 223, 233]\). This architecture allowed us to track the value function of each RL expert separately while using a unique, global action in each trial.

Estimated hyperparameters (learning rate \(\alpha\), forgetting factor \(\gamma\), RPE variance \(\nu\)) were used to compute value functions of participant data, as well as to generate new, artificial choice data and value functions (Figure 4.2C, and Materials and methods). Simulations indicated that the expected value and responsibility were highest for the appropriate Abstract RL, followed by Feature RL, and the two Abstract RLs based on irrelevant features as the lowest (Figure 4.2D). Participant empirical data displayed the same pattern, whereby the value function and responsibility signal of the appropriate Abstract RL were higher than in other RL algorithms (Figure 4.2D, right side). Note that the large difference between appropriate Abstract RL and Feature RL was because the appropriate Abstract RL was an ‘oracle’: it had access to the correct low-dimensional state from the beginning. The RPE variance (hyperparameter \(\nu\)) adjusted the sharpness with which each RL’s (un)certainty was considered for expert weighting. Crucially, the variance \(\nu\) was associated with participant learning speed, such that participants who learned block rules quickly were sharper in expert selection (Figure 4.2E, \(N = 29\), robust regression slope = −1.02, \(t_{27} = -2.59, p = 0.015\)). These modelling results provided a first layer of support for the hypothesis that valuation is related to abstraction.

### 4.4.3 Behaviour shifts from Feature- to Abstraction-based reinforcement learning

The mixture-of-experts RL model revealed that participants who learned faster relied more on the best RL model value representations. Further, the modelling established that choices were mostly driven by either the appropriate Abstract RL or Feature RL, which had higher expected values (but note that the other Abstract RLs had mean values greater than a null level of 0.5), and higher responsibility \(\lambda\). It is important to highlight though that the mixture-of-experts RL might not reflect the actual algorithmic computation used by the participant in this task, but it provides a conceptual solution to the arbitration between representations/strategies. The model comparison
Figure 4.3: Feature RL vs Abstract RL are related to learning speed and the use of abstraction increases with experience. (A) Simulated learning speed and % of failed blocks for both Abstract RL and Feature RL. To make simulations more realistic, arbitrary noise was injected into the simulation, altering the state (see Materials and methods). N = 100 simulations of 45 agents. Right plot: bars represent the mean, error bars the SEM. (B) The relationship between the block-by-block, best-fitting model and learning speed of participants. Each dot represents one block from one participant, with data aggregated across all participants. Note that some dots fall beyond p=one or p=0. This effect occurs because dots were scattered with noise in their x-y coordinates for better visualisation. (C) Between participant correlations. Top: abstraction level vs learning speed. The abstraction level was computed as the average over all blocks completed by a given participant (code: Feature RL = 0, Abstract RL = 1). Bottom: confidence vs abstraction level. Dots represent individual participants (top: N = 33, bottom: N = 31, some dots are overlapping). (D) The learning rate was not symmetrically distributed across the two algorithms. (E) Greediness was not symmetrically distributed across the two algorithms. For both (D and E), each dot represents one block from one participant, with data aggregated across all participants. Histograms represent the distribution of data around the midline. (F) The number of participants for which Feature RL or Abstract RL best explained their choice behaviour in the first and last blocks of the experimental session. (G) The abstraction level was computed separately with blocks from the first half (early) and the latter half (late) sessions. (H) Participants count for the best fitting model, in each block. The experiment was conducted once (n = 33 biologically independent samples). * p<0.05, ** p<0.01, *** p<0.001. Adapted from [230], eLife Sciences Publications.

showed that Abstract RL and Feature RL in many cases offered a more parsimonious description of the participants behaviour. This is unsurprising since Feature RL is a simple model and Abstract RL is an oracle model – knowing which are the relevant feature. Hence, we next sought to explicitly explain participant choices and learning according to either Feature RL or Abstract RL strategy. Given the task space (Figure 4.2A), the only way to solve a block rule faster was to use abstract representations. As such, we
expect to observe a shift from Feature RL toward Abstract RL to occur with learning.

Both algorithms had two hyperparameters: the learning rate $\alpha$ and greediness $\beta$ (inverse temperature, the strength that expected value has on determining actions). Using the estimated hyperparameters, we generated new, synthetic data to evaluate how fast artificial agents, implementing either Feature RL or Abstract RL, solved the learning task (see Materials and methods). The simulations attested that Feature RL was slower and less efficient (Figure 4.3A), yielding lower learning speed and a higher percentage of failed blocks.

Model comparison at the single participant and block levels [224] provided a direct way to infer which algorithm was more likely to explain learning in any given block. Overall, similar proportions of blocks were classified as Feature RL and Abstract RL. This indicates that participants used both learning strategies (binomial test applied to all blocks: proportion of Abstract RL = 0.47 vs. equal level = 0.5, $P(212|449) = 0.26$, Figure 4.3B; two-sided t-test of participant-level proportions: lower, but close to 0.5, $t_{32} = -2.87, p = 0.007$, Figure 4.3B inset).

As suggested by the simulations (Figure 4.3A), the strategy that best explained participant block data accounted for the distribution of learning speed measures in each block. Where learning proceeded slowly, Feature RL was consistently predominant (Figure 4.3B), while the reverse happened in blocks where participants displayed fast learning (Figure 4.3B). Among participants, the degree of abstraction (propensity to use Abstract RL) correlated with the empirical learning speed ($N = 33$, robust regression, slope = 0.52, $t_{31} = 4.56, p = 7.64 \times 10^{-5}$, Figure 4.3C top). Participant confidence in having performed the task well was also significantly correlated with the degree of abstraction ($N = 31$, robust regression, slope = 0.026, $t_{29} = 2.69, p = 0.012$, Figure 4.3C, bottom). In addition to the finding that confidence related to learning speed (Figure 4.1E), these results raise intriguing questions about the function of metacognition, as participants appeared to comprehend their own ability to construct and use abstractions [234].

The two RL algorithms revealed a second aspect of learning. Considering all blocks regardless of fit (paired comparison), feature RL appeared to have higher learning rates $\alpha$ compared with Abstract RL (two-sided Wilcoxon rank sum test against median 0, $z = 14.33, p < 10^{-30}$, Figure 4.3D). A similar asymmetry was found with greediness (Figure 4.3E, two-sided Wilcoxon rank sum test against median 0, $z = 7.14, p < 10^{-10}$). Yet, more specifically, considering only the model (Feature RL or Abstract RL) which provided the best fit on a given block, resulted in Feature RL displaying lower learning rates and greediness. The order inverted entirely when considering the model which provided the worst fit (e.g. using Feature RL in a block that was better fit by Abstract RL): higher learning rates and greediness for Feature RL. These differences can be explained intuitively as follows. In Feature RL, exploration of the task state-space takes longer - in short blocks (best fit by the Abstract RL strategy) a higher learning rate is necessary for the Feature RL agent to make larger updates on states that are infrequently visited. Results also suggest that action selection tends to follow the same principles - more deterministic in blocks that are best fit by Abstract RL (i.e. large $\beta$ for shorter blocks).

We predicted that the use of abstractions should increase with learning progress. Initially, the brain should rely on the full representation of features, as captured by Feature RL, since participants are not familiar with the structure or strategies to learn the task efficiently. With more experience, they
should become proficient in constructing a reduced mapping where only relevant features are represented, as in Abstract RL. To test this hypothesis, we quantified the number of participants using a Feature RL or Abstract RL strategy in their first and last blocks. On their first block, most participants relied on Feature RL, while the pattern reversed in the last block (two-sided sign test, $z = -2.77$, $p = 0.006$, Figure 4.3F). Computing the abstraction level separately for the session median split of early and late blocks also resulted in higher abstraction in late blocks (two-sided sign test, $z = -2.94$, $p = 0.003$, Figure 4.3G). These effects were complemented by a block-by-block analysis, displaying an increase in abstraction from early to late blocks (Figure 4.3H).

Supporting the current modelling framework, the mean expected value of the chosen action was higher for Abstract RL, and model hyperparameters could be recovered in the presence of noise [185]. Given the lower learning speed in excluded participants, the distribution of strategies was also different among them, with a higher ratio of Feature RL blocks. Please see [230] for more details on this analysis.

4.4.4 The role of vmPFC in constructing goal-dependent value from sensory features

The computational approach confirmed that participants relied on both a low-level feature strategy and a more sophisticated abstract strategy (i.e., Feature RL and Abstract RL; Figures 4.2D and 3B). Besides proving that abstract representations were generally associated with a higher expected value, the modelling approach further allowed us to explicitly classify blocks as belonging to either learning strategy. In this section, our objective was to dissociate neural signatures of these distinct learning strategies to show how abstract representations are constructed by the human brain.

First, we reasoned that an anticipatory value signal might emerge in the vmPFC at stimulus presentation [235]. We performed a general linear model (GLM) analysis of neuroimaging data with regressors for ‘High-value’ and ‘Low-value’ trials, selected by the block-level best fitting algorithm (Feature RL or Abstract RL, while controlling for other confounding factors such as time and strategy itself; see Materials and methods). As predicted, activity in the vmPFC strongly correlated with value magnitude (Figure 4.4A). That is, the vmPFC indexed the anticipated value constructed from Pacman features at stimulus presentation time. We used this signal to functionally define, for ensuing analyses, the subregion of the vmPFC that was maximally related to task computations about value when Pacman visual features were integrated. Concurrently, activity in insular and dorsal prefrontal cortices increased under conditions of low expected value. This pattern of activity is consistent with previous studies on error monitoring and processing [236, 237](Figure 4.4D).

For the vmPFC to construct goal-dependent value signals, it should receive relevant feature information from other brain areas and specifically from visual cortices, given the nature of our task. Thus, we computed a psychophysiological interaction (PPI) analysis [238], to isolate regions in which functional coupling with the vmPFC at the time of stimulus presentation was dependent on the magnitude of the expected value. Supporting the idea that the vmPFC based its predictions on the integration of visual features, only connectivity between the visual cortex (VC) and vmPFC was higher on trials that carried large expected values, compared to low-value trials (Figure 4.4B). Strikingly, the strength of this VC - vmPFC interaction
Figure 4.4: Neural substrates of value construction during learning. (A) Correlates of anticipated value at Pacman stimulus presentation time. Trials were labelled according to a median split of the expected value for the chosen action, as computed by the best fitting model, Feature RL or Abstract RL, at the block level. Mass univariate analysis, contrast ‘High-value’ > ‘Low-value’. vmPFC peaks at [2 50 -10]. The statistical parametric map was z-transformed and plotted at \( p_{\text{FWE}} < 0.05 \). (B) Psychophysiological interaction, using as seed a sphere (radius = 6 mm) centred around the participant-specific peak voxel, constrained within a 25 mm sphere centred around the group-level peak coordinate from the contrast in (A). The statistical parametric map was z-transformed and plotted at \( p_{\text{fpr}} < 0.001 \) (one-sided, for positive contrast - increased coupling). (C) The strength of the interaction between the vmPFC and VC was positively correlated with the participant’s ability to learn block rules. Dots represent individual participant data points, and the line is the regression fit. The experiment was conducted once (n = 33 biologically independent samples). (D) Neural correlates of (predicted) low value at visual stimulus presentation time. Trials were labelled according to a median split of the expected value for the chosen option as computed by the best-fitting model, at the participants and block level. The statistical parametric map was z-transformed, and false-positive means of cluster formation (fpr) correction was applied. \( p_{\text{fpr}} < 0.001, Z > 3.09 \). * \( p < 0.05 \). Adapted from [230], eLife Sciences Publications.

was associated with the overall learning speed among participants (N = 31, robust regression, slope = 0.016, \( t_{29} = 2.55, p = 0.016 \), Figure 4.4C), such that participants with stronger modulation of the coupling between the vmPFC and VC also learned block rules faster. The strength of the vmPFC - VC coupling showed a non-significant trend with the level of abstraction (N = 31, robust regression, slope = 0.013, \( t_{29} = 1.56, p = 0.065 \) one-sided). However, this study was not optimised to detect between-subject correlations that normally require a larger number of subjects. Therefore, future work is required to confirm or falsify this result.

4.4.5 A value-sensitive vmPFC subregion prioritises abstract elements

Having established that the vmPFC computes a goal-dependent value signal, we evaluated whether the activity level of this region was sensitive to the strategies that participants used. To do so, we used the same GLM introduced earlier, and estimated two new statistical maps from the regressors ‘Abstract RL’ and ‘Feature RL’, while controlling for idiosyncratic features of the task,
Figure 4.5: Neural substrate of abstraction. (A) Regions of interest for univariate and multivariate analyses. The HPC was defined through automated anatomical labelling (FreeSurfer). The vmPFC was functionally defined as the cluster of voxels found with the orthogonal contrast ‘High value’ > ‘Low value’, at P(unc) < 0.0001. (B) ROI activity levels corresponding to each learning mode were extracted from the contrasts ‘Feature RL’ > ‘Abstract RL’, and ‘Abstract RL’ > ‘Feature RL’. Coloured bars represent the mean, and error bars the SEM. (C) Multivariate (decoding) analysis in three regions of interest: VC, HPC, vmPFC. Binary decoding was performed for each feature (e.g. colour: red vs green), by using trials from blocks labelled as Feature RL or Abstract RL. Colour bars represent the mean, error bars the SEM, and grey dots represent individual data points (for each individual, taken as the average across all three classifications, i.e., of all features). Results were obtained from leave-one-run-out cross-validation. The experiment was conducted once (n = 33 biologically independent samples). (D) Classification was performed for each feature pair (e.g. colour: red vs green), separately for blocks in which the feature in question was relevant or irrelevant to the block’s rule. The statistical map represents the strength of the reduction in accuracy between trials in which the feature was relevant compared to irrelevant, averaged over all features and participants. (E) Classification of the rule (2x2 blocks only). For each participant, classification was performed as fruit 1 vs fruit 2. In (D–E), statistical parametric maps were z-transformed, false-positive means of cluster formation (fpr) correction was applied. p(fpr) < 0.01, Z > 2.33. * p<0.05, ** p<0.01. Adapted from [230], eLife Sciences Publications.
of ‘ROI’ ($t_{128} = 2.16, p = 0.033$), and ‘strategy’ ($t_{128} = 3.07, p = 0.003$), and a significant interaction ($t_{128} = -2.29, p = 0.024$), illustrating different HPC and vmPFC recruitment (Figure 4.5B). Post-hoc comparisons showed vmPFC activity levels distinguished Feature RL and Abstract RL cases well (LMEM: $t_{64} = 2.94, p_{(FDR)} = 0.009$), while the HPC remained agnostic to the style of learning (LMEM: $t_{64} = 0.62, p_{(FDR)} = 0.54$). Alternative explanations are unlikely, as there was no effect in terms of both the correlation between value-type trials and algorithms, and task difficulty, measured by reaction times.

The next question we asked was, ‘Can we retrieve feature information from HPC and vmPFC activity patterns?’ To abstract and operate in the latent space, an agent is still bound to represent and use the features, because the rules are dictated by feature combinations. One possibility is that feature information is represented solely in sensory areas. What matters then is the connection with and/or the read-out of vmPFC or HPC. Accordingly, neither HPC nor vmPFC should represent feature information, regardless of the strategy used. Alternatively, feature-level information could also be represented in higher cortical regions under Abstract RL to explicitly support (value-based) relational computations [239]. To resolve this question, we applied multivoxel pattern analysis to classify basic feature information (e.g. colour: red vs green) in three regions of interest: the VC, HPC, and vmPFC, separately for trials labelled as Feature RL or Abstract RL. We found that classification accuracy was significantly higher in Abstract RL trials compared with Feature RL trials in both the HPC and vmPFC (two-sided t-test, HPC: $t_{32} = -2.37, p_{(FDR)} < 0.036$, vmPFC: $t_{32} = -2.51, p_{(FDR)} = 0.036$, Figure 4.5C), while the difference was of opposite sign in VC ($t_{32} = 1.61, p_{(FDR)} = 0.12$, Figure 4.5C). The increased feature decodability in Abstract RL was significantly larger in the HPC and vmPFC compared to the VC (LMEM model ‘$y \sim ROI + (1 \mid participants)$’, $y$: difference in decodability, $t_{97} = 3.37, p = 0.001$). Due to the nature of the task, the number of trials in each category could vary and thus confounds the analysis. A control analysis equating the number of training trials for each feature and condition replicated the original finding. These empirical results support the second hypothesis. In Abstract RL, features are represented in the neural circuitry incorporating the HPC and vmPFC, beyond a simple read-out of sensory cortices. In Feature RL, representing feature-level information in sensory cortices alone should suffice because each visual pattern mapped to a task-state.

We expanded on this idea with two searchlight multivoxel pattern analyses. In short, we inquired which brain regions are sensitive to feature relevance, and whether we could recover representations of the latent rule itself (the fruit preference). Besides the occipital cortex, significant reduction in decoding accuracy was also detected in the OFC, ACC, vmPFC and dorsolateral PFC when a feature was irrelevant to the rule, compared to when it was relevant (Figure 4.5D). Multivoxel patterns in the dorsolateral PFC and lateral OFC further predicted fruit class (Figure 4.5E).

4.5 DISCUSSION

To accomplish goals, it is fundamental for agents to generate representations of the environment that reduce its complexity and highlight relevant aspects of it. The construction of abstractions, i.e., the identification of simplified representations of the tasks, is an integral part of this process, crucial to allow complex behaviours to appear [240, 201, 241]. In the present study,
we show how humans can effectively generate abstractions from sensory information, employing a strategy that reduces dimensionality according to task demands. We used computational models to track the underlying changes in the value predictions, which operate as a guiding factor in the learning process. We found that participants gained experience during the task which helped them to be more proficient in generating the abstractions. Plus, we found that participants that relied more on the abstract strategy (Abstract RL) had better performance and higher confidence in their actions. This is indicative of the construction of an efficient representation of the task.

The construction of summary representations of the environment in higher-order areas has been supported by many studies [135, 215, 210]. We further endorse this view by pointing out the functional connection between frontal areas involved in value computation, such as vmPFC, and sensory regions relevant to the task. Indeed, we also found that the learning speed was correlated with the coupling between vmPFC and VC. It has been shown that in complex environments, learning the reward of single features is a useful heuristic to facilitate value learning [242]. We complement this finding by showing that irrelevant features of the environment are discarded while the important ones are incorporated into abstractions.

Previous work has highlighted the important role of HPC in the formation and update of conceptual information [212, 209, 210, 86]. While the role of the HPC is to store, index, and update conceptual/schematic memories [210, 243, 211] the ‘creation’ of new concepts or schemas may occur elsewhere. Medial PFC or the vmPFC in humans [244, 243] could play this creative role, which is confirmed by recent findings showing how value signals are modulated by cognitive requirements and goals in the vmPFC [99]. This study suggests that the influence of vmPFC in the valuation of relevant low-level features in VC eventually feeds higher-order areas such as HPC, helping to create these goal-relevant representations. Supporting this view is the finding of higher univariate activation of vmPFC during Abstract RL blocks, while a similar activation of HPC was found in Abstract and Feature RL blocks, i.e., HPC engagement seems constant across feature and abstraction-based strategies. The increase of HPC feature decodability during Abstract RL could also be a sign of the prefrontal “tagging” of relevant features feeding up the higher-order representations. We also found that the decodability of patterns in OFC and DLPFC was enhanced for the goal-relevant features and also for the specific latent rule, supporting the role of pre-frontal regions in the construction and tracking of goal-relevant task states [245, 246, 247, 215, 248].

It is not yet clear if it is feasible for the brain to keep various potential representations in parallel or if a simpler “hypothesis-testing” regime favours the current best model for learning. Our mixture-of-experts RL is a simple algorithm to arbitrate between parallel models, with a trial-to-trial update that depends on the estimated responsibility of each expert. If it was computationally efficient for the brain to keep several models, it would be very convenient from the data perspective to give multiple uses to each one of the samples, like in the mixture-of-experts model. There is circumstantial evidence of multiple strategies being computed in parallel but deployed one at a time [249, 250]. Furthermore, in many scenarios, the brain may have access to only limited data points, while parallel processing is a major feature of neural circuits [251, 252, 253]. Further analysis of our experiment shows that a model comparison between mixture-of-experts and “simple” Feature-RL or Abstract-RL models favours the more parsimonious models.
(see [230] for details), shedding doubts on parallel processing. However, our study may not be ideally designed to capture this given the relatively straightforward and well-defined rules and reduced number of features that encourage a simpler representation. Further studies will need to explore the strategies implemented by the brain during learning, especially in scenarios where goals are not explicit.

The success of RL models relies on the adequate characterisation of the task states, which allows learning through updates on their value. However, how those states are defined is one of the major questions not only in neuroscience but also in machine learning and artificial intelligence [8]. Our study does not directly tackle this question, but it approximates it by using various possible state combinations to define the agent’s actions (e.g., the mixture-of-experts model). However, our approach still includes simple stimuli, with a manageable number of features. In real life, the relevant features may be multiple, and states are far from being well-defined. Deep RL has been used as an alternative method that leverages the representational power of deep learning to organize the features of the task in low dimensional space [254], although the obscure way deep networks create these representations keeps this approach open for further research.

It could be argued that the generation of abstractions as described by our RL model is a mere deployment of attention. If this was the case, we suggest that an increased feature decodability should have been found in the VC during the Abstract RL blocks [255, 256], which was not true. We can further dismiss the attention interpretation based on the results of an additional neurofeedback experiment conducted by our group (see [230] for details). In fMRI neurofeedback, it is possible to train participants to generate specific brain activity via the presentation of cues that reflect the real-time variation of patterns in the brain. For example, a disc of varying size or a thermometer could act as feedback indicating how close the participant’s brain activity is to the desired target pattern [257]. Participants are instructed to “raise” the feedback signal to obtain a monetary reward, which they will try to achieve using various mental strategies [258]. Crucially, human participants can learn to control their brain activity even when they do not have explicit knowledge of the brain area or potential function (e.g., motor or visual areas) of the intended patterns. In this study, participants were trained to generate a target activity in the VC associated with one of the task features, e.g., stripe orientation. This added value to some of the task features. We found that in a learning test after neurofeedback training, participants increased their abstraction level on the blocks where the neurofeedback target feature was relevant for the rule, while no change was found for the blocks where the trained feature was irrelevant. This means that tagging value to certain low-level features facilitates abstraction, a process that is likely to occur early on in visual information processing. Neurofeedback training is constrained to only a VC ROI and is experienced by participants in an unconscious way, discarding the possibility of a mere top-down effect. The manipulation indicates that value tagging of early representation has a causal effect on abstraction and consequently on the learning strategy. That is not to say that attention does not significantly mediate this type of abstract learning; however, attention could most likely be an effector of the abstraction and valuation processes [259]. Indeed, the decoding analysis shows stronger representations of relevant over irrelevant features in the VC (Figure 4.5D). The previous chapter in this thesis (Chapter 3) further
complements the dynamic role that visual attention has in the selection of goal-relevant information.

Overall, this study presents how goal-oriented processes guide the learning of representations to characterise the task in a parsimonious way. Further experience in the task helps participants to create better abstractions, thus impacting performance and internal confidence. These findings illustrate value as having a function beyond the standard role described in decision-making and neuroeconomics. The role of vmPFC valuation seems to drive feature selection to create abstractions in coordination with sensory cortices. Furthermore, our findings tie reports from the decision-making and memory literature, in which vmPFC is attributed to a central position, for reward value computation and schema formation, respectively. The value signal could be more than just a proxy for reward, it could encompass a dynamic representation of the goal-specific demands, acting as a compass in the development of the appropriate learning strategies.
5 GOAL AFFECTS BRAIN REPRESENTATIONS IN PREFERENCE EVALUATION

5.1 SUMMARY

Many decisions rely on extracting perceptual information to identify the most advantageous options, e.g., in the previous chapter, visual features like colour or orientation in a Pacman character cued the most valuable alternative. However, in most value-based decisions, the source of information may not be straightforward, since we can rely on personal memories and experiences to construct the value of items. Assessing our personal preferences involves the creation of representations containing diverse features of the object. The standard view in value-based decisions is that we represent options according to how rewarding they are. A more general perspective is that subjective representations depend on the goals the agent must fulfil. That is to say, object representations should be constrained primarily by goals, rather than depending mainly on visual or hedonic reward features that are invariant to context. To test these views, we designed an fMRI study where participants assessed the positive and negative aspects of pets while looking at naturalistic images of them. Participants completed the tasks in two different frames to decouple goals and rewards. Multivariate analyses show that while representations in the visual ventral stream capture a goal-independent representation of pet identity, activity in prefrontal areas, such as vmPFC or OFC, encoded goal-relevant information. Additionally, hippocampal patterns showed higher similarity between images aligned with the goal of the task frame. Overall, these results align with the construction of brain representations of subjective preferences serving agents’ behavioural demands.

5.2 INTRODUCTION

Imagine that you are planning to get your first pet. You would probably do some research online on the pros and cons of different animals. You will think how fun it will be to play with puppies, observe the personality of cats or remember how beautiful was the parrot you saw during your holiday in Brazil. This may help you to have an idea of the options before making the decision. Unlike perceptual choices that rely on external information, in value-based decisions internal representations are especially important since they will be the main drivers of choice. The current understanding is that internal representations are formed thanks to complex mechanisms involving perception, attention and memory [32, 260, 52, 51].

The study of visual processing in neuroscience has revealed some mechanisms behind the construction of internal representations [261, 161, 262, 263]. The ventral visual stream, starting in the occipital cortex and reaching the
anterior parts of the temporal cortex contributes to the construction of complex and invariant representations using simple retinal input in the posterior areas of the brain [262]. This line of research has mostly been informed by experimental designs where human and non-human animals passively observe varieties of images (from simple objects to celebrities) in decontextualised scenarios, detached from further behavioural demands or goals [264, 265, 266, 267, 161]. In real-life scenarios, visual processing actively interacts with other cognitive operations, such as emotion, attention or motor functions [268, 269, 270].

Studies on the neural mechanisms and networks relevant for value-based choices have been focused on finding representations of reward, revealing the importance of prefrontal areas such as ventromedial prefrontal cortex (vmPFC) and orbitofrontal cortex (OFC) [32, 271, 272]. These experiments typically instruct participants to choose the options they prefer the most, i.e., the alternatives with higher hedonic reward [33, 66, 273, 274]. This approach has overlooked the possibility that these value representations could be coding more than the reward but the ability of the option to fulfil the person’s goals, i.e., choices are not just as an indicator of subjective preferences. Targeting this gap, a new wave of studies has expanded value-based decisions to probe the role of these preferences at the service of the behavioural demands [98, 275, 84]. We have shown that attentional profiles and evidence accumulation in value-based decisions change to satisfy goals, even when the alternatives are identical (Chapter 3 in this thesis and [186]). Areas in the brain’s value network also track goal-congruent values in addition to pure reward during multi-alternative choices [98, 99]. Outside value-based decisions, it has been shown that goals and contextual information play an important role in stabilizing memory encoding and acting as an organizing principle in hierarchical memory structures [91, 86, 276, 95]. Memory processes could be fundamental for value-based decisions since the evidence employed to construct the representations is likely to be extracted from previous experiences and impressions, and not merely from “current” perceptual inputs [51, 53].

These findings suggest that the way the brain represents objects in value choice is flexible and it is reorganised depending on the needs of the task. For example, an owner of multiple pets in financial strain may not be able to maintain all her animals, having to choose one of them to give away. In this scenario, the way she assesses the value of the animals may change relative to what would happen in a normal situation. During this crisis, she may think or remember the negative aspects, like the times her cat woke her up at night sitting on her face or how aloof the hamster was to her friendly gestures. Therefore, while the animals are the same as before (from the perspective of a vision scientist, maybe identical), the shift in the goals will make her appreciate them in a completely different light.

In the present fMRI imaging study, our objective was to study how goal shapes brain representations in the assessment of individual alternatives. Participants in this experiment were instructed to think of themselves in the scenario of searching for a new pet between four options: dog, cat, parrot or hamster. Similarly to the study presented in Chapter 3 (and [98, 186]), we implemented a frame manipulation to decouple reward from goals by asking participants to perform the tasks in two scenarios: like and dislike. In an imagination task, participants were asked to just observe naturalistic images of animals, while they thought of the positive and negative aspects of selecting that pet, in the like and dislike frames, respectively. We presented various
images of different animals, performing different actions and from different angles, pointing to elicit representations that were not stereotypical or dependent on the specific features of a single visual stimulus. Based on the analysis of brain pattern similarities [164], we found that representations in the visual ventral stream encoded animal identities, and potentially included frame information as well. We found the vmPFC and OFC contained goal-relevant value representations during the appraisal of the options. Hippocampus activity also changed its representations according to the behavioural demands, generating stable representation in images aligned with the objective of the task. The experiment also included simple choice trials to check that the participants were consistent in their preferences in both frames: in like trials participants chose the animal they preferred and in dislike trials the animal they wanted to reject. These binary decisions were based on auditory stimuli, to avoid contamination of the visual patterns from the imagination stage. We found that the shift in the relevance of the sensory modality for the task, i.e., an attentional displacement between visual and auditory information, impacted the representations in the specific sensory cortex. Overall, this work shows that object representation is not constructed in isolation, purely dependent on perceptual features, but it is shaped and organised according to one’s own goals.

5.3 METHODS

5.3.1 Experimental paradigm

Participants were presented with four animals (cat, dog, hamster and parrot) in two types of trials inside the MR scanner: imagination and choice. These two types of trials were also presented in two frames: like and dislike (Figure 1A).

In the imagination condition, participants were presented with a variety of photos of different individual animals (e.g., different breeds of dogs), doing various activities (e.g., sleeping, eating), and taken from different angles. A total of 8 photos per animal were presented during the training block, and 24 photos per animal were presented during the experimental blocks. During imagination trials, participants were presented with one picture at a time and were instructed to think of positive aspects (in the like frame) or negative aspects (in the dislike frame) of having that animal as a pet. Importantly, they were instructed not to focus on the specific animal presented in the current trial picture, but to use it as a cue to think about the whole animal category (i.e., they were not asked to think if they liked the specific dog in the presented picture, but if they liked the idea of having a dog). The images were presented with a blue or red frame and the words LIKE or DISLIKE on it. A beeping sound indicated the beginning of an imagination trial with a high (800 Hz) or low pitch (400 Hz) sound depending on the frame. The correspondence of blue-red/high-low pitch tones and like-dislike frames was inverted for half of the participants. On 25% of the imagination trials, participants were also asked to rate on a 1-4 scale (with 4 being the highest value) how much they liked or dislike the animal, depending on the frame. The objective of the rating was to check the participant’s attention and consistency. In the like frame, an animal would be rated 4 if participants had a strong preference for it; whereas in the dislike frame an animal would be rated 4 if they strongly prefer to avoid having it
as a pet. The beginning of rating time was announced by a beeping sound (400 Hz).

In imagination trials the images were presented for 4.5 s and, in case rating was required, participants had extra 3 s to report their rating level using a 4-option button-box, and one second to show feedback of the chosen rating or a message indicating the missed trial. The inter-trial time after imagination trials was 2 s ± 1 s. A total of 192 imagination trials (96 per frame) were presented in the experiment.

During choice trials, participants heard two animal names played in stereo through MR-compatible in-ear headphones. One name was played to the right ear only and the other one to the left ear only, with a delay of 0.5 s of difference between them to avoid interference. Participants selected the animal they preferred to have as a pet in the like frame, and the animal they would prefer to avoid having as a pet in the dislike frame. High or low-pitch sounds, identical to the ones used in imagination trials, were used to indicate the start of a choice trial and its respective frame. The symbol of a speaker was presented in red or blue colour during the trial to indicate the respective frame of the decision, keeping the same frame colour coding employed during imagination trials. Participants were allowed to choose a left or right option after they listened to both sounds, which was indicated by two white circles appearing on the screen. After participants chose an option by pressing a button, the circle corresponding to the side of the selected alternative was coloured light green. Participants had 2.5 s to listen to both alternatives and after that, they had 3.5 s to respond. An extra 1 s were used to highlight the selected option or to show a “too slow” message in case the choice was missed. The inter-trial time after choice trials was 3 s ± 1 s. A total of 144 choice trials (92 per frame) were presented in the experiment.

Participants read the instructions and completed a training block before starting the experiment in the MR scanner. The MR experiment was separated into 6 blocks with 32 imagination trials and 24 choice trials each. The presentation of trials was pseudorandomised to maintain a similar number of trial types across blocks. Within blocks, imagination and choice, like and dislike trials were fully randomised. The full experimental session inside the MR scanner lasted around 1.5 hours. The images of the four animals were selected from various online resources. The sound files with animals’ names were generated using a text-to-sound generator available online (soundof-text.com, UK English selected). The transformation to dichotic stimuli was done using custom MATLAB code. The full MR experiment was coded using the Python toolbox PsychoPy [277].

An additional online questionnaire was filled out by the participants after they completed the experimental session. Participants indicated how much they liked each one of the individual animal images (not the animal category) presented during the experiment. They also reported their familiarity and ranked their preference for the four animals. The questionnaire was implemented using the Gorilla platform (https://gorilla.sc/).

5.3.2 Participants

We recruited 36 healthy volunteers (age: 27.94 ± 4.36 years; 18 females). Three participants were left-handed. All participants were fluent in English and had a normal or corrected-to-normal vision. We excluded 2 participants because of excessive head motion during the scanning; 1 participant because of artefacts in the brain images (in the prefrontal area); and 4 participants
because their behaviour was inconsistent with the changes in frames (e.g., identical rating for like and dislike frames). Therefore, a total of 29 participants were included in the final analysis. The study was conducted in accordance with the Declaration of Helsinki and was approved by the Research Ethics Committee of the University College London. Before starting the experiment, all participants gave written consent. After the experiment, they were compensated £30 for their participation.

5.3.3 fMRI

Brain imaging data were acquired at the Wellcome Trust Centre for Human Neuroimaging using a Siemens (Erlangen, Germany) Prisma 3.0-T MRI scanner with a 64-channel head coil. We measured BOLD activity over six experimental blocks with a T2*-weighted multiband EPI sequence (voxel size = 2 x 2 x 2 mm; matrix size = 106 x 106; repetition time (TR) = 1.450 s; echo time (TE) = 35 ms; acceleration factor = 4, flip angle = 70°) with 72 slices parallel to AC-PC. Each block contained around 400 volumes (the variability given the jitter in trial onsets), resulting in a total block duration of ≈9.6 min. We discarded the first five volumes of each session, during which no stimuli were presented, to allow for the stabilization of the magnetic field. An anatomical T1-weighted MPRAGE scan (1.0-mm isotropic) and field maps were also acquired for each subject.

The pre-processing of brain images was completed using the MATLAB toolbox Statistical Parametric Mapping (SPM12; www.fil.ion.ucl.ac.uk/spm) and custom code. Raw volumes were slice time-corrected, realigned, and then unwarped using the data from field maps. Multivariate analyses were performed using native space images. For the univariate analysis, each subject’s anatomical image was further segmented into grey and white matter, cerebrospinal fluid, bone, soft tissue, and air images. The same deformation map was then used to normalise the EPIs to the MNI template, which were lastly smoothed with a Gaussian kernel with full width at half maximum (FWHM) of 8 mm.

5.3.4 Univariate analysis

Univariate analyses were performed with SPM12 and custom code. In our main general linear model (GLM), the first regressor was an indicator function of the onsets of all the imagination trials in like and dislike frames. The second regressor was an indicator function of all the onsets of the second sound during choice trials, also in both frames. We selected the second sound since at this point participants had available the full information for their decisions. Regressors were not orthogonalized but were left to compete for variance. To extract auditory ROIs, we fitted an additional GLM including the onset of each sound as a regressor. In all GLMs, we also corrected motion artifacts by including six subject-specific parameters from image realignment (corresponding to three rigid-body translations and three rotations) as covariates of no interest. All regressors were then convolved with a canonical hemodynamic response function. A high-pass filter with a cutoff of 1/128 Hz was applied to the time series to remove slow drifts. Temporal autocorrelation was estimated with a first-order autoregressive model (AR-1). Clusters were detected with a cluster-defining threshold of $P < 0.001$ and FWE-corrected for multiple comparisons at $P < 0.05$. The
main GLM was constructed for representational similarity analysis (RSA) and fitted to unsmoothed and unwarped EPIs.

5.3.5 Representation Similarity Analysis (RSA)

Multivariate analysis was performed using RSA. This method allows us to explore how different brain regions represent the information in a task, observing the relationship between a hypothesis model and brain activity patterns [164, 168]. Hypotheses about the neural representations are specified in terms of Representational Dissimilarity Matrices (RDMs). An RDM is a 2D matrix where rows and columns correspond to the experimental conditions, with the elements representing the pairwise dissimilarity (or distance) under those conditions. In this way, conditions that are similarly represented will have low dissimilarity. RDMs can be constructed using activity patterns from participants’ brains during the experimental conditions or created as hypothetic models describing the representational structures. To construct the brain-based RDMs, the pairwise distance was computed using $1 - \rho$ between the brain patterns representative of the experimental conditions. In a second-level similarity analysis, a comparison between RDMs was possible, model and brain-based. This similarity estimation was done using $1 - \rho$ as a measure of distance. The second-level similarity allowed inferring to what extent the activation pattern in a specified ROI correlated with a model prediction.

In the present study, the RSA was carried out with a dedicated MATLAB toolbox [165] plus custom code. In our task, the brain patterns to estimate similarities were extracted from the univariate analysis. Here we defined a GLM using as regressors the onsets for all the imagination and choice trials (details above). Since the task included 192 imagination trials and 144 choice trials means we used 336 separated $\beta$-maps (for each individual participant), generated from the GLM. Unwarped and unsmoothed brain images were used to fit the GLM, with $\beta$-maps in the native subject space.

5.3.5.1 RSA Searchlight

The RSA searchlight analysis explored the entire brain looking for areas with a specific representational structure of the experimental conditions (captured by an RDM model). Searchlight was performed within grey matter masks obtained for each participant. In line with previous approaches [278], these masks were defined as a set of voxels with a probability of including grey matter exceeding 0.3 according to the tissue segmentation step. A spherical ROI with a radius of 15 mm was defined around each voxel in the grey matter mask. Brain-based RDMs were built in each one of these spherical ROIs. For each brain position, the model and brain-based RDMs were reshaped to a vector and compared using Spearman’s $\rho$ correlation coefficient, assigned to the centre of the sphere. We used the imagination condition to run the searchlight. For each subject, we calculated the average $\beta$-maps for each animal from all its exemplars, separated by frame. This produced a single representative pattern for each one of the four animals in like and dislike (8 conditions in total, 4 for each frame). Our main RDM model for searchlight characterised pet identity, meaning that the pairwise elements for the same animal in the matrix were assigned 0 (i.e., high similarity) and all the other elements were 1 (i.e., low similarity) (Figure 5.3A). By comparing this model RDM with the brain-based RDMs we generated a subjective correlation
5.3 METHODS

map (Spearman’s ρ), which encoded the correlation between the model and brain representation across the whole grey matter mask. The maps were lastly normalised to MNI space with the deformation map from the tissue segmentation for each participant and then smoothed with a Gaussian kernel with FWHM of 8 mm. Group-level statistics were performed with a non-parametric cluster-level permutation test implemented in SnPM (statistical nonparametric mapping) [279], with a cluster detection threshold at $P < 0.001$ and no variance smoothing. Clusters were considered significant if they survived an FWE correction at $P < 0.05$ [280].

We additionally performed a searchlight analysis using a model for goal-relevant valuation in imagination trials (Figure 5.5): for each participant we created an RDM using as distance the difference between their animal preferences, i.e., $d_{ij} = |\Delta \text{Ranking}_{ij}| = |R_j - R_i|$, with $R$ the preference ranking for each animal category (not the exemplar value). Importantly, for dislike frame the ranking was the opposite of for like frame, i.e., $R_{\text{like}} = 5 - R_{\text{dislike}}$. Based on previous findings [99] we constrained our search to 20-mm vmPFC ROI, applying a small-volume correction to determine the significant clusters.

5.3.5.2 ROI analysis

We performed a multivariate analysis in regions of the brain important for goal-oriented decisions. Binary masks were created based on ROIs from one of our previous studies on goal-oriented decision making [99]. In that study a multivariate analysis reported goal-dependent representations in the prefrontal cortex: vmPFC: [−10, 54, 2] and OFC: [-32, 38, -14], both relevant in the valuation circuit. From that same study, we also extracted a visual ROI located in Brodmann 37 (right fusiform area, [28, -54, -12]) which was the peak voxel resulting from a multivariate analysis on the perceptual representation of the task stimuli. Given the relevance of sound stimuli in our choice trials, we also created an ROI in the primary auditory cortex. This was extracted from a group-level univariate analysis considering the onset of choice sounds (pet names) (peak activity in BA41: [64, -24, 6]) (details above). We created 10mm spheres centred at the MNI coordinates for all these ROIs and then we warped them to each participant’s native image space. All ROIs were generated using MarsBar toolbox in MATLAB [281]. To extract bilateral hippocampus ROIs, cortical reconstruction and volumetric segmentation were performed for each participant using Freesurfer 7 image analysis suite (http://surfer.nmr.mgh.harvard.edu/). Using the processed images, we used the tool for segmentation of hippocampal subfields and nuclei of the amygdala, also available in Freesurfer [282]. We separated the hippocampus from the segmentation and custom code was used to transform the mask to the participant’s native space in SPM.

We performed an RSA analysis within these ROIs using both, imagination and choice conditions. For imagination, we defined the conditions as in the searchlight analysis, averaging patterns for the 4 animals in the like and dislike frames. For choice condition, we averaged across all the trials that presented the same pair of animals, independent of the side (e.g., all the trials where the choice between dog-cat or cat-dog was presented were averaged together) generating 12 activity patterns (6 for each frame). Given that the number of exemplar trials to average is higher in imagination trials than in choice trials (24 trials vs 12 trials), for the RSA analysis including both tasks simultaneously we subsampled 12 of the imagination trials and calculated the respective average. We constructed an RDM model to use for the ROI
analysis. This model assigned a high similarity within imagination and choice conditions, and low similarity across conditions ignoring the frame, e.g., two imagination trials are similar to each other, and dissimilar to choice trials (justTask model). We constructed other models but we did not find significant relationships with the RDM models: 1) a model that assigned high similarity within the same task and frames, but low similarity across tasks and frames (e.g., the high similarity between all the like imagination trials, and low similarity with dislike frame and choice task) (frameSplitTask model); 2) a model that assigned high similarity to pet identity, independent of the frame and conditions, considering for choice trials the pet identity of the presented animals (petId model); 3) similar to 2) but considering high similarity for conditions within the same frame (framPetId model), and 4) a random value RDM as control. In this report, we did not present the results for all these analyses, but they were mostly non-related with the brain ROIs. We compared the relatedness of the brain RDMs with the models RDMs using Kendall’s rank correlation coefficient $\tau_A$. Randomization tests, with FDR correction for multiple testing (FDR $p<0.05$) were used to infer a single model relatedness with the ROI RDM. Comparison of the difference of relatedness between models was performed using subjectRFXsignedrank (two-sided Wilcoxon signed-rank test) in the RSA toolbox [165] with FDR $p<0.05$ correction for multiple comparisons.

We also used the results from the searchlight analysis to create an ROI. In this case, only imagination trials were considered. With this we calculated a dendrogram and activity patterns MDS scaling (computed with metric stress criterion) using the default settings in the RSA toolbox [165].

In our RSA analysis described above we averaged the patterns of animal and frame categories. This was done mainly to facilitate the analysis given the big number of exemplars (96 images duplicated by the frames) in the imagination task, which made it unfeasible for us to run a full searchlight. As a second approach, to obtain a more detailed view of the effect of frames in the imagination stage, we estimated the similarity between imagination trials using the ROIs described in the previous section. The similarity was calculated using Pearson’s correlation $r$ considering independently the activity patterns for all the trials (photo exemplars), not just an average for each animal. The correlation values considering the trials where the same photo was presented were compared with the similarity between trials showing different photos of the same animal. Additionally, within trials presenting the same pet category (e.g., all the trials presenting dogs), we compared if trial patterns were more similar to other trials with the same frame (congruent trials) vs trials with a different frame (incongruent trials). Further, we separated the imagination trials into high and low exemplar values (i.e., the value preference of the photo, not the animal category), as obtained from the post-test questionnaires. We compared the internal similarity of high vs low-value trials, separately for like and dislike frames. The correlation coefficients between individual trials were averaged at a participant level. We used the Wilcoxon signed-rank test, a non-parametric statistical test that relaxes the normality assumption often violated by correlation coefficients, to compare the differences between samples.
5.4 RESULTS

5.4.1 Behavioural results

Participants were exposed to two frames: like and dislike. For each frame, participants were required to complete two types of trials: imagination and choice (Figure 5.1A). During imagination trials, they were shown naturalistic images of four pet categories: dog, cat, hamster or parrot. There were 8 pictures of different animals for each category; we will refer to each of them as an ‘exemplar’ hereafter. While observing an exemplar, participants were instructed to focus on the positive or negative aspects of having that animal as a pet, in the like or dislike frames, respectively. In choice trials, participants listened to the names of two animals, on the left and right speakers, and they were instructed to report the side (left or right) corresponding to the animal of their preference. In the like frame, they chose the animal they would prefer to have as a pet, and the animal they would prefer to avoid in the dislike frame.

We wanted to make sure that participants performed the tasks correctly during our experimental sessions, considering the changes in the frame. Therefore, in 25% of the imagination trials, we asked them to rate how much they liked or disliked the animals on a 1-4 scale. In the like frame, a high response (e.g., 4) should go to an animal that was much preferred; while in the dislike frame, a non-wanted animal should be the one to get a high response. Importantly, they were asked to rate their preference for the animal category, not the specific animal photo. The main objective of this rating was to measure participants’ engagement with the frames in the imagination task. We found participants responded consistently to the task goal, with ratings reported in the like frame inversely correlated to the rating reported in the dislike frame ($r = -0.622$, $p<0.001$) (Figure 5.1B). We also asked participants to rank the pets according to their preferences in an online questionnaire after the scanning session. There was a clear relationship between in-task and off-task preference reports ($r = 0.68$, $p<0.001$) (Figure 5.1D). In the online questionnaire participants additionally used a rating scale (1–100) to indicate their preference (i.e., how much they liked) each one of the exemplar images presented in the imagination trials (24 exemplars in total). Note that in this case, the report corresponded to the preference for the photo itself, not necessarily for the animal category. Still, the exemplar images of animals with higher preference tended to be reported as preferred (i.e., if dogs were preferred, dog photos were rated higher) ($r = 0.296$, $p <0.001$) (Figure 5.1F). The objective of using different photos during the imagination trials was to elicit a wide range of emotions or memories in participants, showing attractive but also undesirable photos of the animals. From figure 5.1F we can see that participants assigned a wide range of values to the photos, even in the preferred animal category. The variability in exemplar values was similar across animal ranking levels (one-way ANOVA on participants’ exemplar value standard deviation; $F(3, 25) = 0.868$, $p = 0.460$). Overall, dog and cat were the preferred animals for most of the participants, and parrot and hamster the least preferred (considering 1 the least preferred and 4 the most preferred animal, average rank across participants: dog = 3.5; cat = 2.89; hamster = 1.87; parrot = 1.75).

We also checked that participants correctly selected their most or least preferred animals according to the frame in choice trials. Accuracy on the choice was similar for like and dislike frames (accuracy$_{\text{like}}$: 0.909 ± 0.079, accuracy$_{\text{dislike}}$: 0.910 ± 0.079).
Chapter 5: Goal Affects Brain Representations

5.4.2 Representation of imagination and choice tasks

The behavioural data showed that participants were effectively tracking animal preferences and frame contingencies. Our main objective in this experiment was to study how the brain representations flexibly adapt to the experimental demands. For this purpose, we used RSA which characterises the representational content of brain activity [161, 164, 165, 99]. This method uses the similarity of neural patterns across various experimental conditions to infer the structure of representations in specific brain areas. For example, if a brain area represents the type of object that is presented in a task displaying images of houses and faces, then its activity patterns during "houses" trials should be more like other trials of the same type relative to "faces" trials. RSA uses matrices containing similarity information between the experimental conditions called Representational Dissimilarity Matrices (RDMs). Given the simile to distance, dissimilarity measures (i.e., 1 - Pearson’s r) are used to construct these matrices: the higher the value, the more "distant" in the representational space are the conditions. Since RDMs represent the similarity between all the combinations of the selected experimental conditions, it forms a square matrix (N x N) with the dimensions of the N experimental conditions.

While our experiment was focused on the assessment of the animal categories in different frames, the design per se involved a meta-category: the separation of the task in imagination and choice trials. We proceeded to investigate the representation of these two tasks in the brain. We extracted ROIs from a previous independent experiment on goal-relevant decision-making from our group [99]. From here we selected areas involved in a higher-level processing of goal-directed decisions: the ventromedial prefrontal cortex (vmPFC) and orbitofrontal cortex (OFC). Also, we extracted another ROI from that previous study related to visual processing, in particular, the area identified for the perceptual representation of object identity (ROI located in the fusiform area in [99], MNI coordinates: [28, -54, -12]). We will refer to this area as visual ROI in the following analyses. Additionally, given the relevance of the hippocampus in feature representations [283, 276], and evidence sampling for value assessment [51, 53], we performed a subject-based segmentation of this area (see Methods for details on the ROIs). In addition, given that choices were made on the pet names delivered through auditory stimuli, we included an auditory ROI. This was extracted from a univariate analysis based on a GLM using the onset of the pet names audio as regressors. We constructed the RDMs for all these areas including...
Figure 5.1: Task and behavioural analysis. (A) Design of the fMRI experiment: Imagination and choice trials were presented randomly in like and dislike frames. In imagination trials, participants were instructed to think about either positive or negative aspects of having that animal as a pet, in like and dislike frames respectively. In these trials, they observed a single exemplar (animal photo), with a red or blue frame indicating if the trial corresponded to a like or dislike frame (colours were counterbalanced between participants). A high and low pitch sound also indicated if the current trial was part of the like or dislike frame (also counterbalanced between participants). Each exemplar was presented twice, once for each frame. After the presentation of the exemplar, one-quarter of the trials requested the participant to insert a rating on a 1-4 scale to indicate the preference of the animal category (not photo exemplar). These ratings measured how much they liked or disliked the animal, depending on the frame. A distinct beep sound was presented when the rating scale appeared. In choice trials (bottom), participants made binary choices, selecting the animal they preferred (in the like frame) or the one they wanted to discard (in the dislike frame). Importantly, the options were presented using sound stimuli: the name of the animal was played on the left or right side using MR-compatible earplugs. The sounds were presented with a separation of 0.5 s to allow a correct recognition of both sounds. The frame for each trial was indicated using the colour of a speaker icon presented in the centre of the screen. Red or blue indicated the frame in the same way as in imagination trials. The same high and low pitch sound as in imagination trials was presented at the beginning of the choice trials to highlight the current frame. The selected option (right or left) was highlighted for a second before initiating the next trial. (B) Relationship between like and dislike ratings in imagination trials. The negative correlation between like and dislike ratings shows that participants were inverting their reports depending on the frame. In choice trials (bottom), participants made binary choices, selecting the animal they preferred (in the like frame) or the one they wanted to discard (in the dislike frame). Importantly, the options were presented using sound stimuli: the name of the animal was played on the left or right side using MR-compatible earplugs. The sounds were presented with a separation of 0.5 s to allow a correct recognition of both sounds. The frame for each trial was indicated using the colour of a speaker icon presented in the centre of the screen. Red or blue indicated the frame in the same way as in imagination trials. The same high and low pitch sound as in imagination trials was presented at the beginning of the choice trials to highlight the current frame. The selected option (right or left) was highlighted for a second before initiating the next trial. (B) Relationship between like and dislike ratings in imagination trials. The negative correlation between like and dislike ratings shows that participants were inverting their reports depending on the frame. (C) Logistic regression for choice trials in like and dislike frames. The plots show the probability of choosing the right option with respect to the difference in preference for each animal (\(\Delta \text{Ranking} = \text{rankright} - \text{rankleft}\)). Each line represents an individual participant, and blue and red lines indicate the group average. The change from positive to negative slope indicates that participants were correctly shifting their responses with the frames. (D) The animal ranking was extracted from an online questionnaire completed after the fMRI session. Each participant arranged the animals from highest to lowest preference. Subject-level rankings were correlated with the ratings given during the fMRI experiment, after the imagination trials. (E) Additionally, participants completed another questionnaire with their preference for each one of the exemplar images (not the animal as a category). We found that the preference for the photo was correlated with the preference of each one of the animals.

imagination and choice trials (Figure 5.2A). We found across these ROIs separate representations of imagination and choice tasks, with higher similarity within tasks relative to between tasks. This can be observed from the RDMs obtained for these areas (Figure 5.2A). To check this finding, we compared the ROIs with a model RDM that captured high similarity within imagination and choice tasks, and low similarity between tasks. We found significant relatedness between this model and all the selected ROIs (ROI
chapter 5: goal affects brain representations

RDM relatedness with model RDMs tested using randomization test, p<0.05, with multiple testing FDR adjustment). Note that brain and model RDM relatedness was calculated using rank correlations, to avoid assuming a linear relationship between the dissimilarities [165]. Therefore, the brain organises imagination and choice tasks separately, even when the trials were randomized to exclude potential block effects or temporal correlations [284].

An important feature of our experiment was that each task relied on a different stimulus type (visual in imagination and auditory stimuli in choice). We checked whether the representational structure in visual and auditory ROIs reflected this difference. As can be seen from the MDS plot (Figure 5.2B), we found that the internal similarity between the imagination trials was higher than between choice trials in the visual ROI (fusiform area). We confirmed this by comparing the dissimilarity extracted from the brain RDM showing significantly lower values (i.e., more similar patterns) in imagination relative to choice trials (F(2, 27) = 19.728, p < 0.001; Imagination vs Choice: FDR p<0.001) (Figure 5.3D). The opposite pattern was observed in the auditory ROI: choice trials’ patterns were closer to each other, in comparison with the internal relationship observed within the imagination trials (Figure 5.2C). This was also found in the dissimilarity comparison (F(2, 27) = 82.979, p < 0.001; Imagination vs Choice: FDR p < 0.001) (Figure 5.3E). These results show how behavioural demands cue the reorganisation of brain patterns, orienting the activity of the relevant sensorial cortices. Previous studies have reported higher pattern similarity as a result of the stabilization of brain activity product of visual attention [285, 286], and neurons in the auditory cortex can sharpen their spatial tunning following attentional adaptations useful to capture relevant features [287]. Additionally, neural gain has been associated with focused attention, higher functional connectivity and tightly clustered patterns across the brain networks [288].

5.4.3 Representational structure in imagination task

5.4.3.1 Animal category analysis

After describing how the brain captures the differences between the two tasks, in the following section we present in more detail the brain representations of the animal assessment during the imagination task. In addition to the ROI approach used in the previous section, RSA also allows us to look for brain areas that contain specific representational structures. This methodology is known as a volumetric searchlight. This requires defining a model RDM that contains the desired relationship structure between the experimental conditions and to correlate it with RDMs extracted from voxels sampled covering the entire brain. In particular, we used a moving sphere to estimate multiple RDMs and check which regions contained the desired representational structure.

Firstly, we searched for a representation of the pet alternatives in imagination trials, independent of the frames. We constructed a model RDM where each row of the matrix (or column since it is symmetrical) corresponds to one of the pets in one of the frames, e.g., column one corresponds to a dog in a like frame (Figure 5.3A). We used that model RDM in a searchlight analysis. Due to computational constraints, for each one of the selected experimental conditions we averaged the brain activity patterns related to all the exemplars of the same animal into a single pattern and used it to calculate the brain RDMs. We used a volumetric searchlight in which the
correlation between model- and brain-based RDMs was computed within a 15-mm-radius spherical region centred in each voxel of the participant-level grey matter mask. The obtained correlation coefficient (Spearman’s ρ) was then assigned to the central voxel, generating subject-level correlation maps. Group-level statistics were performed using a nonparametric permutation test: all the reported clusters were identified with a cluster detection criterion of $P < 0.001$ and FWE-corrected at $p < 0.05$. We found activity in a large cluster covering various bilateral occipital regions reaching structures in the ventral stream like the fusiform gyrus (peak voxel in MNI space: $[12, -82, 4]$; $t_{28} = 9.04$, $p < 0.001$) (Figure 5.3B). These areas are relevant to visual processing and have been reported to play a role in object recognition [262, 99]. As a reference, the visual ROI used in the previous analysis obtained from

---

**Figure 5.2**: Imagination and choice trials RSA in selected ROIs. (A) RDMs calculated for the selected ROIs. Imagination and choice conditions were included in the same matrix. Blue and red squares represent the conditions for like and dislike frames, respectively. The blue, red and grey areas represent the conditions inside the tasks, imagination and choice. The separation between conditions is observed for most of the ROIs, which can be seen as areas of high dissimilarity in the elements of the matrix corresponding to the choice and imagination relatedness. (B) MDS plot for the visual ROI (fusiform area). A separation between the choice and imagination conditions was found. The choice condition was characterised as the pair of animals presented, independent of the side (e.g., hamVpar, hamster vs parrot). The blue circle in the brain insets indicates the sphere ROI. (C) MDS plot for the auditory cortex. Red: like frame; blue: dislike frame. Par: parrot. Ham: hamster. (D) Distance between activity patterns extracted from RDM for choice and imagination tasks in the visual ROI. Bar plot presents dissimilarity measures calculated for the patterns representing the imagination conditions only (Within Imag), choice condition only (Within Cho) and including the dissimilarity between imagination and choice conditions. A higher similarity was found in imagination for visual ROI. (E) Distance between conditions in the auditory ROI. A higher similarity was found in the choice of auditory ROI. Dots represent the participant average of dissimilarity measures. Black bars represent the standard error of the mean (95% confidence interval). Like and dislike frames were included together to calculate the dissimilarity measures. *** $p < 0.001$. Brain image templates, copyright 1993–2004 Louis Collins, McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University.
[99] had a more anterior location in the ventral pathway than the peak voxel found in this searchlight.

Figure 5.3: Imagination searchlight RSA. Pet identity. (A) RDM model for animal identity used in the RSA searchlight. (B) Brain regions following the activation pattern indicated in the model as obtained from the searchlight procedure. Brain image templates, copyright 137 (C) 1993–2004 Louis Collins, McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University.

To further analyse this activity, we extracted a 10-mm sphere centred in the peak voxel of the searchlight and extracted the RDM for the brain activities in the ROI (Figure 5.4A). We found that this area not only represented animal identity but also frame information. ROI patterns similarity within frame conditions was higher than across frames (Wilcoxon signed-rank test, $z = -2.3029; p<0.05$) (Figure 5.4D). This can be also appreciated from the darker colours in the elements of the RDM corresponding to like-like or dislike-dislike animal similarities. A hierarchical binary cluster tree analysis represented in the dendrogram [289, 86, 215] also found a clear grouping of animal pattern representations by frame (Figure 5.4B). A multidimensional scaling (MDS) analysis also showed two clusters, one for like and another for dislike frame (Figure 5.4C). This result does not constitute double-dipping since we constructed the searchlight purely based on animal identity, agnostic to like and dislike frames. However, it is important to consider that in our task each frame was presented with different colours (e.g., blue and red for like and dislike frames, respectively). This means that the encoding of colour and not frame could be driving this specific similarity pattern. We keep this result tentative, although a frame coding could be supported by previous findings of areas involved in early visual processing affected by top-down goal-directed settings [230].

Since the imagination task also involved the evaluation of the animals (i.e., thinking about positive or negative features), we searched for areas representing the value relationships between the options across frames. We constructed another model RDM using the differences in subject-level animal rankings, a measure of the valuation of each animal. Importantly, in the dislike frame we used the opposite ranking to capture a goal-relevant valuation, e.g., the animal with the highest ranking (rank 4) became the lowest ranked option (rank 1) in the dislike frame. This means that each animal had different rankings depending on the frame, i.e., the distance between the representations of the same animal in the like and dislike frame was not zero. We run a searchlight analysis using this goal-relevant value RDM. In previous work from our group [99] we have performed a similar analysis, capturing in an RSA searchlight the usefulness of various everyday items to
Figure 5.4: Frame representations in the occipital cortex. (A) We constructed an ROI selecting the peak of the searchlight cluster on pet identity in imagination trials. The blue circle in the brain inset indicates the ROI. We see the high similarity for the elements of the matrix corresponding to like-like and dislike-dislike conditions. (B) The dendrogram on dissimilarity measures extracted from the ROI analysis shows the clustering of animal patterns within the two frames, like and dislike. (C) MDS plot presenting the dissimilarity as distances between the animal representations. (D) Dissimilarity within the frame was found significantly lower than between frames. Notation in the dendrogram and MDS plot: par: parrot; ham: hamster; L: like frame; D: dislike frame.

solve two distinct goals (i.e., light a fire or anchor a boat). In this work, frontal brain regions, in particular vmPFC, were found to represent goal-relevant usefulness. The role of vmPFC in value computation \[290, 291, 98\] and the deployment of goal-relevant associative networks (often termed schemas) \[208\] have been extensively reported, supporting these goal-dependent neural representations. We constrained our searchlight to a 20 mm. sphere around the vmPFC peak reported in \[99\] (MNI coordinates: \([-10, 54, 2]\)). We found significant similarity with the model RDM in the vmPFC (peak voxel in MNI space: \([-16, 60, -2]\); \(t_{28} = 3.48, p_{FWE-corr} < 0.05\), small volume corrected). We run a whole-brain searchlight but no significant clusters were identified. Our results are in line with previous reports of vmPFC encoding, not simply a hedonic reward (how much participants like an alternative), but also the relevance of the option within the current task demands (see also Chapter 4; \[98, 99, 230\])

5.4.3.2 Image exemplar analysis

In the previous analyses, we have shown task representations across our selected ROIs in imagination and choice trials. However, in the case of imagination trials, due to the averaging at the animal level, we lose an important part of the richness of our experimental design: the various exemplars of the animals presented in both frames. Although we instructed participants to think about the general category of animals, we still expected that the presentation of multiple animal exemplars elicited different features or aspects of the pets in the participant’s evaluation. Therefore, in the following section, we investigated the brain representations during the imagination condition considering each individual trial separately. For this analysis, we used the ROIs employed in the previous section using the individual activity patterns for each one of the trials (in both frames) of the imagination task. We focused in particular on the hippocampus and OFC as areas of interest in the representation of goal-relevant information, and visual areas for the stimuli representation. First, we checked if these areas displayed goal-independent representation of the visual stimuli, i.e., areas that encoded with a high similarity the presentation of the same photo exemplar relative
116 Chapter 5: Goal Affects Brain Representations

Figure 5.5: Imagination searchlight RSA. Goal-relevant value. (A) Using participant ranking of the animals we estimated the goal relevance of each option in the like frame (i.e., the preferred animal to have as a pet). We inverted the ranking in the negative frame to capture the shift in task relevance (e.g., the animal with lower preference in the like frame becomes the animal with the higher ranking in the dislike frame). (B) Using frame-dependent rankings, we estimated the dissimilarity ($d_{ij} = |\Delta \text{Ranking}_{ij}|$) between each animal in the separate frames. (C) A searchlight in a vmPFC mask found significant similarity with the goal-relevant value RDM.

Brain image templates, copyright 137 (C) 1993–2004 Louis Collins, McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University.

to other photos of the same animal. Note that in this experiment each photo was repeated twice (once in each frame). We found that only a visual ROI, located in the fusiform area (extracted from [99]), showed a high similarity in activity patterns when the same image vs a different image (of the same animal) was presented (Wilcoxon signed-rank test, Visual: $z = -3.75; p<0.001$) (Figure 5.6A). This is consistent with the result found in searchlight and not surprising regarding visual processing.

We then checked whether we could find goal-dependent representations in our selected ROIs from the analysis of the animal exemplars. Within the the same animal, we compared the activity patterns in the same frame (congruent exemplars) and across frames (incongruent exemplars). We expected that the similarity of animals represented in goal-dependent areas should be higher in the same frame, e.g., the activity pattern of a dog presented in the like frame must be closer in the representational space to another dog in the like frame than to a dog presented in the dislike condition. We found indications of this pattern of activity in the OFC (Wilcoxon signed-rank test, $z = -2.38; p<0.05$),
that showed a significantly higher pattern similarity between images of the same animal in the congruent frame relative to the incongruent frame (Figure 5.6B). We did not find a significant effect in the hippocampus (considering an ROI including both, left and right sides), however, a significant effect appeared for left hippocampus only (Wilcoxon signed-rank test, $z = -2.21$; $p<0.05$). We focused this analysis on areas known to potentially play a role in context processing (e.g., [86, 215]), however, the results should be taken with caution given the small size of the sample ($n = 29$).

**Figure 5.6:** Pattern similarity in imagination trials. (A) Comparison of trials with the same exemplar (i.e., the same photo) vs trials with the same animal, but a different photo. Only the visual ROI presented a significant similarity for the same exemplar relative to the different pictures of the animal. (B) Comparison of trials with the same frame (congruent) vs trials with different frames (incongruent), within the same animal. Hippocampus (HPC) and OFC presented higher similarity for the congruent relative to incongruent trials. Dots represent participant average similarity. Black bars represent the standard error of the mean (95% confidence interval). ***: $p<0.001$; **: $p<0.01$, *: $p<0.05$.

Previous studies have shown that brain representations flexibly adapt to behavioural demands [99, 98]. The hippocampus, traditionally associated with memory, has been found to have a central role in the construction of representations guided by context [292, 86, 95]. Furthermore, hippocampal involvement has been suggested for the construction of value, acknowledging that it is likely that the information used to construct this abstract value is fed from previous experiences [52, 51, 293, 53, 294]. Further findings have shown that hippocampus processing can be influenced by attentional shifts, focusing on the representation of relevant features for the demanded task [295, 286, 276, 296]. Our imagination task required participants to assess the value of animals as pets, in two contexts, like and dislike. This means that participants had to evaluate their alternatives, focusing on positive or negative aspects depending on the frame. We presented a variety of animal photos hoping to elicit in participants various memories or features of each animal. In the post-test questionnaire, we asked participants to report their preference ratings for each one of the images in isolation. We expected that these exemplar value
ratings could be a good proxy to characterise the type of memory/features that the photo evoked in the participants, e.g., a photo of a dirty dog or an aggressive cat with a low score by participants may be associated with negative aspects of having the animal as a pet. Therefore, we separated all the images by high and low exemplar values using a median split for each participant. Previous studies have reported that attention and successful memory retrieval stabilises hippocampal representations, which is translated into a higher similarity between brain patterns of the same condition \cite{297, 285, 286}. In our task, we hypothesised that high-value exemplars captured the relevant features for the like frame, while in the dislike frame, the low-value exemplars should elicit features that were worthy of attention. We calculated the pattern similarities between pairs of hippocampal patterns corresponding to high-high or low-low value exemplars separately in both frames (Figure 5.7A). We expected that patterns of the frame-relevant condition (i.e., high-high the similarity in like and low-low similarity in dislike frame) should be more similar to each other, as a consequence of the attentional effect driven by the goal. We found that the right hippocampus encoded a significant interaction between the difference in pattern similarity between high-high and low-low value exemplars ($\Delta$Similarity) and the frames (Wilcoxon signed-rank test, like: $z = 2.45; p<0.05$, dislike: $z = -2.37; p<0.05$, paired samples Wilcoxon signed-rank test, like vs dislike: $z = 2.58; p < 0.01$) (Figure 5.7B). Specifically, we found that in the like frame, the high-value exemplars were more similar to each other than the low-value exemplars; while in the dislike frame, the signature was inverted, with a higher similarity between low-value exemplars. In other words, more stable representations were found in the images that were coherent with the active frame, e.g., photos of cute puppies and cats generated similar representations in the like frame, while photos of ugly parrots or hamsters were more closely represented during the dislike frame. We repeated this same analysis considering instead the full hippocampus ROI and we also found a significant difference between like and dislike frames similarities (like vs dislike: $z = 2.36; p < 0.05$). Overall, this result supports the role that goal-directed attention has on hippocampal representations, which may be an important mechanism to help the construction of value encoding the usefulness to fulfil the objective of the task \cite{98, 99}.

### 5.4.4 Goal-relevant representations in choice task

The information to make value-based choices does not necessarily depend on the perceptual characteristics of the stimuli but it should be extracted from internal sources, potentially using memory to accumulate information on the suitability of the alternatives \cite{51}. Activity in the hippocampus \cite{53} and prefrontal cortex areas, such as OFC \cite{49} have been suggested to support neural reinstatement or the appearance of brain patterns associated with the decision alternatives during the choice deliberation period. To capture this possibility, we estimated the similarity between imagination patterns, which we used as a proxy for the memory and experiences of the participants with each one of the animals, and the binary choice trials. An additional advantage of having multiple images during the imagination trials was to capture pattern representations that did not rely on particular features of a single photo, but that could capture some generality of the animal category. We expected that during choice, those brain patterns would appear again, with a prevalence of activity associated with the chosen animal. A
Figure 5.7: Goal-directed pattern similarity in the hippocampus during imagination trials. (A) Each participant was asked to rate how much they liked each one of the photos presented during imagination trials, independent of task frames or animal categories. Using this information, we separated all the imaginings trials into two groups: 1) trials with high-value exemplars and 2) trials with low-value exemplars. These two groups were specific for each participant, depending on their preferences. For each frame, we separated the imagination trials in those two groups, depending on the value of the exemplar presented. We calculated the similarity of hippocampal activity patterns for pairs of imagination trials, extracting specifically the similarity between high-high and low-low value exemplar trials. (B) We estimated the difference in the similarities between high-high value exemplars and low-low value exemplars (∆Similarity). In the right hippocampus, during the like frame, the imagination trials with high-value exemplars were “closer” to each other (i.e., higher pattern similarity) than the trials with low-value exemplars. On the other hand, in the dislike frame, the similarity of low-low value exemplars presented higher similarity to each other, relative to the high-high value exemplars’ similarities. The brain inset shows the hippocampus ROI in blue. Dots represent participant average similarity. Black bars represent the standard error of the mean (95% confidence interval). **: p<0.01; *: p<0.05.

key aspect of our choice trials was that they relied on auditory stimuli, to avoid the confounding factor of visual stimuli presented across imagination and choice trials. This is because if we presented images in imagination and choice trials, the pattern similarities could be due to the deployment of online visual attention toward specific features of the animal (e.g., the bright colours of parrots) and not an actual memory reinstatement. We expected that the similarity between imagination and choice trials would be based on the actual recovery of the patterns relevant to the choice. Therefore, using the dissimilarity values obtained from the RDM (including imagination and choice trials) we estimated the similarity of brain activity for each animal imagination (at a participant level) with each choice trial pair (Figure 5.8A). Previous reports have indicated that neural activity associated with the chosen option is preferentially reinstated during the decision [49]. Therefore, we
estimated the dissimilarity of the chosen and unchosen animals with their respective imagination patterns. For this analysis, we joined the dissimilarities for both frames. We could not find any significant differences between the similarity of the choice trials and imagination templates associated with the chosen and unchosen animals. In the hippocampus, we only found a trend showing higher similarity for the chosen than the unchosen animal (two-sided Wilcoxon signed-rank test, like vs dislike: $z = 1.632, p = 0.10$) (Figure 5.8B). We additionally compared the dissimilarities between the animals shown in each binary choice trial and the animals not presented, but we did not find significant differences in hippocampus or OFC. We performed a similar analysis for chosen and unchosen options, this time considering each animal and frame (e.g., for a choice of dog or cat in the like frame, we compared with imagination templates only in the like frame vs templates in the dislike frame). We calculated if choice trial patterns were more similar to the imagination patterns of the chosen option in the respective frames (e.g., the choice of a dog in dislike, is more similar to the imagination pattern of a dog in dislike than to like frames). Again, we did not find significant differences. The lack of significant results may be due to the high variability of the patterns we elicited in imagination trials and the insufficient number of samples in imagination trials and choice. Furthermore, as presented above, the patterns elicited by visual and auditory modalities were quite distinct making even more difficult to identify a significant reinstatement.

Figure 5.8: Choice trials brain patterns reinstatement. (A) We used the RDM matrix containing imagination and choice trials for each participant. We extracted the dissimilarity value between the imagination templates and the animals presented in each binary choice combination. For example, for a binary choice of dog versus cat, we selected in the RDM the value of the elements in green (chosen) and yellow (unchosen) in the matrix corresponding to the dissimilarity of the choice patterns with the imagination patterns observed for chosen (dog) and unchosen (cat) animals, respectively. (B) Comparing the dissimilarity between chosen and unchosen animals and their respective imagination templates, we did not find a significant difference, although the hippocampal patterns for the chosen animals were more similar to the imagination activity than the unchosen animals. In the presented analysis we included both frames together. A posterior analysis performed on patterns separated by the frames did not report significant differences in pattern dissimilarity.
Assessing our personal preferences involves pondering the diverse features of an object. While perceptual information in the environment can inform the subjective evaluation (e.g., if an apple is rotten, we will not rate it highly) most of the information used to construct these representations likely comes from internal sources (e.g., our previous experience eating apples). The retrieval of internal information is a process far from understood and likely to be highly susceptible to the conditions in which it occurs [51]. In this study, we investigated how the brain represents information during the appraisal of options when exposed to contexts that change the goal of its evaluation. In our study, participants observed diverse naturalistic images of animals, while considering how good or bad would be to have them as pets (imagination trials). In some additional trials, participants made choices between animals only using auditory information (choice trials). We found that the brain was involved in tracking the change in the experimental context, using the perceptual information adaptively to satisfy the internal goals and preferences.

We found that in our imagination task, the visual cortex was encoding information regarding the identity of the animals or photos presented during the imagination task. This activity was spread across areas in the ventral stream traditionally related to item recognition. In standard visual perception tasks in neuroscience, stimuli (simpler or complex) are tracked passively by the experimental subjects, like monkeys observing big arrays of animate and inanimate objects [161]. In our study, we included a variable behavioural goal (imagining positive or negative aspects of the animals) that we expected could generate differences at the level of object representations, even when the images were identical. This is in line with results suggesting that goals more than retinotopic maps could be the organizing principles in visual processing areas [161, 262]. We found that frames seemed to be organising object representations in frame-dependent structures (Figure 5.4), which could hint at the effect of higher-level processes, and top-down control over the activity of sensory cortices [278, 298, 299, 230, 300]. However, this result should be taken with caution since the frame information in our experiment was cued using a colour coding (e.g., blue for like, red for dislike). Colour can be found along the visual dorsal stream, with representations of chromatic retinal stimulus in the early visual cortex and perceptual experienced colours in higher regions [301]. The location of frame colour at a higher level of the representation hierarchy could be a manifestation of the cognitive importance of the frame over other object features, i.e., animal identity (please check Figure 5.4B to see the dendrogram). This adaptation of sensory cortices would be appropriate for the specific demands we introduced in this task: representing the frames [230]. Therefore, although our analysis suggests an impact of goals on visual processing, we cannot assert this with more certainty given the confounds with colour coding. The motivation to include distinct colours to identify the contexts was to avoid participants from confusing the current frame, since we did not use a block design and the trial frames were fully randomized.

On the other hand, anterior regions of the brain encoded higher-level aspects of the assessment process, incorporating goal-relevant information. We found higher similarity of brain patterns between trials within the same frame in the OFC during the imagination trials. This is in line with reports of the OFC representing task states and context in decision-making [302, 215,
Additionally, we found that a model of goal-relevant value preferences was similar to the structure of the representations found in the vmPFC, specifically in an area reported in one of our previous studies [99]. These findings confirm that activity patterns in this area could represent beyond purely hedonic reward, they could capture value information in a context-aware manner, i.e., encoding the usefulness of the options in the vmPFC.

The hippocampus is important for the encoding of information in distinct attentional states [286, 276]. In those studies, the representation of the same stimuli (e.g., virtual rooms with paintings) changed depending on the objective of the task (e.g., to search for a similar type of art or identical spatial layout) and impacted eventual memory retrieval. In our experiment, each frame generated a different attentional state in participants, i.e., thinking in positive or negative features. While our study was not a memory task (not requiring the encoding of the stimuli for later retrieval), the involvement of the hippocampus in the construction of value representations has been hinted as an important piece in the subjective value decision process [51, 53]. This role of the hippocampus in goal-relevant valuation is supported by our results that indicate an enhanced similarity between the activity patterns that are relevant to the task. We found that photo exemplars with high preference values were closer to each other in the representation space for the like frame, while exemplars with low values were more similar to each other in the dislike frame. We considered that high-value exemplars (e.g., photos of charming puppies) could elicit positive memories or features in participants, which would be features that facilitate the valuation process in the like frame. On the other hand, low-value exemplars (e.g., images of dirty or violent cats) could foster participants’ thoughts of negative aspects of the animal, which was beneficial in the dislike frame. Therefore, the images presenting goal-relevant features (e.g., photos of playful kittens in like frame) may have generated less conflictive representations (relative to the task frame), which translated into more stable brain patterns (i.e., patterns more similar to each other). On the other hand, conflictive scenarios (e.g., a photo of a cute puppy in the dislike frame) could have produced more diffuse representations. Previous studies have highlighted the importance of the hippocampus in the integration of goal-relevant features [210]. Plus, the high similarity between patterns has been reported as a result of the stabilization of brain activity product of attention or successful memory retrieval [297, 285, 286].

Of particular interest is the interaction of prefrontal regions with the hippocampus in the memory encoding and retrieval of valued alternatives within specific contexts [286]. OFC and hippocampus have an important role in the implementation of cognitive maps intimately linked to adequate context-aware performance [241]. Furthermore, the vmPFC and its interaction with the hippocampus and posterior neocortex have been associated with the processing of memory schemas [305, 306]. We did not explore these interactions in this work, but it is a key point to be revisited in future analysis of this study.

Since we utilized distinct sensory modalities in our imagination and choice trials, we found an additional adaptation of brain processing to these task demands. While we reported high similarity in the visual cortex patterns during imagination trials, patterns in the auditory cortex were closer to each other during the choice tasks. Note that the imagination task was based on the visual presentation of images and the choice task required participants
to discriminate between sounds (animal names). This means that the sensory modality associated with the focus of attention generated higher pattern stability in its respective sensory cortex. This result is coherent with enhanced cognitive processing of the relevant type of stimuli [285, 286], only that this time it was captured across sensory modalities. Neurons in the auditory cortex can sharpen their spatial tuning by attentional adaptations to the task [287]. In a visual-semantic task, the neural gain has been associated with focused attention, generating tightly clustered patterns across networks [288], which implies a higher similarity of brain activity. Therefore, higher engagement of the visual cortex in imagination trials and auditory cortex in the choice trials may generate more stable activity patterns in those regions reflecting directed processing of information. A previous study in the visual modality reported that more consistent and reproducible patterns are related to conscious cognitive processing vs more scattered patterns in response to non-conscious stimuli [307, 308]. Our results complement this finding, showing that top-down attentional control can help drive brain activity to generate representations adapted to task demands.

One of the objectives of our study was to characterise the brain pattern reinstatement during choice. This means that activity associated with the options would appear again during the deliberation stage of the decision, with a predominance of the chosen option [49]. We expected to use imagination trials showing single animals as templates to test the reappearance of animal-related brain patterns during the binary choice. In the same way, we hypothesised that brain patterns associated with the frame representations should be similar across the imagination and choice tasks. Although we predicted that brain patterns between imagination and choice trials would be different because of the visual and auditory modalities, we anticipated we could find at least some level of similarity across experimental tasks. However, none of this could be confirmed in our results. One of the reasons for this could be the small sample that we had for choice and imagination tasks. We could have used a bigger sample to create more robust category templates, as in other visual studies where hundreds of images are presented to a low number of subjects (e.g., [161, 309]). Another potential reason could be that process that participants perform during imagination trials (e.g., thinking about positive aspects) was not representative of the mechanism employed during the choice task (e.g., deciding that a dog is preferred over a cat). Alternatively, choice trials may have been too simple and repetitive (since we had only four possible animals), not requiring major assessments at the moment of the decision, i.e., it was more like an automatic report of a brief list of preferences than an actual consideration of the features of both alternatives. It has been described that choice and appraisal may involve distinct processes, which in our case may contribute to the difficulty in tracking cross-task frame effects [310, 98].

Overall, our study offers some insights into how goals and context information affects brain processing during the subjective assessment of options. Although we could not support the full-exempt of our predictions, we still found some interesting effects that suggest that tasks and frames permeate the assessment process at different levels in the brain, modulating the selection of goal-relevant information to generate internal representations.
Part III

CONFIDENCE AND GOALS IN HUMAN DECISION
With every decision comes an internal measure that reflects the level of uncertainty in our choice. Humans and other animals can generate an assessment of the quality of their choice and outcomes. This capacity to introspect is not merely anecdotal since it is fundamental to guide future choices and changes of mind. The factors that are integrated into the confidence signal and the specific role that it plays in human behaviour is an important open question in cognitive neuroscience. In the third part of this dissertation, I studied the post-decision confidence stage. The study presented in Chapter 6 exemplifies the relevance of confidence in the decision process. Here I focus on confirmation bias or the tendency of agents to overweight information related to their previous choices. I found that confidence is instrumental in this bias, guiding the occurrence of changes of mind and the exploration of alternatives in future decisions. Indeed, when participants were more confident in their choice, they tended to keep exploring the information that aligned with their preliminary selection, instead of paying attention to evidence that could prove them wrong. This intriguing behaviour could have its root in subtle changes in the participant’s endogenous context: the objective during a revision stage is not only to be correct but to validate their previous decision. The standard studies in confidence have not considered scenarios where the goals change, which can help to understand the overall role of these metacognitive reports. In Chapter 7, I presented more directly the effect that goals have on the generation of confidence. Based on the experiment presented in Chapter 3, I showed that decision frame manipulation affected confidence reports. In particular, I presented that confidence was enhanced by the overall amount of information aligned with the goal of the task. A typical view of confidence is that it reflects the accuracy of choices and the difficulty of the decision independently of other contextual factors. The modulation by goal-coherent evidence is in line with novel studies that suggest a more nuanced view of the generation of confidence. In Chapter 8, I connected my findings with another phenomenon recently described in the literature on confidence and metacognition: the positive evidence bias. This bias reported in perceptual studies considers that confidence is unbalanced towards the evidence supporting the decision, disregarding the unchosen evidence. I showed in a series of experiments that this effect can be also found in the value-based decision and it depends on the goals of the agent. I found that this effect does not depend on the presence of evidence per se but it could be driven by the absence of evidence (i.e., a “negative” evidence bias) if the task requires it. I developed a computational model to capture the effect of goals on confidence, based on a framework of signal detection theory and goal-dependent asymmetry in the variance of state representations. Finally, I hint that this goal-relevant behaviour in confidence can be a consequence of processes occurring during the deliberation process, such as variation of attention and internal expectations.
ATTENTION AND CONFIDENCE IN CONFIRMATION BIAS

6.1 SUMMARY

No one likes to be wrong. Committing to a decision changes the way we appreciate the initial problem. For example, if we decided to move to London, it is unlikely we will try to look for reasons that New York is a better city. The choice event seems to change an endogenous context in the decision-maker, which affects the future integration of information. Confirmation bias is the phenomenon where people tend to underweight information incompatible with previous choices. In the present work, we argue that a similar bias exists in the way information is actively sought, as this signature of human choice could be rooted in the basic mechanism of the decision process. We investigate how choice influences information gathering using a perceptual choice task and find that participants sample more information from a previously chosen alternative. Furthermore, the higher the confidence in the initial choice, the more biased information sampling becomes. As a result, when faced with the possibility of revising an earlier decision, participants are more likely to stick with their original choice, even when incorrect. Critically, we show that agency controls this phenomenon. The effect disappears in a fixed sampling condition where the presentation of evidence is controlled by the experimenter, suggesting that how confirmatory evidence is acquired critically impacts the decision process. These results suggest active information acquisition and confidence in the decisions play an important role in the propagation of strongly held beliefs over time.

6.2 INTRODUCTION

As presented in previous chapters, how we explore our environment and sample information is intimately connected with our goals, e.g., attending the snack we would like to eat or reject; identifying the features of a video game character that will maximize the score. In these examples the goals have been set by external circumstances (i.e., the experimenter), however, subtle changes taking place in the agents’ internal context, can also have an impact on their decision process. A behavioural measure that reflects these internal processes is the sense of confidence that accompanies every choice. It is known that confidence plays an important role in the modulation of decisions [311, 58, 57, 312, 313]. Reducing uncertainty is central to making efficient choices, so this should guide the selection of information [314, 315, 316]. At the same time, biases and information sampling that deviates from the proposed optimal behaviours have been widely reported [317, 318, 319, 203, 320, 321, 322, 323].

Confirmation bias is defined as the tendency of agents to seek out or overweight evidence that aligns with their beliefs while avoiding or un-
derweighting evidence that contradicts them [324, 325, 326, 327, 328, 329]. Recently, cognitive scientists have realised confirmation bias may reveal a fundamental property of how the brain drives information search. In this line, confirmation bias has been found not only in complex decisions where people choose media channels to get their news and information [330, 331, 332] but also in much simpler perceptual or value-based choices [333, 334, 335, 336]. The origin of the bias has been suggested to be in the weighting of incoming information or the search for confirmatory evidence [319, 337, 338, 329]. However, the role of information sampling generated autonomously by the agent has not been studied. This offers a way to explicitly explore how the sampling strategy is implemented to fulfil the internal objectives of the decision.

Confidence likely influences information sampling before choice, potentially interacting with the confirmation bias. [339] recently showed that confidence indeed affects a neural signal of confirmatory evidence integration. The standard view of confidence implies that the higher the decision confidence, the stronger the decision-makers’ belief is in the correctness of this choice [70, 340]. The objective of this study is to explore the relevance that confidence has on this phenomenon, in particular the role that it has on information sampling. We predict that confirmation bias in information sampling will be stronger after choices are made with high confidence.

To approach this question, we designed a perceptual binary decision task that includes two choice phases separated by a sampling phase. We tracked sampling patterns using eye-tracking. The purpose of the sampling step was to test our hypotheses that confirmation bias arises from biased information sampling, that this effect influences future choice and, crucially, that confidence mediates this behavioural tendency. We additionally explored whether the participant’s autonomous generation of the sampling patterns affects confirmation bias. We found that freely allocated sampling was more connected with confirmation bias, which was also reflected in the computational modelling of attention in the decision. Overall, our findings highlight a central role played by confidence reflecting participants’ beliefs, which is echoed in the sampling of the alternatives.

6.3 Methods

6.3.1 Participants

Twenty-three participants completed this experiment, of which two were excluded from the analysis because they gave the highest possible confidence rating on more than 75% of trials. Another three participants were excluded because their confidence was a poor predictor of their accuracy in the task (lower than two standard deviations under the mean coefficient predicting accuracy in a logistic regression). Participants were reimbursed £30 for their time as well as an additional amount between £0 and £20 that could be gained in the task. Participants had a normal or corrected-to-normal vision, and no psychiatric or neurological disorders. We obtained written informed consent from all participants before the study. This experiment was approved by the University of Cambridge Psychology Research Ethics Committee.
6.3.2 Experiment

Participants completed a perceptual decision task where two sequential choices were made. Each trial involved a first stage where two dot patches were presented, and participants were instructed to choose the one with a higher number of dots. In the second stage, participants could resample the patches before confirming or changing their initial decision. For the first choice, the stimulus was presented for 500ms each, with a randomised presentation order for the left and right options. Participants' responses were elicited through eye movements. To make a choice they looked at one of the patches and to rate their confidence they looked at a position inside a rating scale. After this first choice, the sampling phase between the two choices randomly varied duration, and it was either 3, 5, or 7 s. This was not cued to the participants, so at the start of the sampling phase, they did not know how much time they would be given to sampling the alternatives. Sampling was completely gaze-contingent meaning that the dots in a circle were only visible when the participant fixated inside that circle and participants could only see one circle at a time. This was done by tracking when the participant’s gaze was within one of two pre-defined square areas of interest (AI) centred on the two stimuli. Between each phase of the trial, participants had to fixate on a central fixation cross. Furthermore, we introduced a control condition in which participants were not free to sample the circles however they liked during the sampling phase. Instead, in one-third of trials the patches were shown for an equal amount of time each and in two-thirds of trials one patch was shown three times longer than the other (50% of trials the left was shown longer, 50% of trials the right was shown longer). Participants were constantly reminded of their initial choice by the circle surrounding the chosen patch changing colour. Participants took part in two sessions, each consisting of 189 trials. In one session, they performed the main task and in the other the control condition of the task. The order of these two sessions was pseudo-random. This experiment was programmed using the SR Research Experiment Builder version 1.10.1630 (SR Research Experiment Builder, 2017).

6.3.3 Eye-tracking

Eye movements were recorded at a rate of 1000 Hz using an EyeLink 1000 Plus eye-tracker. Areas of Interest (AI) for the eye tracking analyses were pre-defined as two squares centred on the gaze-contingent circles in the experiment. The sides of the squares were the same as the diameter of the gaze-contingent circles. For each decision period, we derived the total dwell time in each AI from the eye-tracking data. The computer used in this experiment had a screen size of 68.58 × 59.77 cm and participants were seated 60 cm away from the screen.

6.3.4 Analyses

We studied the effect of choice on the time spent on each of the two stimuli using paired sample t-tests on the mean sampling times spent on each stimulus from each participant. Trials with the shortest sampling phase length of 300ms were excluded from all analyses because it became apparent that this time was too short for participants to be able to saccade to each circle more than once.
6.3.5 Hierarchical models

Hierarchical regression models were conducted using the lme4 package in R [183, 102]. All models allowed for random effects (at the participant level) intercepts and slopes. We computed degrees of freedom and p-values with the Kenward-Roger approximation, using the package pbkrtest [341]. We predicted the sampling time difference between the two circles using a hierarchical linear regression model. To predict choice in the second-choice phase, hierarchical logistic regressions were used to predict the log odds ratio of picking the left circle on a given trial. Confidence and sampling time were z-scored on the participant level.

6.3.6 Attentional model - GLAM

The Gaze-weighted Linear Accumulator Model [35, 123] is part of the family of linear stochastic race models in which different alternatives (i; left or right) accumulate evidence (Ei) until a decision threshold is reached by one of them, determining the chosen alternative (the details of this model have been presented in the Methods section of Chapter 3).

The model fit with GLAM was implemented at a participant level in a Bayesian framework using PyMC3 [184]. Uniform priors were used for all the parameters:

\[ v \sim \text{Uniform}(1^{-6}, 0.01) \]

\[ \gamma \sim \text{Uniform}(-1, 1) \]

\[ \sigma \sim \text{Uniform}(1^{-6}, 5) \]

\[ \tau \sim \text{Uniform}(0, 5) \]

We fitted the model for each individual participant and for free and fixed sampling conditions separately. To model participants’ behaviour, we used as input for GLAM the reaction times (RT) and choices obtained from phase 3, and relative gaze for left and right alternatives for each trial during sampling phase 2. For fixed sampling trials, the presentation times of the dot patches were used to calculate the relative gaze time. For both conditions, model fit was performed only on even-numbered trials using Markov-Chain-Monte-Carlo sampling, we used implementation for No-U-Turn-Sampler (NUTS), four chains were sampled, 1,000 tuning samples were used, and 2000 posterior samples were used to estimate the model parameters. The convergence was diagnosed using the Gelman-Rubin statistic (\(|\hat{R} - 1| < 0.05\)) for the four parameters (\(v, \gamma, \sigma,\) and \(\tau\)). Considering all the individual models (18 participants), we found divergences in \(\sim20\%\) of the estimated parameters (\(\sim16\%\) in free; \(\sim25\%\) in the fixed condition). We removed the participants that presented divergent parameters (7 participants) to check whether the results we found were driven by these data. The significantly higher gaze bias in free-viewing condition was maintained even after removing these participants. Model comparison was performed using Watanabe-Akaike Information Criterion (WAIC) scores available in PyMC3, calculated for each individual participant fit. Note that in the fixed condition a model without gaze bias was more parsimonious than the model including the \(\gamma\) parameter.
The fact that the gaze model was not the most appropriate to capture the data in the fixed condition may explain why we observed more parameter divergences in that case.

To check how well the model replicates the behavioural effects observed in the data [185], simulations for choice and RT were performed using participants’ odd trials, each one repeated 50 times. For each trial, number of dots and relative gaze for left and right items were used together with the individually fitted GLAM parameters to simulate the trials. Random choice and RT (within a range of the minimum and maximum RT observed for each particular participant) were set for 5% of the simulations, replicating the contaminating process included in the model as described by [35].

6.4 RESULTS

To test the hypothesis that confirmation bias is echoed in the sampling profiles, we designed a perceptual experiment where participants would have to possibility to confirm or change their minds on previous decisions. Participants performed a two-alternative forced choice (2-AFC) on the presentation of two patches of dots (Figure 6.1). In the first stage (phase 1), participants had to choose whether the left or right-side patch had a higher number of dots, after a quick presentation of 500 ms for each alternative. To explore the relevance of internal certainty estimates in confirmation bias, participants rated their confidence in the phase 1 choice immediately after their choice. After their initial decision, we tested how participants behaved when they could interact again with the alternatives (phase 2). Crucially, during this second stage participants could explore again the same patches as in phase 1 in a sampling phase, after which a new choice and confidence were reported. The length of the sampling phase was manipulated to be either 3s, 5s, or 7s, to control for exposure time. Furthermore, we wanted to check whether biased evidence accumulation was caused by differential exposure to the perceptual evidence, or if the sampling choices themselves drive the effect. In other words, would the same choice bias appear if participants were passive recipients of biased sampling, or does the choice bias require that participants make their own sampling decisions? Therefore, the phase 2 sampling stage considered two conditions: in a free sampling condition participants could explore the patches without constraints; in the fixed sampling conditions the time of exposure to the dots patches was predefined by the experimenter. Eye-movement information was recorded during the whole experiment, in both conditions.

6.4.1 Free sampling condition

Firstly, participants were sensitive to the difficulty of the given trials (Figure 1B) and were more accurate on the second choice compared to the first choice (t_{17} = 6.80, p < 0.001). In the free sampling condition, patch presentation was completely gaze-contingent, showing that participants spent more time viewing the patch they just chose (Figure 6.1C; t_{17} = 3.52, p < 0.01). Furthermore, the size of this sampling time bias was proportional to the total amount of sampling time available (please check [342] for details).

Confidence in a choice reflects the strength of the participant’s belief that their choice was correct. This experiment also showed a mediating effect of confidence on how sampling bias affects the second choice. A significant interaction between choice and confidence in the first choice was found
(Figure 6.2A–B; \(t_{16.07}=4.29, p < 0.001\)), indicating that higher confidence leads to a higher sampling bias towards the chosen option. We also found a main positive effect of choice and evidence difference (difference in the number of dots) on sampling time difference (Figure 6.2B; main effect of choice: \(t_{16.07}=2.90, p < 0.01\); main effect of evidence difference: \(t_{16.07}=9.21, p < 0.001\)). Confidence was also shown to negatively predict the total amount of time spent sampling, that is, the total time actually spent gazing at the two stimuli during the sampling phase (rather than the central fixation cross; \(t_{22.99}=4.01, p < 0.001\)). We also explored the effects on second choice confidence. In a regression analysis, we found that change of mind negatively predicted confidence in the second choice phase (\(t_{16.59}=6.39, p < 0.001\)), i.e., participants reported lower confidence after they changed their choice in the second stage. No effect of sampling time difference on the second confidence rating was found (\(t_{161.30}=-0.78, p = 0.44\)).

Our choices are determined by the evidence we accumulate before making a decision. Therefore, biased sampling in the free sampling phase (phase 2) was expected to affect decisions in the second decision phase. Specifically, we hypothesised that the more strongly participants preferred sampling their chosen patch, the more likely they were to choose it again. In other words, participants would not change their mind when they had a higher sampling bias. Following this hypothesis, we found a negative effect of sampling bias on subsequent change of mind (Figure 6.2C; \(z = -7.20, p < 0.001\)) as well as the main negative effects of evidence difference and (first choice) confidence on change of mind (Figure 6.2C; main effect of evidence difference: \(z = -2.66, p < 0.01\); main effect of confidence: \(z = -8.73, p < 0.001\)).

It has been shown that the uncertainty around internal estimates scales with numerosity [343]. As such, an alternative explanation for the sampling biases found in this experiment might be that participants were minimising uncertainty by sampling the option with more dots (the correct choice alternative) for longer. To exclude this scenario, we run additional regression analysis predicting the difference in sampling time, including numerosity (the total number of dots present on the screen) as a predictor. Choice and evidence difference were also included as predictors, as in the analysis presented above. We found that neither total numerosity (\(t_{25.07}=0.62, p = 0.54\)) or the number of dots of the chosen option (\(t_{87.40}=0.25, p = 0.81\)) had a significant effect on sampling bias, meaning that participants’ sampling was not biased by a drive to reduce uncertainty by sampling the option with more dots. Furthermore, we reanalysed data from a previous perceptual decision study ([186] and Chapter 3 of this thesis) where participants also had to choose between two circle stimuli with dots (but were not requested to resample the stimuli and choose for a second time). In this study, on some trials participants had to choose the option with the most dots (the ‘most’ frame), and on other trials the option with the least dots (the ‘fewest’ frame). In this dataset, we found a significant effect of choice on sampling time difference in both frames (‘most’ frame: \(t_{57.54}=24.01, p < 0.001\); ‘fewest’ frame: \(t_{13.72}=14.97, p < 0.001\) and no significant effect of the total number of dots on sampling time difference (‘most’ frame: \(t_{31.41}=-0.37, p = 0.72\); ‘fewest’ frame: \(t_{10.01}=1.49, p = 0.14\), meaning that participants’ sampling was not biased by numerosity. Overall, these results seem to indicate that numerosity is not significantly affecting the sampling process in this task.

The sequential order of presentation in the initial sampling phase before the first choice might also be expected to affect sampling. To exclude this possibility, we performed a regression analysis predicting sampling time
difference as a function of presentation order and found no effect \( t_{49.65}=0.08, p = 0.93 \). It is also important to note that the stimulus chosen in the first choice phase was highlighted throughout the trial. This was done to reduce the likelihood of working memory being a confound on this task, but we recognise the possibility that it may have interacted with the main effect of choice on sampling.

### 6.4.2 Fixed sampling condition

An important point is to distinguish whether the difference in sampling effect is a product of a biased evidence accumulation caused by differential exposure to the perceptual evidence, or if the sampling choices themselves drive the effect. In other words, would the same choice bias appear if participants were passive recipients of biased sampling, or does the choice bias require that participants make their own sampling decisions? We addressed this question by including a separate set of trials in which we introduced a control task, a 'fixed-viewing condition'. Here participants did the same two-stage decision task but did not have the possibility to freely sample the patches in phase 2. Instead, the dot patches were shown for a set amount of time. In one-third of trials, the patches were shown an equal amount of time; in two-thirds of trials, one patch was shown three times longer than the other. Participants completed during the same experiment fixed-sampling and free-sampling sessions, with the order pseudo-randomised between participants. If the effect of confirmation bias on subsequent choice is observed when confirmatory evidence is passively presented, we would expect to see the same effect in the fixed-viewing condition (in which asymmetric information is provided by the experimenter) as in the free-sampling condition. Alternatively, if it is required that confirmatory evidence is actively sought by the decision-maker to observe a confirmation bias, we would expect that the effect of biased information sampling on the subsequent choice disappears in the fixed-viewing condition.

In line with the second prediction, we found that in the fixed-viewing condition, contrary to the free-sampling condition, the amount of time spent viewing the patches in the sampling phase did not significantly affect subsequent choice. In a hierarchical logistic regression predicting the change of mind from the first to the second choice within a given trial, the main effect of sampling bias on change of mind was completely offset by the positive effect of the interaction term between sampling bias and a dummy variable that was set to 1 if a trial was in the fixed-viewing condition (Figure 6.2C–D; \( z = 6.77, p < 0.001 \)). This means that there was no effect of sampling bias on change of mind in the fixed-viewing condition. To check that participants were engaging in this version of the task, we looked at whether the number of saccades made within each patch during the sampling phase was similar between the two tasks. We found that the number of saccades was actually higher in the fixed-viewing condition than in the main experiment \( t_{17} = -4.22, p < 0.001 \), which means participants were indeed exploring the information in this condition. Furthermore, no significant difference in accuracy was observed between the two conditions \( t_{17} = 1.51, p = 0.14 \), though sensitivity to decision evidence was slightly higher in the second choice in the free sampling condition compared to the fixed sampling condition. The number of changes of mind was also equal between the two conditions \( t_{17} = 0.75, p = 0.47 \) as well as both confidence ratings (confidence in first choice: \( t_{17} = -1.38, p = 0.19 \); confidence in second choice: \( t_{17} = 0.5, p = 0.62 \)).
Figure 6.1: Task design and participant behaviour for the experiment. (A) Task structure. Participants had to choose which of two dot patches contained the most dots after viewing each for 500ms (phase 1) and rate their confidence in this choice. Then participants were given 3000ms, 5000ms, or 7000ms to view the dots which they could allocate between the two patches in whichever way they liked (phase 2) by looking inside the circles. Finally, participants made a second choice about the same set of stimuli and rated their confidence again (phase 3). (B) Participants effectively used the stimuli to make correct choices and improved their performance on the second choice. This psychometric curve is plotting the probability of choosing the left option as a function of the evidence difference between the two stimuli for each of the two choice phases. (C) In the free sampling condition during the sampling phase (phase 2) participants spent more time viewing the stimulus they chose in the preceding choice phase than the unchosen option. Data points represent individual participants. Reprinted from [342], eLife Sciences Publications.
6.4 RESULTS

Figure 6.2: The effect of choice on sampling behaviour is mediated by confidence in experiment. Participants were less likely to change their minds if they showed a strong sampling bias for their initially chosen option in the sampling phase, but this was only the case in the free sampling condition. (A) Sampling bias in favour of the chosen option increases as a function of confidence in the initial choice. Confidence and sampling bias towards the chosen option are both normalised at the participant level in this plot. (B-C) Plotted are fixed-effect coefficients from hierarchical regression models predicting the sampling time (how long each patch was viewed in the sampling phase) difference between the left and right stimuli. Data points represent regression coefficients for each individual participant. (B) There is a significant main effect of choice on the sampling time difference, such that an option is sampled for longer if it was chosen, and a significant interaction effect of Choice x Confidence, such that options chosen with high confidence are sampled for even longer. (C) There is a main negative effect of sampling bias on change of mind, such that participants were less likely to change their mind in the second decision phase (phase 3) the more they sampled their initially chosen option in the free sampling phase (phase 2). The main effect of sampling bias on change of mind disappears in the fixed sampling condition, which can be seen by the positive interaction term Sampling bias x Fixed sampling which entirely offsets the main effect. The analysis includes a dummy variable ‘Fixed Sampling’ coding whether the trial was in the fixed-viewing condition. (D) The probability that participants change their minds on the second-choice phase is more likely if they looked more at the unchosen option during the sampling phase. The plot shows the probability that participants changed their minds as a function of the time spent sampling the initially chosen option during phase 2. The lines are polynomial fits to the data, while the data points indicate the frequency of changes of mind binned by sampling bias. Note that the actual gaze time of the participants is plotted here for both task conditions. The same pattern can be seen when instead plotting the fixed presentation times of the stimuli for the fixed task condition. Reprinted from [342], eLife Sciences Publications.

6.4.3 Attentional evidence accumulation modelling

To further investigate how attention, when freely allocated, shapes the accumulation of evidence and choice bias, we modelled the data from both viewing conditions using the Gaze-weighted Linear Accumulator Model (GLAM; [123, 186, 35]). GLAM belongs to the family of race models with an additional modulation by visual attention (Figure 6.3A). It is an approximation of a widely used class of models – the attentional Drift Diffusion
Model (aDDM; [33, 34]) in which the full dynamic sequence of fixations is replaced by the percentage of time spent fixating the choice alternatives. Even-numbered trials were used to fit the model while odd-numbered trials were used to test it. See the Materials and methods section for further details.

GLAM is defined by four free parameters: \( \nu \) (drift term), \( \gamma \) (gaze bias), \( \tau \) (evidence scaling), and \( \sigma \) (normally distributed noise standard deviation). The model correctly captured the reaction times (RT) and choice behaviour of the participants at group-level both in the free-sampling (Figure 6.3B) and fixed-viewing conditions (Figure 6.3C). More specifically, we found that the model predicted faster RTs when trial difficulty was low (\( |\Delta \text{Dots}| \) is high; Figure 6.3B–C, top left). The model also reproduced overall choice behaviour as a function of the number of dots in the patches (\( \Delta \text{Dots} = \text{Dots}_{\text{Left}} - \text{Dots}_{\text{Right}} \)) in both conditions (Figure 6.3B–C, top right). Furthermore, we found gaze allocation (\( \Delta \text{Gaze} = g_{\text{Left}} - g_{\text{Right}} \)) predicted the probability of choosing the correct patch in the free-sampling condition (Figure 6.3C, bottom left). However, to properly test how predictive gaze allocation is of choice, we must account for the effect of evidence (\( \Delta \text{Dots} \)) on choice. As such, we used the gaze influence (GI) measure [35], which reflects the effect of gaze on choice after accounting for the effect of evidence on choice. GI is calculated by taking the actual choice (0 or 1 for right or left choice, respectively) and subtracting the probability of choosing the left item as predicted by a logistic regression with \( \Delta \text{Dots} \) as a predictor estimated from behaviour. The averaged residual choice probability reflects GI. We found GI estimated purely from the participant’s behaviour was higher in the free-sampling than in the fixed-viewing condition (comparing average GI by participant, free-sampling condition: Mean = 0.148, SD = 0.169; fixed-viewing condition: Mean = 0.016, SD = 0.14; \( t_{17} = 2.708, p < 0.05 \)). This suggests the effect of visual attention on choice was higher in the free-sampling condition. In line with this, the model also predicted a higher GI on corrected choice probability in the free-sampling condition (comparing average GI by individual model predictions, free-sampling condition: Mean = 0.112, SD = 0.106; fixed-viewing condition: Mean = 0.033, SD = 0.028; \( t_{17} = 2.853, p < 0.05 \); Figure 6.3B–C, bottom right).

We then tested whether attention affected information integration more when information was actively sought (i.e., the free-sampling condition) compared to when information was given to the participants (i.e. the fixed-viewing condition). We compared the parameters obtained from the individual fit in the free-sampling and fixed-viewing conditions (Figure 6.3D). We found a significant variation in the gaze bias parameter (Mean \( \gamma_{\text{Free}} = 0.81 \), Mean \( \gamma_{\text{Fixed}} = 0.98 \), \( t_{17} = -3.934; p < 0.01 \)), indicating a higher influence of gaze on choice in the free-sampling condition. Note that during the fixed-viewing condition, the parameter \( \gamma \approx 1 \) indicates that almost no gaze bias was present in those trials. Conversely, there was no significant difference for the other parameters between the conditions (Mean \( t_{\text{Free}} = 1.44, t_{\text{Fixed}} = 1.13, t_{17} = 1.003; p = 0.32, \text{n.s.} \), Mean \( \nu_{\text{Free}} = 0.0077, \nu_{\text{Fixed}} = 0.0076, t_{17} = 0.140; p = 0.89, \text{n.s.} \), Mean \( \sigma_{\text{Free}} = 8.07 \times 10^{-5}, \sigma_{\text{Fixed}} = 8.70 \times 10^{-5}, t_{17} = -1.201; p = 0.24, \text{n.s.} \)). These results suggest that gathering information actively (i.e., free-sampling condition) does not affect the overall speed at which information is integrated, but it specifically modulates the likelihood of choosing the gazed-at option. Finally, to test that the identified effect did not depend on a less variable gaze time range we resampled the data from the free-sampling condition to match the gaze time range in the fixed-viewing condition and fitted the GLAM again. We replicated our finding even when the gaze time range in
Figure 6.3: Gaze impacted evidence accumulation (for the second choice) more strongly in the free than in the fixed sampling condition. (A) Free and fixed sampling condition trials were fitted separately using a Gaze-weighted Linear Accumulator Model (GLAM). In this model, there are two independent accumulators for each option (left and right) and the decision is made once one of them reaches a threshold. The accumulation rate was modulated by gaze time when the gaze bias parameter is lower than 1 (γ < 1). In that case, the accumulation rate will be discounted depending on γ and the relative gaze time to the items, within the trials. Gaze information from the free sampling trials and presentation times from the fixed sampling trials were used to fit the models. The panel depicts an example trial: patch sampling during phase 2 (left panel) is used to estimate the relative gaze for that trial (central panel), and the resulting accumulation process (right panel). Note the GLAM ignores fixation dynamics and uses a constant accumulation term within a trial (check Methods for further details). (B) GLAM predicted the behaviour in free and fixed sampling conditions. The four panels present four relevant behavioural relationships comparing model predictions and overall participant behaviour: (top left) response time was faster (shorter RT) when the choice was easier (i.e. bigger differences in the number of dots between the patches); (top right) probability of choosing the left patch increased when the number of dots was higher in the patch at the left side (∆Dots = DotsLeft – DotsRight); (bottom left) the probability of choosing an alternative depended on the gaze difference (∆Gaze = gLeft – gRight); and (bottom right) the probability of choosing an item that was fixated longer than the other, corrected by the actual evidence ∆Dots, depicted a residual effect of gaze on choice. Note that in the free condition, the model predicted an effect of gaze on choice in a higher degree than in the fixed condition. Solid dots depict the mean of the data across participants in both conditions. Lighter dots present the mean value for each participant across bins. Solid grey lines show the average for model simulations. Data was z-scored/binned for visualisation. (D) GLAM parameters fitted at the participant level for free and fixed sampling conditions. Free sampling condition presented a higher gaze bias than fixed sampling, while no significant differences were found for the other parameters. γ: gaze bias; τ: evidence scaling; ν: drift term; σ: noise standard deviation. **: p < 0.01. Reprinted from [342], eLife Sciences Publications.
the free-sampling condition was reduced to match that in the fixed-viewing condition.

6.5 Discussion

In this work, we have demonstrated that a form of confirmation bias exists in active information sampling, and not just in information weighting as previously thought. Using a novel experiment we showed that this effect is robust for simple perceptual choice (for an additional study confirming the presented findings please check [342]). Critically we show that the sampling bias affects future choice and that the presence of confirmation bias is also modulated by participants’ internal beliefs, captured by decision confidence. Furthermore, we demonstrated that this effect is only present in the free-sampling condition, showing that agency is essential for biased sampling to affect subsequent choice.

Preference for confirmatory evidence has been previously studied in the context of strongly held political, religious, or lifestyle beliefs, and not in perceptual decision-making [330, 324, 325, 326, 327, 328, 329, 330]. Our results, together with the recent work of others [334, 345, 336], show that confirmation bias in information search is present even in simple perceptual decisions that have no meaningful impact in participant’s daily lives. This supports that confirmation bias might be a fundamental property of information sampling and the decision process, existing irrespective of how important the belief is to the agent.

We show that confidence modulated the confirmation bias effect: choices made with higher confidence led to the increased sampling of the chosen option and an increased likelihood of choosing the same option again in the second choice phase. This shows that the strength with which a belief is held determines the size of the confirmation bias in active information sampling. Confidence has been shown to affect the integration of confirmatory evidence as reflected in MEG recordings of brain activity during evidence accumulation [339]. Even more, recent studies in economics and neuroscience have given theoretical and experimental proof of a relationship between overconfidence, and extreme political beliefs [346, 3]. Our results suggest that altered information sampling could be the missing link between confirmation bias and overconfidence. Specifically, given that we have shown that increased confidence leads to increased confirmation bias, it follows that overconfidence in a belief would lead to an increased sampling of confirmatory evidence in line with that belief, which in turn would lead to even higher confidence. We also found that high confidence impacted negatively on the time that participants spent in overall sampling, reflecting a decreased urgency by participants for sampling new information. In other words, confidence plays an important role in specifying future sampling behaviours and, consequentially, in future choices.

Recent findings suggest that biases in information sampling might arise from the Pavlovian approach, a behavioural strategy that favours approaching choice alternatives associated with reward [319, 347]. Furthermore, the number of hypotheses an individual can consider in parallel is likely to be limited [348]. As such, it may be advantageous to first attempt to rule out the dominant hypothesis before going on to sample from alternative options. In this vein, the sampling bias we see could be the solution to an exploit-explore dilemma in which the decision-maker must decide when to stop ‘exploiting’ a particular hypothesis (on which stimulus has the most
6.5 Discussion

An important aspect of this task is that participants were able to freely sample information between choice phases, providing a direct read-out of confirmation bias in the active sampling decisions made by the participants. Previous accounts of confirmation bias in perceptual choice have instead focused on an altered weighting of passively viewed information as a function of previous choice [333, 339, 334]. However, from these findings, it remained unclear to what extent this bias manifests in the processing of information compared to the active sampling of information. Our findings show that active information sampling plays a key role in the amplification of beliefs from one decision to the next and that changes in evidence weighting likely only account for part of observed confirmation bias effects.

The results from the GLAM model show that, in line with previous studies [33, 186, 121, 35], a specific boost in the accumulation of evidence of the visually attended items was found in the free sampling condition. Conversely, a disconnection between an item’s sampling time and evidence accumulation was found in the fixed condition (i.e., the absence of gaze bias in GLAM implies that visual fixations did not affect evidence integration when the sampling was not controlled by the decision-maker). One explanation for this result is that attentional allocation itself is directed towards the options that the participants reckon are more relevant for the task to be performed, i.e., aligned with the goal (Chapter 3 and [186]). In our experiment, the goal of the task was theoretically identical for the first and second choice (i.e., to find the patch with more dots). However, as discussed above, it could be that participants perceived the goal of the second choice to be slightly different from the goal of the first: in the second case, they had to verify whether their initial choice was correct, as well as to find the patch with the most dots. This resembles the economic model of confirmation bias mentioned above [342]. Accordingly, the chosen option is bestowed with higher relevance and then more attention (consequently boosting evidence accumulation). On the other hand, since in the fixed sampling condition participant’s attention allocation is not necessarily associated with their goals, the difference in display time of the items is ignored or cannot be consistently integrated into their decision process. A complementary perspective has been given in recent work by Jang et al. [349] and Callaway et al. [350]. Their models characterise attention’s role as lowering the variance of the accumulated evidence towards the attended option, which in turn updates the internal belief. Crucially,
they have characterised attention as an internally generated factor that is allocated following optimal sampling policies, unlike other models that take attention as an exogenous factor [33, 34, 35]. Perhaps, in a similar way to the proposition above concerning internal goals, it could be the case that the exogenous sampling pattern we imposed in our experiment was not aligned with the optimal evolution of the internal beliefs of participants, therefore, the ‘offered’ evidence was misaligned and therefore did not impact choice. Further modelling work testing Jang and Callaway’s proposals in regimes where attention is exogenously controlled can give insight into the relevance of attentional agency. An alternative hypothesis is that gaze allocation itself is used as additional evidence in favour of a choice alternative. This would mean that when an agent has previously attended to a choice alternative, this is used as evidence in favour of that option in and of itself. Therefore, when gaze allocation is not under the agent’s control, as in the fixed-viewing condition, it is not used to inform choice.

Our findings imply that agency plays a clear role in evidence accumulation, and consequently in confirmation bias. It also suggests that common experimental designs in which information is provided by the experimenter and passively sampled by the participant might not be an ecological way to study decision-making. These tasks mask the potentially large effect of active information search on belief formation and choice. [351] recently showed that people were less likely to share false information online if they had been asked to rate the accuracy of a headline just previously. It may therefore be possible to reduce confirmation bias in information search in a similar way by priming participants to attend more to accuracy instead of confirmatory evidence. More complex behavioural tasks are required to improve our understanding of the different drivers of information sampling and how sampling in turn guides future choices [352].

To summarise our findings, we observed that participants sampled more information from chosen options in a perceptual choice paradigm and that this sampling bias predicted subsequent choice. Asymmetric sampling in favour of the chosen alternative was stronger when participants’ confidence in their first choice was higher. We discuss how this variation in the sampling bias could be driven by changes in participants’ internal goals given that they need to corroborate their previous choices. The more committed participants are to their previous selection, as reflected by confidence, the more changed their expectations and sampling behaviour for the second choice. This is in line with our findings in Chapter 3 of this thesis, sampling biases were associated with the setting of goals. Furthermore, the effect of information on subsequent choice was only seen in a version of the task where participants could sample freely, suggesting agency plays an important role in the propagation of strongly held beliefs over time. In other words, if the evidence offered to participants was not aligned with their internal goals, their decision process did not integrate that information. Overall, these findings suggest that confirmatory information processing might stem from a general information sampling strategy used to seek information to strengthen prior beliefs rather than from altered weighting during evidence accumulation only and that active sampling is essential to this effect. Biased sampling may cause a continuous cycle of belief reinforcement that can be hard to break, which could be manifested more extremely in psychiatric conditions, such as major depression [353]. Improving our understanding of this phenomenon can also help us better explain the roots of extreme political, religious and scientific beliefs in our society.
As presented in previous chapters, goals have a critical role in the modulation of attentional deployment and brain representations during the deliberation period before choice. In the present chapter, we show that goal’s influence extends into the post-decision processes as well. Each decision generates an internal estimate of its certainty, a confidence measure. Although typically confidence is considered as the probability the choice is correct, recent findings have shown a more nuanced perspective on the factors involved in its generation. In the following experiment, we changed the decision frame to probe how confidence is affected by goals, in value-based and perceptual binary choice experiments. We show that confidence was modulated not only by the difficulty of the choice but also by the overall amount of evidence presented for the decision. Confidence was enhanced specifically by the goal-relevant evidence, showing that frame manipulation played a role in confidence generation. We simulated confidence using accumulator models, capturing participants’ behaviour only if the model included an attentional bias. These results show that goals have a pervading influence on the decision process, reaching participants’ estimation of uncertainty.

Humans can introspect on their decisions by monitoring their uncertainty and reporting their level of confidence [340, 57, 75, 66]. According to the standard view, confidence is defined as the probability a choice is correct. This definition implies that confidence refers to a single trial and the evidence available for this choice, not to a general ability to perform a task or other contextual information. From this perspective, confidence is typically obtained as a measure of the difficulty of the choice, e.g., the difference in evidence between the alternatives (ΔEvidence). Decisions with more distinct alternatives (e.g., a rotten banana vs a fresh apple) should be reported with higher confidence, while closer alternatives in terms of evidence (e.g., a ripe banana vs a fresh apple) should generate more difficult choices with lower confidence.

As presented in previous chapters, goals and context are important factors that influence the decision process. We have described how simple frame manipulations have an impact on evidence accumulation, affecting attentional allocation during the deliberation process [186]. In the experiments presented in Chapter 3, we introduced a positive and negative frame where participants had to choose the best (i.e., the option with higher evidence) or worst alternative (i.e., the option with lower evidence). According to the canonical definition of confidence, such framing manipulation should not
influence confidence. However, since most of the perceptual and value-based studies of confidence have not manipulated goals, this hypothesis has not been tested. Additionally, dynamic factors in the decision, such as attentional allocation, most of the time are not included in the analysis of decision confidence (although see [354, 355]). From our previous findings on attention and goals, we expected that goal-oriented gaze biases could be relevant to the construction of participants’ confidence reports.

The main objective of this chapter is to probe whether goals affect the construction of confidence in value-based and perceptual choices. The results presented here are the second part of the findings on our value-based and perceptual experiments presented in Chapter 3, i.e., this is new data obtained from the same experiment. Here participants completed binary choices in two frames: they indicated the food item that was liked or disliked in a value-based decision; and selected the patch with the most or fewest number of dots in a perceptual experiment. Participants reported their confidence after each choice. We found a significant effect of overall evidence (i.e., the evidence presented together in the chosen and unchosen alternatives) on confidence reports, which was dependent on the goals. This result is in opposition to the canonical understanding of confidence where frames should be irrelevant.

The second objective of this chapter is to capture the computational underpinnings of goal-relevant choices and the generation of confidence. Sequential sampling models have been extremely successful in giving a computational account of the decision process, including accuracy and reaction times (RT) measures. The Drift Diffusion Model (DDM) describes the stochastic sampling of evidence represented by a single accumulator of evidence [110, 20]. In binary choices, this accumulator, after gathering enough evidence, reaches an upper or lower boundary, which indicates the chosen option and RT. Although widely popular, this model cannot capture confidence straightforwardly, since the terminating criteria always request the same level of evidence, without making a distinction between difficult and easy decisions. Alternatively, in race models, each alternative is described by separate accumulators, with all of them sharing a single boundary. The first accumulator crossing the boundary finalizes the evidence integration and defines the choice and decision time. At the endpoint, the difference between the accumulators can be used as a model representation of the confidence level, denominated the Balance of Evidence [76, 23, 66]. In this chapter, we used this methodology to calculate confidence in our goal-dependent task. We used a race model that included an attentional parameter modulating the accumulation of evidence [35] to test the hypothesis that attention played a role in the goal-dependent confidence effects. Using balance of evidence simulations, we captured the goal-dependent pattern of results found for confidence. Furthermore, we showed that this behaviour was present only if visual gaze modulation was included in the model. Considering these results, we discovered that goals affected not only the deliberation stage but also the post-decision processes such as the construction of metacognitive reports.
7.3 METHODS

7.3.1 Experiment Design

The data used for the analysis in this chapter was collected together with the experiment presented in Chapter 3. In this experiment, a group of participants completed binary choices in value-based choices (food item preferences) and another group of individuals performed a perceptual choice task (dot numerosity). Two frames were defined in both experiments: in value-based like and dislike frames cue participants had to choose the snack they would prefer to eat or the one they wanted to reject, respectively; in the perceptual task, participants completed most and fewest frames where they had to choose the dot patch with a higher or a lower number of dots, respectively. After each choice, at the end of the trial, participants reported their confidence in the decision using a slider. Participants moved the slider to the left to indicate low confidence trials and to the right side to report high confidence. Visual attention was measured using eye tracking during all the trials. For more details on the experimental setup and on choice behaviour and its associated visual attention patterns, please check Chapter 3. In this study, we focused on the confidence reports given by participants. Eye tracking was not considered during the confidence report stage, but gaze allocation information obtained during the choice stage was included in the analysis.

7.3.2 Participants and Exclusion criteria

Participants used in this analysis are the same as in Chapter 3. Please check Chapter 3 for the details on the exclusion criteria. In the value-based experiment, 31 participants passed the exclusion criteria and were included in the analysis (16 females, 17 males, aged 20–54, mean age of 28.8). For the perceptual experiment, 32 participants (22 females, 10 males, aged 19–50, mean age of 26.03) were included in the behavioural and regression analyses. Due to instability in parameter estimation (problem of MCMC convergence), four additional participants were removed from the GLAM modelling analysis in the perceptual experiment.

7.3.3 Data analysis: behavioural data

The analysis of the behavioural data for confidence was performed using hierarchical regression analysis. Fixed-effects confidence intervals were calculated by multiplying standard errors by 1.96 [102]. The value of 1.96 is extracted from the standard normal distribution, where $P(-1.96 < Z < 1.96) = 0.95$, i.e., there is a 95% probability that $Z$, a standard normal variable, will fall between -1.96 and 1.96. Because these confidence intervals are estimates that do not take the covariance between parameters into account, they should not be interpreted too strictly, but rather serve to give the reader a sense of the precision of the fixed-effect coefficients [57]. We used the mixed-effect parameters obtained for each participant to calculate the significant differences between frames in both experiments. Predictors were all z-scored at the participant level. Matplotlib/Seaborn packages were used for visualisation. All the hierarchical analyses were performed using lme4 package [183] for R integrated into a Jupyter notebook using the rpy2 package (https://rpy2.readthedocs.io/en/latest/).
Chapter 7: Decision Goals and Their Impact in Confidence

7.3.4 Modelling

7.3.4.1 Value Experiment

We fitted the GLAM model [35] to capture participants’ behaviour. For details on the model, please check the Methods section of Chapter 3. Choices, item preference values and RT were used as input to the model, separately for like and dislike frames, as presented in Chapter 3. Note that for the dislike frame, we fitted the model using the opposite value of the items, the goal-relevant evidence. In Chapter 3 we simulated choice behaviour using the model fitted to each participant’s data. In this analysis, using those same model parameters at the participant level, we simulated the accumulation process in each trial to obtain a measure of the balance of evidence [76, 124]. Balance of evidence in accumulator models has been used previously as an approximation to the generation of confidence in perceptual and value-based decision experiments [124, 356, 66]. The purpose of this analysis was to replicate the effect of the overall evidence (ΣValue) on confidence (please see the Results section for the behavioural results on confidence). Given that GLAM includes a parameter biasing evidence accumulation by the gaze allocation, we also checked whether the behaviour observed at the level of confidence depends on considering the information of the attentional patterns. Consequently, using the value of the items and gaze ratio from odd-numbered trials, we simulated two accumulators (Equation 1 in Chapter 3), one for each alternative. Our simulations used the GLAM parameters obtained from the participant’s fit. Once the boundary was reached by one of the stochastic accumulators (fixed boundary = 1), we extracted the simulated RT and choice. The absolute difference between the accumulators when the boundary was reached (Δe = |E_{right}(t_{final}) - E_{left}(t_{final})|) delivered the balance of evidence for that trial. In total 37,200 trials were simulated (10 repetitions for each one of the trials completed by the participants). A linear regression model to predict simulated Δe using |ΔValue|, simulated RT and ΣValue as predictors was calculated with the pooled data from the simulations. This model was selected because it was the most parsimonious in predicting participants’ confidence in the Value Experiment, i.e., model with lower Bayesian information criteria (BIC) in the model comparison. The best model predicting confidence includes the gaze shift frequency (GSF) as a predictor in the regression, but since GLAM does not consider the gaze dynamics we removed it from the model. Δe simulations using a GLAM without gaze influence (i.e., equal gaze time for each alternative) were also generated, to check if gaze difference was required to reproduce ΣValue effect over confidence. The parameters fitted for individual participants were also used in the no-gaze difference simulation. The same linear regression model (Δe ~ |ΔValue| + simulated RT + ΣValue) was used with the data simulated with no-gaze difference. Note that our simulations of the decision process generated synthetic response times and Δe for the same decisions performed by human participants.

7.3.4.2 Perceptual experiment

We performed confidence simulations for the Perceptual Experiment in a similar way to the Value-based Experiment, fitting the GLAM model separately for most and fewest frames. We adapted the evidence input to the model to capture goal-relevant information in the fewest frame, as indicated in the Methods section in Chapter 3. We tested the effect of ΣDots
on confidence in the simulations with a similar linear regression model to the
one used in the Value Experiment. The regression predicted the simulated
Δe using |ΔDots|, simulated RT and ΣDots as parameters, pooling the data
across simulations. Δe simulations using a GLAM without gaze asymmetry
were also calculated in this case. All the figures and analysis were done in
Python using GLAM toolbox and custom scripts.

7.4 RESULTS

The results presented here correspond to the analysis of the confidence re-
sponses in the choice tasks presented in Chapter 3. That chapter describes
two experiments where participants completed binary choice decisions on
value-based (i.e., food items) or perceptual input (i.e., dot numerosity).
Choices were made in two frames: in the value experiment, a like or dislike
frame; in the perceptual experiment, the most or fewest dot frames. Addi-
tionally, eye-tracking information was recorded during the deliberation stage
for each decision. After each choice, participants reported their confidence in
having made the correct decision. In this section, the results of the analysis
of confidence reports are presented. For details on the deliberation process
and choice, please check Chapter 3.

7.4.1 Which factors determine confidence?

7.4.1.1 Value experiment

To analyse how participants’ goals affected the generation of confidence, we
fitted a hierarchical linear model (Figure 7.1A). We presented here the most
parsimonious model for value-based decisions (i.e., the model with lower
BIC in the model comparison, with alternative models including partial
configurations of the same parameters, please check [186] for the details).
In the model, confidence was predicted by decision difficulty (ΔValue, the
difference in preference value between the two alternatives), reaction time
(RT), the summed value considering both presented alternatives (ΣValue),
and the gaze shifts made before choice (gaze shift frequency, GSF). As was
the case for the results presented in Chapter 3 for the choice regression, the
results for the confidence regression in the like frame replicated all the effects
reported in a previous study from our lab [57]. We found that the magnitude
of ΔValue (|ΔValue|) had a positive influence on confidence in like (z = 5.465,
p<0.001) and dislike (z = 6.300, p<0.001) frames, indicating that participants
reported higher confidence when the items had a larger difference in value;
this effect was larger in the dislike frame (t(30) = -4.72, p<0.01). RT had a
negative effect on confidence in like (z = -6.373, p<0.001) and dislike (z =
-7.739, p<0.001) frames, that is, confidence was lower when the RTs were
longer. Additionally, we found that, in both conditions, a higher number of
gaze switches during the decision time predicted lower values of confidence
in like (z = -2.365, p<0.05) and dislike (z = -2.589, p<0.05) frames, as reported
in [57].

We then looked at the effect of the summed value of both options, ΣValue,
on confidence. As in [57], we found a positive effect of ΣValue on confidence in
the like frame (z = 3.206, p<0.01); that is, participants reported a higher
confidence level when both options were high in value. Interestingly, this
effect was inverted in the dislike frame (z = -4.492, p<0.001), with a significant
difference between the two frames (t(30)=9.91, p<0.001). Contrary to what
Chapter 7: Decision Goals and Their Impact in Confidence

Figure 7.1: Hierarchical linear regression model to predict confidence. (A) In Value Experiment, a flip in the effect of ΣValue over confidence in the dislike frame was found. (B) In Perceptual Experiment a similar pattern was found in the effect of ΣDots over confidence in the fewest frame. The effect of the other predictors on confidence in both experiments and frames coincides with previous reports [57]. All predictors are z-scored at the participant level. In both regression plots, bars depict the fixed-effects and dots the mixed-effects of the regression. Error bars show the 95% confidence interval for the fixed effect. In Value Experiment: ΔValue: difference in value between the two items (Value\textsubscript{Right} - Value\textsubscript{Left}); RT: reaction time; ΣValue: summed value of both items; ΔDT: difference in dwell time (DT\textsubscript{Right} - DT\textsubscript{Left}); GSF: gaze shift frequency; ΔDT: difference in dwell time. In Perceptual Experiment: ΔDots: difference in dots between the two circles (Dots\textsubscript{Right} - Dots\textsubscript{Left}); ΣDots: summed number of dots between both circles. **: p<0.001, ***: p<0.01, *: p<0.05. Reprinted from [186], eLife Sciences Publications.

happened in the like frame in which confidence was boosted when both items had high value, in the dislike frame confidence increased when both items had low value. This novel finding reveals that the change in context also triggered a reassessment of the evidence used to generate the confidence reports, i.e., confidence also tracks goal-relevant information.

7.4.1.2 Perceptual experiment

We repeated the same regression analysis in the perceptual decision experiment. In this task, evidence was defined instead of subjective value by dot
numerosity. The difficulty was calculated using the absolute difference in the number of dots ($\Delta Dots$) and the sum of evidence from the overall number of dots displayed on the screen ($\Sigma Dots$). We directly replicated all the results of the Value Experiment, generalising the effects we isolated to the perceptual realm (Figure 7.1B). Specifically, we found that $\Delta Dots$ had a positive influence on confidence in most ($z = 3.546, p<0.001$) and fewest frames ($z = 7.571, p<0.001$), indicating that participants reported higher confidence when the evidence was stronger. The effect of absolute evidence $\Delta Dots$ on confidence was bigger in the fewest frame ($t(31)=-4.716, p<0.001$). RT had a negative effect over confidence in most ($z = -7.599, p<0.001$) and fewest frames ($z = -5.51, p<0.001$), that is, faster trials were associated with higher confidence. We also found that GSF predicted lower values of confidence in most ($z = -4.354, p<0.001$) and fewest ($z = -5.204, p<0.001$) frames. Critically (like in the Value Experiment), the effect of the sum of evidence ($\Sigma Dots$) on confidence also changes its sign depending on the frame. While $\Sigma Dots$ had a positive effect on confidence in the most frame ($z = 2.061, p<0.05$), this effect is the opposite in the fewest frame ($z = -7.135, p<0.001$), with a significant difference between the parameters in both frames ($t(31)=14.621, p<0.001$). The magnitude of the $\Sigma Dots$ effect was stronger in the fewest frame ($t(31)=-10.438, p<0.001$). This means that the effect we observed is not something specific to the value-based choice, but it is a robust effect across decision modalities.

7.4.2 Attentional Model: GLAM

To understand the potential mechanisms behind the effect of context on confidence, we used computational models. Accumulator models characterise the decision as stochastic integration of evidence which can capture various features of the process, such as choice accuracy, response time, and confidence. The model employed in this analysis, the GLAM [35], belongs to the family of race models in which evidence is independently accumulated for each option. This integration is modulated by the allocation of visual attention, boosting the speed of evidence accumulation towards the fixated items (for the modelling details on choice data, please see Chapter 3). We hypothesised that goal-relevant integration of evidence could capture the signature effects we observed in participants’ confidence reports. Based on our previous findings that highlight the effect of goals on evidence sampling patterns (Chapter 3), we also hypothesised that the preference of attention for goal-relevant information might drive the context effects we found in confidence.

7.4.2.1 Balance of evidence and confidence

Balance of evidence is a proxy for decision confidence used in accumulator models [124, 76, 23, 66]. Balance of evidence is defined as the absolute difference between the accumulators for each option at the moment of choice, which is when one of them reaches the decision threshold (i.e., $\Delta e = |E_{\text{right}}(t_{\text{final}}) - E_{\text{left}}(t_{\text{final}})|$) (Figure 7.2A). We used the GLAM to estimate the balance of evidence from model choices and check if it captured the pattern of confidence observed in participants. To estimate $\Delta e$, we performed a large number of computer simulations using the fitted parameters for each participant in both experiments.
Figure 7.2: Balance of evidence ($\Delta e$) simulated with GLAM reproduces $\Sigma$Value and $\Sigma$Dots effects over confidence. (A) GLAM is a linear stochastic race model in which two alternatives accumulate evidence until a threshold is reached by one of them. $\Delta e$ has been proposed as a proxy for confidence and it captures the difference in the evidence available in both accumulators once the choice for that trial has been made. (B) Using $\Delta e$ simulations, we captured the flip of the effect of $\Sigma$Value over confidence between like and dislike frames. $\Delta e$ simulations were calculated using the model with parameters fitted for each participant. A pooled linear regression model was estimated to predict $\Delta e$. The effects of $\Sigma$Value predicting $\Delta e$ are presented labelled as ‘Model Sim’. A second set of simulations was generated using a model in which no asymmetries in gaze allocation were considered (i.e., no attentional biases). This second model was not capable of recovering $\Sigma$Value effect on $\Delta e$ and is labelled as ‘Model Sim No Bias’. $\Sigma$Value coefficients for a similar model using participants’ data predicting confidence are also presented and labelled as ‘Human’ for comparison. (C) A similar pattern of results is found in the Perceptual Experiment, with the model including gaze bias being capable of recovering $\Sigma$Dots effect on $\Delta e$. This novel effect may suggest that goal-relevant information is also influencing the generation of second-order processes, such as confidence. This effect may be originated from the attentional modulation of the accumulation dynamics. Coloured bars show the parameter values for $\Sigma$Value and $\Sigma$Dots and the error bars depict the standard error. The solid colour indicates the Value Experiment and the striped colours indicate the Perceptual Experiment. All predictors are z-scored at the participant level. Reprinted from [186], eLife Sciences Publications.

7.4.2.1.1 Value experiment

To confirm that the relationship between confidence and other experimental variables was captured by the balance of evidence simulations, we constructed a linear regression model predicting $\Delta e$ as a function of the trial values and the simulated RTs ($\Delta e \sim |\Delta Value| + \text{simulated RT} + \Sigma\text{Value}$). We found that this model replicated the pattern of results we obtained experimentally (Figure 7.2). We then explored whether the model was able to recover the goal-dependent effect of $\Sigma$Value on confidence (Figure 7.2B). Similarly to the participant’s behaviour, $\Sigma$Value boosted $\Delta e$ in the like frame ($\beta_{\Sigma\text{Value}} = 0.071$, t(37196) = 14.21, p<0.001) and reduced $\Delta e$ in the dislike frame ($\beta_{\Sigma\text{Value}} = -0.061$, t(37196) = -12.07, p<0.001). In other words, the effect of $\Sigma$Value on confidence was replicated in the simulations with an increase of $\Delta e$ when high value options are available to choose from. In the dislike frame, the fitted model also replicated the adaptation to the context which predicts higher $\Delta e$ when both alternatives have low value.
We additionally explored the relevance that the attentional modulation had on ΣValue effect. We found the replication of the effect for ΣValue over Δe with GLAM did not hold when the gaze bias was taken out of the model in like ($β_{\Sigma\text{Value}} = -0.007$, $t(37196) = -1.495$, $p=0.13$, ns) and dislike ($β_{\Sigma\text{Value}} = -0.002$, $t(37196) = -0.413$, $p=0.679$, ns) frames (Figure 7.2B). We also found that the effect of $|\Delta\text{Value}|$ on confidence was replicated by the simulated balance of evidence, increasing Δe when the difference between item values was higher (i.e., participants and the model simulations were more ‘confident’ when the items had a higher difference in value). Higher Simulated RT predicted lower Δe, similarly to slower participant’s RT predicted lower confidence. $|\Delta\text{Value}|$ and simulated RT effects were also present in the simulations with no gaze effect (Figure 7.3A and D).

7.4.2.1.2 Perceptual experiment

We conducted a set of similar analyses and model simulations in the Perceptual Experiment (Figure 7.2C). We found that ΣDots boosted Δe in the most frame (Most: $β_{\Sigma\text{Dots}} = 0.029$, $t(33596) = 4.71$, $p<0.001$) and reduced Δe in the fewest frame (Fewest: $β_{\Sigma\text{Dots}} = -0.088$, $t(33596) = -14.41$, $p<0.001$). As in the Value Experiment, this effect disappeared when the gaze bias was taken out of the model (Most: $β_{\Sigma\text{Dots}} = -0.0002$, $t(33596) = -0.04$, $p=0.96$, ns; Fewest: $β_{\Sigma\text{Dots}} = -0.006$, $t(33596) = -1.03$, $p=0.29$, ns). $|\Delta\text{Value}|$ and simulated RT effects were also present in the simulations with and without gaze bias, following the patterns observed in participant’s behaviour (Figure 7.3B and E).

![Figure 7.3: Pooled linear regressions to predict the balance of evidence (Δe) simulations. Here the full model results from Figure 7.2 are displayed. In the Value Experiment, the full simulations of Δe replicated the pattern of results obtained in human data (confidence results), that is, there was a flip in the sign of ΣValue effect on confidence between like (A) and dislike (D) frames. However, if the gaze asymmetry is removed, we found the effect of ΣValue over Δe disappears. The results in the Perceptual Experiment, most (B) and fewest (E) frames, mirror the findings in the Value Experiment. Reprinted from [186], eLife Sciences Publications.](image)

Overall, these results show how the model is capable of capturing the novel empirical effect on participants’ confidence reports, giving computational support to the hypothesis that goal-relevant evidence is fed to second-order processes like confidence. It also hints at a potential origin to the effects of the sum of evidence (i.e. ΣValue, ΣDots) on confidence: asymmetries in the accumulation process, in particular the multiplicative effect of attention over-accumulation of evidence, may enhance the differences between items.
that are more relevant for the frame. This consequentially boosts the level of confidence that participants have in their decisions.

7.5 DISCUSSION

The objective of this chapter was to gain a deeper insight into the impact of goals on the generation of post-decision measures of uncertainty. Here we analysed the second part of the data collected in the decision experiments introduced in Chapter 3. After each choice, we measured the trial-by-trial fluctuations of confidence in value-based and perceptual decisions. Most of the studies on confidence and metacognition have used tasks with single goals, where participants choose the option with higher evidence, i.e., an item with a higher subjective value or the patch with more dots [66, 56, 68]. In this study, to manipulate the goals we included a frame where participants had to choose the options with lower evidence. We found that the role of confidence goes beyond that of simply tracking the probability of an action being correct [340], as proposed in standard signal detection theory. Instead, it was also influenced by the perceived sense of uncertainty in the evaluation process [357, 358], and contextual cues [219]. In previous work [57], we reported how, in value-based choice, confidence was related not only to the difference in value between the two items, but also to the summed value (ΔValue and ΣValue using the current notation), and we found that confidence was higher if both items had a high value [57]. Here, we replicate this effect in both experiments in the like and most conditions. However, this effect flips in the dislike or fewest frame: in these cases, confidence increases when the summed value or number of dots is smaller. This result is particularly striking since the frame manipulation should be irrelevant for the purpose of the decision and has little effect on the objective performance. This suggests that similarly to attention (as reported in Chapter 3), the sense of confidence is also shaped by the behavioural goal that participants are set to achieve. The complex nature of confidence is relevant since it influences future behaviours and information-seeking patterns ([57, 359, 127, 3, 342] and Chapter 6).

To capture the potential mechanisms behind this goal-relevant pattern in confidence we used accumulator models. In this type of model, evidence supporting each one of the choices was integrated across time, until one of the options crosses the decision threshold. In particular, we used a model (the GLAM, [35]) that included an attentional component that decreases the speed of integration of the unattended options. In both experiments, the incorporation of goal-relevant evidence to fit the GLAM resulted in a better model fit compared with the model in which the value or perceptual evidence was integrated independently of the frame (please see Chapter 3 for details). We then modified the GLAM to include a measure of confidence defined as the balance of evidence (Δe) [124, 23, 66]. Δe was obtained from the difference between accumulators at the endpoint of the evidence integration (i.e., choice time): a higher Δe was interpreted as higher decision confidence. By calculating this measure, we confirmed that our model can replicate all the main relations between confidence, choice and RT. We then tested if the model simulation was also recovering the flip in the relationship between confidence and summed evidence (ΣValue or ΣDots) triggered by the frame manipulation. We found the model captures this effect only if the attentional bias is included in the simulations. The boost in Δe when goal-relevant evidence in both alternatives is high can be attributed to the architecture of the model: in the GLAM, gaze has a multiplicative effect over evidence
accumulation. For example, consider a case with two items of value $A_1 = 2$ and $A_2 = 1$, and a discount factor for the unattended item $u = 0.3$. Assuming the item with a higher value is gazed more we could express, in a very simplified way, the $\Delta e$ for this choice as $\Delta e_A = A_1 - A_2^*u = 2 - 1 \times 0.3 = 1.7$. Consider now two new items with identical $\Delta$Value but a higher magnitude of the $\sum$Value, $B_1 = 10$ and $B_2 = 9$. Note that since $\Delta$Value is the same, this choice in absence of attentional effect should be considered of identical difficulty than in case A ($A_1 - A_2 = B_1 - B_2 = 1$), and therefore the agent should be neither more, nor less confident. But, keeping the same attentional factors as for the first set, we have that the $\Delta e_B$ between the items increases, $\Delta e_B = B_1 - B_2^*u = 10 - 9 \times 0.3 = 7.3$ ($\Delta e_A < \Delta e_B$). This effect would not be observed if attention affected evidence accumulation in an additive way ($A_1 - (A_2 - u) = B_1 - (B_2 - u)$). Our empirical confidence data, therefore, provide further support for a multiplicative instead of an additive effect of attention into goal-relevant information [179]. Overall, these data speak in favour of a coding scheme in which the goal sets, from the beginning of the task, the allocation of attention and, by doing so, influences first-order processes such as choice, but also second-order processes such as confidence. Further empirical data will be required to test this idea more stringently.

Overall, our results show that the influence of goals goes beyond direct choices and reaches post-decision processes as the generation of confidence. This confirms confidence as an integral part of the decision pipeline fed by information the same way the decision is. Our previous findings show that the sampling of information is dependent on changes in the context, which generates asymmetries in evidence integration. Here our analysis suggests that those same goal-directed asymmetries can be at the root of the $\sum$Evidence bias on confidence. In the next chapter, this idea will be further explored. Understanding the mechanisms used by these supposedly trivial contextual manipulations to influence choice uncertainty can be of great relevance in tasks where confidence in previous stages feeds the following decisions (e.g., transactions in the stock market).
GOALS MODULATE THE EVIDENCE BIAS IN HUMAN CONFIDENCE

8.1 SUMMARY

In standard signal detection theory, the degree of confidence in a decision should be solely driven by the absolute difference in evidence between the available alternatives. However, recent work has shown that confidence is also strongly modulated by the overall amount of evidence available – an effect named the positive evidence bias - and associated with a disproportionate focus on evidence in favour of one’s choice. In the first instance, these findings are extremely puzzling. Why, in a binary decision, should one be more confident when there is stronger evidence for both options? And why should people neglect evidence in favour of an unchosen option in their confidence estimates? Well-tuned metacognitive confidence is critical for guiding behaviour; so why do such biases exist? In this work, we show through a series of experiments that these so-called ‘anomalies’ or ‘biases’ are not driven by positive evidence per se, but are instead modulated by the behavioural goal (i.e., task) of the decision-maker. We then develop a novel model (grounded in signal detection theory) that allows asymmetries in internal belief variance. We show that such a model can capture all the idiosyncratic patterns of confidence we observe in our data. Using pupillometry analysis in an auditory task, we found that goals seem to change the expectations in the processing of incoming information, reflected in lower pupil sizes when choices are aligned with the task’s objective. We suggest a novel computational mechanism by which the facilitated processing of information in line with goals modulates metacognitive confidence.

8.2 INTRODUCTION

Our decisions, on perceptual features or personal preferences, are always accompanied by an internal sense of certainty. Confidence plays a fundamental role in allowing us to assess past actions (e.g., “I’m sure I did not miss the red light!”) and to inform future choices (e.g. “I’m quite confident I’ll come back to this restaurant”). An adequate generation of confidence could be disrupted in many psychiatric disorders [360, 358, 73, 361, 362]. Therefore, understanding the factors driving confidence is of central importance in the study of human decisions.

Standard models of decision confidence characterise it as the probability a decision was correct [346], commonly captured by the contrast of evidence in favour and against the decision [23, 363]. In these scenarios, the central assumption is that confidence incorporates equal evidence related to the selected and unselected options. Recent studies in perceptual decisions have nuanced this interpretation showing that human confidence deviates from
this interpretation [364, 68, 365, 366, 75]. One of these biases shows that confidence tends to overweight the evidence that supports the chosen option relative to the unchosen (Positive Evidence Bias, PEB) [63, 367, 368, 369, 370, 371].

One of the implications of the positive evidence bias is that confidence is modulated by the mere presence of evidence. Recent observations support this view showing that the overall decision evidence (ΣEvidence) indeed boosts confidence reports [57, 186, 125]. ΣEvidence is usually captured by the sum of evidence of all the options available in the decision (e.g., the total number of dots displayed in a random dot kinetogram). This indicates that subjects tend to report higher confidence when the available alternatives have overall higher levels of evidence. Value-based and perceptual decisions present a positive effect of ΣEvidence on confidence (Chapter 7 of this thesis; [57, 186, 125]). Some of the models proposed to capture the positive evidence bias also predict this ΣEvidence effect in confidence [368, 372]. These models attribute this phenomenon to a detection-like behaviour in discrimination decisions, which is commonly interpreted as dependent on the features of the stimuli, such as the strength of perceptual evidence (e.g., visibility of Gabor patches; [373, 368, 372]).

While the positive evidence bias has been described in a variety of perceptual tasks, these experiments usually consider decision experiments with a single goal: to choose the alternative with higher evidence. Goals and contexts impact powerfully information sampling and evidence accumulation [186]; therefore, it is expected they also play a role in the estimation of the confidence signal. The main target of the present work is to show the impact that task goals have on the evidence bias in confidence. In particular, we devised experimental frames that make it possible to dissociate positive evidence from behavioural choice: in addition to the standard decisions where participants pick the option with more evidence, we included further choices where the option with lower evidence was relevant. We tested this in the value-based and perceptual domains.

We both reanalysed previous datasets and conducted new experiments. We show that frames in which participants choose high and low evidence, display distinct evidence effects on confidence. While confidence was boosted by the presence of evidence in the High Evidence frame, confidence was diminished by evidence in the Low Evidence frame. Our findings highlight that the impact of evidence on confidence may not arise from the purely perceptual features of the alternatives, since in our “negative” frames the lack of evidence seems to be driving an increase in confidence. In addition, we show that this evidence bias in confidence is not valid only for the perceptual decision but is found in value-based choices as well. We propose a computational framework based on signal detection theory [13], where asymmetries in evidence assessment allow the model to capture the goal-relevant evidence effect on confidence.

In previous work, we reported that attention may be playing an important role in generating the asymmetries that could drive the evidence bias (Chapter 3 and [186]). These experiments consider a value-based and perceptual decision-making task both reanalyzed in this chapter. In our third experiment included in this chapter, we show that this effect on confidence is robust and does not rely on spatial attention displacements, hinting at a
more general attentional mechanism. To further support the hypothesis that
attention plays a role in gating the evidence effect in confidence, we analysed
pupil changes in a decision-making auditory task. Pupilometry has been
shown to track attentional processes \cite{374} with pupil variations related to
uncertainty, surprise, and expectancy \cite{375, 376, 377, 378}. In our last exper-
iment, we found a frame-dependent pupil variation, with a reduction in pupil
size when the chosen evidence was in line with the objective of the task.
This could indicate that goals orient attention in subtle ways, generating
expectations towards relevant aspects of the choice, which ultimately impacts
confidence.

Overall, our results highlight that the goals impact the decision process
generating asymmetries in information processing, which is reflected in the
construction of confidence. These changes could be indicative of a system
that adapts to the demands of real situations, more than a simple “bias”.

8.3 Methods

8.3.1 Procedure

8.3.1.1 Experiment 1- Value-based decision

The datasets from experiments 1 and 2 have been employed in a previous
study by our group \cite{186} presented in chapters 3 and 7. At the beginning
of this experiment, participants were asked to report on a scale from £0 to
£3 the maximum they would be willing to pay for each of the 60 snack food
items. They were informed that this bid would give them the opportunity
to purchase a snack at the end of the experiment, using the BDM \cite{182}. The
BDM procedure incentivizes participants to report their true valuation of
the items. Participants were asked to fast for 4 hr previous to the experiment
following an established protocol for this type of experiment \cite{33, 2, 57}.

Following the bid process, participants completed the choice task: in
each trial, they were asked to choose between two snack items, displayed
on-screen in equidistant boxes to the left and right of the centre of the screen
(Figure 8.1A). After each binary choice, participants also rated their subjective
level of confidence in their choice. Pairs were selected using the value ratings
given in the bidding task: using a median split, each item was categorised
as high-value or low-value for the participant; these were then combined to
produce 15 high-value, 15 low-value, and 30 mixed pairs, for a total of 60
pairs tailored to the participant’s preferences. Each pair was presented twice,
inverting the position to have a counterbalanced item presentation.

The key aspect of this experimental setting was that all participants
executed the choice process under two framing conditions: (1) a like frame,
in which participants were asked to select the item that they liked the most,
that is, the snack that they would prefer to eat at the end of the experiment
and (2) a dislike frame in which participants were asked to select the item
that they liked the least, knowing that this is tantamount to choosing the
other item for consumption at the end of the experiment.

After four practice trials, participants performed a total of 6 blocks of 40
trials (240 trials in total). Like and dislike frames were presented in alternate
blocks and the order was counterbalanced across participants (120 trials per
frame). An icon in the top-left corner of the screen (‘thumbs up’ for like and
‘stop sign’ for dislike) reminded participants of the choice they were required
to make; this was also announced by the investigator at the beginning of
every block. The last pair in a block would not be first in the subsequent block.

Eye movements were recorded throughout the choice task and the presentation of food items was gaze-contingent: participants could only see one item at a time depending on which box they looked at. This was done to reduce the risk that participants while gazing at one item, would still look at the other item in their visual periphery.

Once all tasks were completed, one trial was randomly selected from the choice task. The BDM bid value of the preferred item (the chosen one in the like frame and the unchosen one in the dislike frame) was compared with a randomly generated number between £0 and £3. If the bid was higher than the BDM generated value, an amount equivalent to the BDM value was subtracted from their £20 payment and the participant received the food item. If the bid was lower than the generated value, participants were paid £20 for their time and did not receive any snacks. In either case, participants were required to stay in the testing room for an extra hour and were unable to eat any food during this time other than the food bought in the auction. Participants were made aware of the whole procedure before the experiment began.

8.3.1.2 Experiment 2 - Perceptual decision (dot number patch)

Experiment 2 had a design similar to the one implemented in Experiment 1, except that alternatives were visual stimuli instead of food items. In this task, participants had to choose between two circles filled with dots, again in two frames. In the most frame, they had to pick the one with more dots; and the one with fewer dots in the fewest frame. The total number of dots presented in the circles could have three numerosity levels (=50, 80 and 110 dots). For each pair in those three levels, the dot difference between the circles varied in 10 percentage levels (ranging from 2% to 20% with 2% steps). To increase the difficulty of the task, in addition to the target dots (blue-green coloured), distractor dots (orange coloured) were also shown. The number of distractor dots was 80% of that of target dots (40, 64, and 88 for the three numerosity levels, respectively). Pairs were presented twice and counterbalanced for item presentation. After 40 practice trials (20 initial trials with feedback, last 20 without), participants completed 3 blocks of 40 trials in the most frame and the same number in the fewest frame; they faced blocks with alternating frames, with a presentation order counterbalanced across participants. On the top left side of the screen a message indicating ‘Most’ or ‘Fewest’ reminded participants of the current frame. Participants reported their confidence level in making the correct choice at the end of each trial. As in the previous experiment, the presentation of each circle was gaze contingent. Participants received £7.5 for 1 hr in this study.

Experiments 1 and 2 were programmed using Experiment Builder version 2.1.1.40 (SR Research). Although gaze information was gathered in these tasks, it was not considered for the current analyses.

8.3.1.3 Experiment 3 - Perceptual decision (random dot motion)

A new web-based random dot kinematogram (RDK) experiment was designed to support our findings in experiments 1 and 2. In traditional RDK tasks, a cloud of moving dots is displayed on the screen with a certain proportion of dots moving coherently in a single direction while the rest drifts randomly. To keep better control of the evidence in favour and against the
decision, we constrained the dot movement to the left and right directions only. The coherence measure in RDK tasks characterises the proportion of dots moving in the same direction; note that in our particular setup, a coherence of 50% would mean an equal number of dots moving in the left and right directions. Participants were instructed to select the direction of dot movement considering two frames. In a high motion frame, participants selected the direction where most of the dots were moving; and the direction where fewer dots were moving in a low motion frame. The total number of dots presented on the screen was 20, 40, 60, 80, 100, 120, and 140 dots, with an equal number of trials for each category. For each dot numerosity level, 8 coherence levels were used: 51%, 52%, 55%, 60%, 65%, 70%, 80%, and 90%. For each one of these combinations of dot number and coherence, the movement direction was set to left and right. Therefore, each experimental block contained 112 trials with unique dot movement patterns. The full experiment consisted of 2 blocks for each one of the frames (224 trials per frame, 448 trials in total). At the beginning of each block, participants were informed of the type of frame they had to consider for their answers and the screen background during the trials was modified: blue background (hex colour #b5cce9) for high motion frame and red background (hex colour #e9b5b5) for low motion frame. The background colours were selected to keep similar brightness in both frames. The moving dots were presented at the centre of the screen inside an aperture with an ellipse shape (500 x 350 pixels), the dot radius was 2 pixels, the dot life was set to 20 frames and the dot colour was black in both frames. The presentation of the experimental blocks (112 trials) was interleaved between high motion and low motion frames: half of the participants were shown a high motion block first and the other half started with a low motion block. In each trial, the cloud of moving dots was presented until the participants made their choice. Participants reported their selected direction using their keyboards ('a' key for left and 'l' for right) in a self-paced fashion within 5 seconds. If the participant did not answer the trial was considered as missed and not repeated. After each choice participants were asked to report their confidence in their decision using the numerical keys (1 to 9). Before the task, participants were given the instructions and exposed to 32 training trials, with the option of repeating the full training in case some aspect was not clear. Online participant performance was assessed when half of the trials had been completed; if the participants’ accuracy was under 60% the experiment was stopped and the participant was rejected from the study. Participants were reimbursed £6 for participating and an extra £4 for completing the full experiment. All payments were done through the Prolific recruitment platform (https://www.prolific.co/). This experiment was implemented online using the Gorilla research platform (https://gorilla.sc/) and JavaScript code JsPsych (de Leeuw, 2015). RDK plugin for JsPsych [379] was modified to restrain dot movement to only left and right directions. Other parameters of dot motion were kept as the default settings given by the plugin.

8.3.1.4 Experiment 4 - Perceptual decision (auditory clicks)

This new experiment considered a two-alternative forced choice on auditory stimuli [376]. Participants used headphones during the experiment while they observed the screen located at a distance of 1 mt. Participants listened to two streams of auditory clicks of duration 1s, presented simultaneously one to the left and the other to the right ear. Clicks were generated as
periods of 40ms of silence on a white noise background. The number of clicks presented on each stream varied between 1 and 10. Pairs of click audio streams were generated using all the potential combinations, excluding the case in which both audio streams contained the same number of clicks. Participants made choices in two decision frames: in the high click frame, participants selected the side that contained more audio clicks, while in the low click frame participants reported the side with the lower number of clicks. An arrow pointing up or down located at the centre of the screen indicated whether the trial was a high or low click frame. After each choice, participants were shown a scale to indicate their level of confidence in their decision. Frames were presented in a block design with 2 blocks of high clicks and 2 blocks of low clicks trials. Participants completed in total 180 experimental trials (45 trials per block, 90 trials per frame) with short breaks in between. Twenty practice trials (10 for each frame) were presented before the experimental trials. Practice trials included feedback (correct/error). Each stimulus pair was presented twice during the experimental trials in each frame to counterbalance the position. The sequence of frame-blocks was randomly assigned across participants.

Participants were seated with their chin and forehead resting on a head support, to maintain a stable pupil recording during the experiment. At the beginning of each trial, participants needed to fixate on the centre of the screen to initiate the presentation of the stimuli (drift correction was estimated using this fixation). For the choice phase of the trial, participants were instructed to always focus on the centre of the screen. After a 2s period (used to calculate pupil baseline) the pair of sounds was presented for 1s. Participants made their choices without time restrictions using the M and Z keys for the right and left options, respectively. After the choice key was pressed a 3s delay was presented. The arrow cue indicating the frame was presented at the centre of the screen during the entire choice phase of the trial. After the decision, participants could move on the confidence scale using the keyboard (M and Z keys) and select their confidence level by pressing the spacebar. No time limit was imposed on confidence reports.

In this experiment, pupil diameter was sampled at 500 Hz using an EyeLink 1000 eye-tracker system (SR Research). The experiment was programmed using Experiment Builder version 2.3.38 (SR Research). The display resolution was 1024 × 768 pixels. Sound stimuli were created using custom Matlab scripts. The clicks were located randomly inside the 1s duration of the audio stimuli sampled at 44.1kHz.

8.3.2 Participants

8.3.2.1 Experiment 1

Forty volunteers gave their informed consent to take part in this research. Of these, 31 passed the exclusion criteria and were included in the analysis (16 females, 17 males, aged 20–54, mean age of 28.8). One participant was excluded for not using the full value scale when bidding for the snack items. A second participant was excluded given they frequently reported the same bid value. A further four participants were excluded due to reporting the same confidence in most of their choices. Three participants were excluded due to not complying with the instructions of the experiments. In the latter case, one participant's eye-tracking data showed the highest number of blink events and made choices without fixating on any of the items; the other two
did not comply with the frame manipulation. All participants reported not being under the treatment for mental health disorders. To ensure familiarity with the snack items, all the participants in the study had lived in the UK for 1 year or more (an average of 17 years).

8.3.2.2 Experiment 2

Forty healthy volunteers were recruited for the second experiment. Thirty-two participants (22 females, 10 males, aged 19–50, mean age of 26.03) were included in the behavioural and regression analyses. Three participants were excluded for the repetition of the confidence rating across trials. Five participants were removed for not complying with the instructions of the task: four of them had performance close to chance level or did not follow the frame modification, and one participant presented difficulties with eye-tracking.

Please check [186] and Chapter 3 (Methods) for full details on the exclusion criteria for experiments 1 and 2. All participants signed a consent form, and both studies were done following the approval given by the University College London, Division of Psychology and Language Sciences ethics committee.

8.3.2.3 Experiment 3

Experiment 3 was conducted online with participants recruited using the Prolific platform [380]. Participants were screened for age (minimum age 18 years old), not having a diagnosed or ongoing mental health condition and normal (or corrected to normal) vision. Forty healthy participants in total were recruited, but three were rejected online (i.e., they did not complete the experiment) due to low performance. The full experimental data were available for 37 participants (7 females, 29 males, aged 18–19, mean age of 18.18). Of the 37 participants, 3 extra participants were excluded due to low performance once the full experiment was completed. Four participants were excluded due to repetitive confidence ratings (>60% of trials with the same confidence value) and 1 participant was excluded because of not using the full range of confidence (<50% of the full range). After the exclusion, 29 participants were included in the hierarchical regression analysis and model fit (4 females, 24 males, aged 18–19, mean age of 18.20).

Since we excluded a considerable number of participants from the analysis of the online experiment, we repeated the hierarchical regression analysis presented in figure 8.1F including all 37 participants to check the evidence effect on confidence is still present despite the lower performance in those subjects. The main effect (i.e., change in sign of the effect of ΣMotionStr on Confidence between frames) was still present in the analysis of the whole dataset, hinting at the robustness of the finding.

All participants completed an online consent form before starting the experiment. This online experiment was completed following the approval given by the University College London, Division of Psychology and Language Sciences ethics committee. All data were collected during April 2021.

8.3.2.4 Experiment 4

Experiment 4 was conducted as an in-person experiment in an eye-tracking setup. Thirty-eight subjects participated in this study, but 32 participants
were included for analysis (24 females, 8 males, aged 18–52, mean age of 28.8). Exclusion criteria considered:

1. Participants had performance inferior to 65% accuracy.
2. Participants used less than 25% of the choice confidence scales.
3. Participants gave the same confidence rating for more than 50% of their choices.
4. Participants did not comply with the requirements of the experiment (i.e., participants that failed to change the frame).

Six participants were excluded due to having low performance (criteria 1). All participants completed a consent form before starting the experiment. This experiment was completed following the approval given by the University College London, Division of Psychology and Language Sciences ethics committee. All data were collected from April to October 2022.

8.3.3 Data analysis: behavioural data

Behavioural measures during like/dislike, most/fewest, high/low motion, and high/low clicks frames were compared using statistical tests available in SciPy. Sklearn toolbox in Python was used to perform logistic regressions on choice data. All the hierarchical analyses were performed using lme4 package [183] for R integrated into a Jupyter notebook using the rpy2 package (https://rpy2.readthedocs.io/en/latest/). For choice models, we predicted the log odds ratio of selecting the item appearing at the right. Additionally, we predicted confidence using a linear mixed-effects model. Fixed-effects confidence intervals were calculated by multiplying standard errors by 1.96. Predictors were all z-scored at the participant level. Matplotlib/Seaborn packages were used for visualisation.

8.3.4 Bayesian Models

We developed a series of models inspired by standard bidimensional signal detection theory to describe choice and confidence behaviour in our experiments. In this type of model, the alternatives are commonly designated as target and non-target, with two distributions characterising each one of these stimuli [381]. These two distributions have been traditionally modelled as Gaussian distributions with a mean (μ) and variance (σ²). In this standard approach, the mean of the target option is larger than the non-target (μ_{Target} > μ_{Nontarget}) and the variance of both distributions is equal (σ_{Target} = σ_{Nontarget}).

In the typical use of these models, the goal of the task is not modified, which means the target option is always the one with more evidence (e.g., the number of dots, visual contrast, etc.). Since in our experiment the manipulation of the goal was the key aspect of the task, target and non-target options were not fixed categories from the perspective of the amount of evidence. Therefore, in this paper, we refer to the two distributions, target and non-target, as high or low evidence, respectively, with μ_{High} > μ_{Low}. Besides this change in terminology, the central modification in our models was to allow high and low evidence distributions to have different variances, to capture the goal-oriented changes of the Evidence effect on confidence.

In this section, we describe the details of Bayesian graphical models used for the simulations and to fit each one of the four experiments. All model simulations and fits were implemented in Bayesian framework using Python library PyMC3 [184].
Model Simulations

We simulated a Bayesian observer in a generic binary decision experiment (figure 8.2B). The agent is exposed to two options, with different amounts of evidence (right and left evidence, \(E_r\) and \(E_l\) respectively). The objective of the task was for the simulated agent to report the option (left or right) with higher evidence in a \textit{High Evidence} frame and the alternative with lower evidence in a \textit{Low Evidence} frame. For each simulated trial, we generated the side of the item with higher evidence (\text{side}; 0: left; 1: right) from a Bernoulli distribution (High Ev Opt variable in Figure 8.2B). The evidence for each option was extracted from a Gaussian distribution for the right and left possible directions:

\[
\begin{align*}
    \text{side} &\sim \text{Bernoulli}(0.5) \\
    E_r &\sim \mathcal{N}(\text{side}, 1) \\
    E_l &\sim \mathcal{N}(1 - \text{side}, 1)
\end{align*}
\]

The observer also had two internal distributions that characterise their beliefs, a Gaussian distribution to describe high evidence samples and another one for low evidence samples (figure 8.2A), \(E_{\text{high}}\) and \(E_{\text{low}}\), respectively:

\[
\begin{align*}
    E_{\text{high}} &\sim \mathcal{N}(\mu_{\text{high}}, \sigma_{\text{high}}) \\
    E_{\text{low}} &\sim \mathcal{N}(\mu_{\text{low}}, \sigma_{\text{low}})
\end{align*}
\]

Therefore, for each trial, the simulated agent observed the pair of evidence \(<E_r, E_l>\) and determined the likelihood they were generated from a distribution \(<\text{High Evidence, Low Evidence}>\) or \(<\text{Low Evidence, High Evidence}>\). The logarithm of the ratio of these two likelihoods (L) was used as the decision variable for the observer:

\[
L = \log \frac{p(E_r, E_l \mid <\text{HighEv, LowEv}>)}{p(E_r, E_l \mid <\text{LowEv, HighEv}>)} = \log \frac{p(E_r \mid E_{\text{high}})p(E_l \mid E_{\text{low}})}{p(E_r \mid E_{\text{low}})p(E_l \mid E_{\text{high}})}
\]

\[
\begin{align*}
    &= \log \left[ \frac{1}{\sigma_{\text{high}} \sqrt{2\pi}} \exp \left( -\frac{1}{2} \frac{(E_r - \mu_{\text{high}})^2}{\sigma_{\text{high}}^2} \right) \right] + \log \left[ \frac{1}{\sigma_{\text{low}} \sqrt{2\pi}} \exp \left( -\frac{1}{2} \frac{(E_l - \mu_{\text{low}})^2}{\sigma_{\text{low}}^2} \right) \right] - \log \left[ \frac{1}{\sigma_{\text{low}} \sqrt{2\pi}} \exp \left( -\frac{1}{2} \frac{(E_r - \mu_{\text{low}})^2}{\sigma_{\text{low}}^2} \right) \right] - \log \left[ \frac{1}{\sigma_{\text{high}} \sqrt{2\pi}} \exp \left( -\frac{1}{2} \frac{(E_l - \mu_{\text{high}})^2}{\sigma_{\text{high}}^2} \right) \right]
\end{align*}
\]

Note that for all these simulations we assumed \(\mu_{\text{High}} = 1\) and \(\mu_{\text{Low}} = 0\). Depending on the frame that was simulated, the observer selected the option with lower or higher evidence. When high evidence was to be picked, the agent chose the right option if \(L \geq 0\) and the left option when \(L < 0\). In the frame in which low evidence had to be selected, right option was chosen if \(L \leq 0\) and left option when \(L > 0\).

We simulated three different models to study how the generation of confidence is affected by decision parameters. Our main model to probe, the asymmetric variance model (AVM), considers that \(\sigma_{\text{high}}\) and \(\sigma_{\text{low}}\) can vary independently. Crucially, the asymmetry was dependent on the goal of the task, in high evidence frame \(\sigma_{\text{high}} > \sigma_{\text{low}}\) and in low evidence frame \(\sigma_{\text{high}} < \sigma_{\text{low}}\).
\[ \text{confidence} = |L| (2) \]

Simulations were performed considering various combinations of \( \sigma_{\text{High}} \) and \( \sigma_{\text{Low}} \) variances. Both, \( \sigma_{\text{High}} \) and \( \sigma_{\text{Low}} \), were assigned values in a range between \([0.5, 1.9]\) with a step of 0.2. One thousand samples were extracted from each simulated model for each \( \sigma_{\text{High}} \) and \( \sigma_{\text{Low}} \) combination. For the analysis presented for a single noise combination in the Results section, we selected \( \sigma_{\text{High}} = 1.2 \) and \( \sigma_{\text{Low}} = 1 \) to represent the high evidence frame and \( \sigma_{\text{High}} = 1 \) and \( \sigma_{\text{Low}} = 1.2 \), for the low evidence frame. In this case, 2000 samples were extracted from the posterior distribution.

To compare our model with the standard approach used in decision-making studies, we simulated an equal variance model (EVM) with \( \sigma_{\text{High}} = 1 \) and \( \sigma_{\text{Low}} = 1 \). Choice decision variable (equation 1), decision threshold and confidence (equation 2) were calculated in a similar way to the AVM. Simulated evidence was also generated identically to the AVM simulations. Two thousand samples were extracted from the posterior to generate the data for analysis. This model would be commonly associated with a Balance of Evidence model which cannot reproduce the effects generated by evidence (i.e., overweighting chosen evidence and \( \Sigma \text{Evidence} \)) on confidence.

Finally, we simulated a third model, the response-congruent heuristic model \cite{368} which was proposed to characterise evidence biases on confidence. In this model, confidence is generated by discarding completely the evidence supporting the unchosen option. In other words:

\[ \text{confidence} = E_{\text{chosen}} (3) \]

With \( E_{\text{chosen}} \) the evidence of the chosen option. For the HM simulations, the model architecture is identical to EVM, except by the confidence definition. In all the models, simulated choice and confidence were characterised using logistic and linear regressions using sklearn and statsmodels toolboxes in Python. All the predicted and predictors values were normalised (z-scored) before fitting to the linear models.

### 8.3.4.2 Model fitting

We used Markov chain Monte Carlo (MCMC) methods implemented in PyMC3 \cite{184} to sample posterior distributions of the parameters considering the inputs: actual evidence supporting the left and right option (\( EV_{\text{left}} \) and \( EV_{\text{right}} \), respectively), subjects’ confidence reports (\( r \)) and choices (\( c \)). To account for the stochasticity of human responses, we included in the AVM an extra parameter to characterise the noise in the encoding of the options (sampling variance, \( \sigma^2_s \)) \cite{382, 383, 125} and a logit function that transforms the decision variable \( L \) in the choice \( c \) \cite{127, 3}.

\[
\begin{align*}
\sigma_{\text{High}} &\sim \text{Uniform}(0.5, 5) \\
\sigma_{\text{Low}} &\sim \text{Uniform}(0.5, 5) \\
\sigma_s &\sim \text{Uniform}(0.5, 5) \\
E_l &\sim N(EV_{\text{left}}, \sigma_s) \\
E_r &\sim N(EV_{\text{right}}, \sigma_s) \\
\beta &\sim N(0, 31.62)
\end{align*}
\]
The lower bound was set to 0.5 to exclude the convergence of noise parameters to 0, which generated deterministic choices that failed to replicate human behaviour. The parameter controlling the prior of $\beta$ was defined in terms of precision instead of standard deviation (precision $\tau = 0.001 \rightarrow \sigma = 31.62$). Choice was determined from the Bernoulli logit function defined as:

$$c \sim \text{Bernoulli} \left( \frac{1}{1 + \exp(\beta L)} \right)$$

With $L$ the log-likelihood term calculated as in equation 1. To construct confidence in the model we additionally accounted for the effect of empirical reaction time (RT) which was included together with the estimated value of $|L|$ (the magnitude of the log-likelihood term):

$$\beta_{\text{conf}} \sim N(0, 31.62)$$
$$\beta_{\text{RT}} \sim N(0, 31.62)$$

The value of $|L|$ and RT went into a logit function to generate a confidence rating in 0-1 scale. The two additional weighting parameters ($\beta_{\text{conf}}$ and $\beta_{\text{RT}}$) were included in a deterministic expression to account for the contribution of both terms to the final reported confidence:

$$\text{conf} = \frac{1}{1 + \exp(\beta_{\text{conf}} |L| + \beta_{\text{RT}} \text{RT})}$$

$$r \sim N(\text{conf}, 0.025)$$

Following [127], the mapping between model confidence and observed confidence allowed a small degree of imprecision ($\sigma = 0.025$) in subjects’ ratings, roughly equivalent to grouping continuous ratings made on a 0-1 scale into ten bins.

The fitting of the model proceeded in two stages [384]. In the first stage, $\sigma$ and $\beta$ were fitted using choice behaviour from participants. The mean value of the posterior distribution for both parameters were used as fixed parameters in the second stage. The remaining parameters were fitted to participants’ choice and confidence in the second stage. The fitting was performed at a pooled-participant level. $\text{EV}_{\text{left}}$, $\text{EV}_{\text{right}}$, and RT were normalised (z-scored). Confidence $r$ was normalised at the participant level and transformed to the 0-1 range.

In experiment 1, the model used as evidence input into the model ($\text{EV}_{\text{left}}$ and $\text{EV}_{\text{right}}$) the willingness to pay reported by the participants for each item. In experiment 2, $\text{EV}_{\text{left}}$ and $\text{EV}_{\text{right}}$ corresponded to the number of dots presented in the left and right circles, respectively. Since in experiment 3, the evidence was controlled by dot coherence (coh) but also by dot numerosity (we called this summed motion strength in our analysis, $\Sigma \text{MotionStr}$) we slightly modified the way evidence was input into the model. Coherence information for each trial was considered to calculate the proportion of dots moving in each one of the two possible directions: $\text{coh}_{\text{left}} = \text{coh}$, and $\text{coh}_{\text{right}} = 1 - \text{coh}$, when the higher movement of dots was towards the left; and $\text{coh}_{\text{left}} = 1 - \text{coh}$, and $\text{coh}_{\text{right}} = \text{coh}$, when most dots were moving in the right direction. Note that coh could take values 0.51, 0.52, 0.55, 0.60, 0.65, 0.70, 0.80, and 0.90; the proportion of dots moving in the specified direction. Remember in our task, dot movement was fixed to left or right directions, therefore 1-coh describes the proportion of dots moving in the opposite (lower movement) direction.

$$E_{\text{Lo}} \sim N(\text{coh}_{\text{left}}, \sigma)$$
\[ E_{L0} \sim N(\text{coh}_{\text{right}}, \sigma) \]

Over and above the evidence obtained from coherence we included \( \Sigma \text{MotionStr} \) evidence, modulated by a parameter (\( \eta \)).

\[ E_1 = E_{L0} + \eta \Sigma \text{MotionStr}(6) \]
\[ E_r = E_{r0} + \eta \Sigma \text{MotionStr}(7) \]

The remaining parameters and priors in the model used for experiment 3 were identical to experiments 1 and 2. Finally, in experiment 4 \( EV_{\text{left}} \) and \( EV_{\text{right}} \) correspond to the number of auditory clicks presented in the left and right speakers, respectively. Model structure and priors for the parameters in experiment 4 were identical to experiment 1. While \( \sigma_{\text{high}} \) and \( \sigma_{\text{low}} \) were considered free parameters in the fitting process, \( \mu_{\text{High}} \) and \( \mu_{\text{Low}} \) were assigned as \( \sigma_{\text{Bel}} = \sigma_{\text{high}} - \sigma_{\text{low}} \) for all the samples in the trace resulting from the fitting process. We found that although the convergence value was not clear the sign of the parameter was always in line with our hypothesis: \( \sigma_{\text{Bel}} \) was positive in the high evidence frame and negative in the low evidence frame of the experiments. In Figure 8.6, for the comparison of \( \sigma_{\text{high}} \) and \( \sigma_{\text{low}} \) for each model, 1,000 samples were drawn from the posterior. Regression analysis for simulated trials and participants trials was presented.

The number of posterior samples for the analysis depended on the number of trials used to fit the model: experiment 1: 93,000 total samples (1,860 trials x 50); experiment 2: 96,000 total samples (1,920 trials x 50); experiment 3: 161,000 total samples (3,258 trials x 50), experiment 4: 72,000 total samples (1,440 trials x 50).

8.3.5 Data analysis: pupillometry experiment

8.3.5.1 Pupil dilation pre-processing

Eye data were organised using DataViewer 4.2.1 (SR Research) generating reports containing details of the pupil area variations for all the participants. Pupil time series were subsampled to 100Hz. Pre-processing of eye data included the interpolation of eye blinks [376, 377]. The effect of blinks and saccades on pupil response was estimated through deconvolution and removed using linear regression [385]. Each trial was z-scored and baseline corrected subtracting the average pupil area in the 2s before the appearance...
of the sound stimuli. For visualisation, the pupil was epoched around relevant onsets: stimuli presentation and response time. The analysis was completed using the pre-processed pupil data.

8.3.5.2 GLM analysis

We fitted a General Linear Model (GLM) at each time point to track the modulation of the pupil trace by decision evidence. We focused our analysis on the time participants reported their choice, considering a 4s window (2s before and after this event). We separated trials for the High and Low Clicks frames. At a participant level, we extracted the pupil trace around choice time across trials. We estimated a GLM to predict the relative pupil size at each time point within the trials, with parameters (1) chosen evidence, (2) unchosen evidence, (3) reaction time of the trial and (4) the position of the fixation at that time point (expressed as X-Y coordinates). Evidence was considered as the number of clicks presented on the selected and non-selected sides. The latter predictors were included to control for nuisance covariates [377]. Group-level parameters for chosen evidence predicting pupil size are presented in Figure 8.9. Statistical test was performed using a permutation analysis: for each time-sample (400 samples in total) and participant (n = 32) we fitted 50 GLMs from different permutations of the dataset. We used the individual parameters estimated from the permutation to calculate a null distribution of the group-level parameters in both frames. The subtraction of the chosen evidence parameters between High Clicks and Low Clicks frames was calculated to characterise the null distribution and used to estimate the significance of the difference between the frame parameters, i.e., the effect of chosen evidence is higher in the low clicks frame. The p-values for the parameter difference at each time sample were corrected using a false-discovery rate (α = 0.01). A cluster size of at least 6 contiguous time samples was considered for significance [186].

8.4 Results

In this work, we present the re-analysis of two datasets from a previous study by our group [186] together with the findings from two new experiments. In these four experiments, participants had to perform either value-based or perceptual decisions within two task frames and then report their confidence level in each one of their choices.

In the first value-based experiment (n = 31) hungry participants made binary choices between food items in like and dislike frames (Figure 8.1A). In the former, participants indicated the item they preferred to eat; in the latter, the snack they wanted to avoid eating. After each choice participants reported their confidence in the decision using a slider. Before starting the decision task, participants indicated their subjective value rating for each one of the available food items, using a standard incentive compatible Becker-DeGroot-Marschak mechanism (BDM) [182].

In the second perceptual experiment, a different group of participants (n = 32) performed a binary perceptual discrimination task (Figure 8.1C). Two circles with a different number of dots were presented in each trial. During the “most” frame, participants reported the circle with a higher number of dots whereas, in the “fewest” frame, they reported the circle with a lower number of dots. Confidence was reported after each choice, as in Experiment 1.
The third perceptual experiment was completed using a web-based platform (n=29) and consisted of a binary choice task involving decisions about random dot kinetograms (RDK). Participants visualized a single aperture at the centre of the screen where dots were moving in two directions: left or right (Figure 8.1E). Unlike standard RDKs we constrained the dot motion to those two directions to be able to control for the evidence in favour and against the selected option. In the “high motion” frame participants indicated the direction in which a majority of dots were moving, and in the “low motion” frame they selected the direction in which a minority of dots were moving. They reported their confidence level for each motion judgment.

The fourth perceptual experiment used sound evidence as the basis for the choice, while participants’ pupil variation was tracked (n = 32). Participants visualized an invariant screen of constant luminosity while they listened to sounds (click sequences) coming from their left and right speakers (Figure 8.1G). Participants had to indicate the side with the higher number of clicks in a “high clicks” frame, or the option with a lower number of clicks in the “low clicks” frame.

8.4.1 Evidence bias on confidence depends on frame

8.4.1.1 Experiment 1

In this section, we performed regression analysis to show the effects that experimental evidence had on participants’ confidence and the impact that frames have on them. As reported in previous work from our group [186, 57] and others [125], the difference in evidence (as predicted by SDT) but also the overall evidence for both alternatives in a trial (ΣValue), modulated confidence reports. In the like frame (blue), confidence increased when ΣValue was higher (bΣValue = 0.09 ± 0.03, p<0.01) (Figure 8.1B). This effect has been interpreted in different ways in those articles, but here we propose that ΣValue effect is just another way of observing a positive evidence bias [63, 367, 368, 369, 371, 370]. For completeness, in this regression analysis we also found that an increase in the value difference between the items (|ΔValue|, controlling for the difficulty of the trials) had a positive effect on confidence (b|ΔValue| = 0.13±0.12, p<0.001), while increases in reaction time (RT) reduced the level of confidence (bRT = -0.30±0.04, p<0.001). For more details on the PEB effect in this experiment, we unpacked the ΣValue on the contributions of chosen and unchosen evidence components. We fitted a hierarchical linear model for confidence including as separated predictors the evidence supporting the chosen and the unchosen snacks (chosen and unchosen item preference value, respectively). We found the magnitude of the effect of chosen value was significantly higher than unchosen value, indicating an overweighting of the selected alternative on the confidence reports (mean |bChosenVal| = 0.22; mean |bUnchosenVal| = 0.09; t(30) = 4.819; p <0.001). Importantly, while an increase in chosen value boosted confidence (positive effect of the predictor on confidence), an increase in the unchosen value decreased the participant’s confidence level. In other words, the imbalance in the contribution of chosen and unchosen options eventually leads to an overall ΣValue confidence boost. This suggestion will be also supported by our model simulations in the following section.

Our key analysis was to test how the changes in task goals impacted the positive evidence bias. In previous work [186] we show that goal manipulation influences attention and consequently shapes the decision process,
including the generation of confidence. Considering only the dislike frame, we fitted another hierarchical linear model for confidence that reflected a flip of the $\Sigma$Value effect on confidence in the dislike frame ($b_{\Sigma \text{Value}} = -0.15 \pm 0.03$, $p<0.001$; like vs dislike $b_{\Sigma \text{Value}}$: $t(30) = 9.91$; $p < 0.001$). This means that in the dislike frame, an increase in overall evidence reduced the reported confidence in the choice, a result first reported in [186] (Figure 8.1A). As in the like frame, we separated chosen and unchosen values to predict their individual influence on confidence. We found that again the chosen alternative was overweighted in the confidence reports (mean $|b_{\text{ChosenVal}}| = 0.27$; mean $|b_{\text{UnchosenVal}}| = 0.1$; $t(30) = 5.783$; $p < 0.001$), however, in this case an increase in the chosen item value generated a decrease in confidence. The opposite was found for the non-selected option: the higher the value of the
unchosen item, the higher the confidence. Note this change in the sign of
the effect was not found for |ΔValue| and RT (dislike: \( b_{|ΔValue|} = 0.16±0.03, p<0.001; b_{RT} = -0.31±0.04, p<0.001 \))

In conclusion, our results show that the positive evidence effect is not
necessarily driven by “positive” evidence, i.e., higher value of the option.
Instead, we found that it was modulated by how evidence was used to
achieve the goal of the task. In the dislike frame of our task, an increase in
the value of the chosen option decreased participant’s confidence, leading to
a negative evidence bias. Some models to explain PEB have considered that
positive evidence depends on the stimulus, associated with a physical feature,
e.g., stimulus visibility. Therefore, the evidence effect would be driven by a
decision system that is prone to detection operations [368, 372, 373]. More
recent work has suggested biologically inspired models, where neurons
with lower normalisation-tunning are responsible of confidence signal [372].
However, this model is still driven by bottom-up stimulus strength. Our
findings offer another perspective, suggesting that the effect could be driven
by the negative evidence, i.e., the lack of evidence or stimulus, when the
goal of the task requires it. It could be argued that this negative effect
occurs because in the value domain the “detection” operation is not relevant
for the choice. To rule out this possibility, we used a similar experimental
design but in perceptual decisions. It is important to highlight that, to our
knowledge, this experiment shows for the first time that PEB is not exclusive
to perceptual decisions but can be extended to value-based decisions.

8.4.1.2 Experiment 2

We found a similar pattern of behavioural results in our second experiment in
the perceptual domain. In this dot numerosity task, the frame modulated the
effect of \( ΣDots \) on confidence, with a positive impact of the overall evidence
on the most frame confidence, and a negative effect in the fewest frame (most:
\( b_{ΣDots} = 0.04 ± 0.02, p<0.001; \) fewest: \( b_{ΣDots} = -0.17 ± 0.02, p<0.001; \) most vs
fewest \( b_{ΣDots} \): \( t(31) = 14.58, p<0.001, \) Figure 8.1B). Participants’ confidence
overweighted the evidence (number of dots) in favour of the chosen option
vs unchosen in the most (mean \( |b_{ChosenDots}| = 0.49; \) mean \( |b_{UnchosenDots}| = 0.44; \)
\( t(31) = 8.776; p<0.001 \) and fewest frames (mean \( |b_{ChosenDots}| = 0.6; \)
mean \( |b_{UnchosenDots}| = 0.43; \) \( t(31) = 13.271; p<0.001 \)). However, while an
increase in the number of dots in the chosen patch boosted confidence in
the most frame, the same increase of chosen evidence in the fewest frame
reduced the level of confidence. The inverse pattern was observed for the
unchosen alternative. This result shows that the evidence bias on perceptual
confidence is also affected by goal manipulations, similarly to the results
of the value experiment. When less physical evidence was presented (i.e., a
lower number of dots) in the fewest frame, confidence was boosted. As in
experiment 1, the frame did not change the sign of the effect of trial difficulty
and RT on confidence (most: \( b_{|ΔDots|} = 0.08±0.02, p<0.001; b_{RT} = -0.32±0.03, p<0.001; \) fewest:
\( b_{|ΔDots|} = 0.14±0.02, p<0.001; b_{RT} = -0.32±0.05, p<0.001 \))

8.4.1.3 Experiment 3

The third experiment was to test if the effects we observed in our previous
experiments were not exclusive to experimental setups requiring spatial
allocation of attention (i.e., gaze displacement to the right or left on the
screen). In this case, we used a dot motion task, which has been reported to
also elicit PEB effects [63, 367, 339]. Given that previous findings hinted at an important role of attention in the generation of goal-dependent evidence effects [186], we thought that the use of a task that does not require spatial attentional displacement might make this effect disappear. Using motion discrimination in a single aperture allows a subtler probe of the frame effects since supporting and opposing evidence were presented at the same screen location. We constrained motion in this task to only right and left directions, to capture the amount of evidence in favour and against choice.

As in our previous analyses, we fitted a hierarchical linear model predicting confidence controlling for reaction time (RT) and task difficulty. We characterised difficulty using motion coherence, i.e., the proportion of dots moving in the same direction. The evidence on each trial was captured by $\Sigma$MotionStr, overall motion strength, which accounts for the total number of dots presented. We found again a significant interaction of frame with the $\Sigma$Evidence effect on confidence, increasing confidence when more dots were displayed in the high motion frame (high motion: $b_{\Sigma\text{MotionStr}} = 0.06 \pm 0.02, \ p < 0.001$), and when fewer dots were on screen in the low motion frame (low motion: $b_{\Sigma\text{MotionStr}} = -0.09 \pm 0.02, \ p < 0.001$; most vs fewest $b_{\Sigma\text{MotionStr}}$: $t(28) = 8.671, \ p < 0.001$, Figure 8.1C). Separating $\Sigma$MotionStr evidence in each alternative (chosen and unchosen motion direction), we found that confidence overweighted the chosen option over the unchosen in the high motion (mean $|b_{\Sigma\text{ChoMotionStr}}| = 0.32$; mean $|b_{\Sigma\text{UnchoMotionStr}}| = 0.23$; $t(28) = 4.628; \ p < 0.001$) and low motion frames (mean $|b_{\Sigma\text{ChoMotionStr}}| = 0.43$; mean $|b_{\Sigma\text{UnchoMotionStr}}| = 0.27$; $t(28) = 6.812; \ p < 0.001$). The effect of the chosen evidence on confidence was found to be positive in the high motion frame (i.e., more confidence when more dots are moving in the chosen direction) and negative in the low motion frame (i.e. lower confidence when more dots are moving in the chosen direction), replicating experiments 1 and 2. The inverse pattern was observed in the unchosen option. This means that while in the high motion frame confidence was boosted in decisions involving movement of a higher number of dots, in the low motion frame confidence increased when there was motion of fewer dots on the screen. Again, the frame did not change the sign of the effect of trial difficulty and RT on confidence (high motion: $b_{\text{Coherence}} = 0.23 \pm 0.02, \ p < 0.001$; $b_{\text{RT}} = -0.31 \pm 0.03, \ p < 0.001$; low motion: $b_{\text{Coherence}} = 0.35 \pm 0.02, \ p < 0.001$; $b_{\text{RT}} = -0.29 \pm 0.03, \ p < 0.001$). These results show that the goal-dependent evidence effect was robust, even in a design with subtler frame manipulations.

8.4.1.4 Experiment 4

The fourth experiment was an auditory decision experiment. The aim of this task was to further explore the PEB beyond visual tasks, in particular, to focus on attentional control through the modulation of pupil size in a goal-relevant fashion. For more details, please check section 4 below. We fitted a hierarchical linear model predicting confidence, again using evidence ($\Sigma$Clicks, the sum number of clicks presented in left and right speakers), RT and difficulty as predictors. Choice difficulty was captured by the difference in auditory clicks $\Delta$Clicks. We found a significant interaction of frame with the $\Sigma$Clicks effect on confidence, (high clicks: $b_{\Sigma\text{Clicks}} = -0.06 \pm 0.03, \ p = 0.03$; low: $b_{\Sigma\text{Clicks}} = -0.15 \pm 0.02, \ p < 0.001$; high vs low clicks, $b_{\Sigma\text{Clicks}}: t(31) = 5.36, \ p < 0.001$, Figure 8.1H). In this experiment, $\Sigma$Clicks did not generate a significant positive effect in the high clicks frame. We characterised the evidence in favour of chosen and unchosen options using the number of
clicks in each one of the alternatives. In the high clicks frame, we did not find that the chosen evidence was overweighted over unchosen, considering the group estimates (mean $|b_{ChoClicks}| = 0.18$; mean $|b_{UnchoClicks}| = 0.26$; $t(31) = -2.52$; $p < 0.05$). In the low clicks frame we found the asymmetry, with confidence being affected mostly by the chosen evidence (mean $|b_{ChoClicks}| = 0.23$; mean $|b_{UnchoClicks}| = 0.006$; $t(31) = 9.42$; $p < 0.001$). The effect of the chosen evidence on confidence was found to be positive in the high clicks frame (i.e., higher confidence when more clicks are presented in the chosen alternative) and negative in the low clicks frame (i.e. lower confidence when more clicks are presented in the chosen direction), replicating the previous experiments. The unchosen evidence had a negative effect on confidence only in the high clicks frame, we did not find a significant effect of the unchosen evidence on confidence in the low clicks frame.

We attribute the lack of a positive effect of $\sum$Clicks on confidence (and the absence of overweighted chosen evidence) in the high clicks frame to a slightly different design of the auditory task relative to the previous experiments. While in visual tasks participants could sample both options freely until the choice was delivered, in the auditory experiment the presentation of evidence was constrained to one second. After that, free response time was allowed, but without sampling new evidence. It has been reported that restriction on decision time may interfere with the relationship of confidence with other behavioural measures, such as RT [68]. We consider this could be the case for evidence in the positive frame (please see discussion for further comments).

As in previous experiments, we found that the frame did not change the sign of the effect of trial difficulty and RT on confidence (high clicks: $b_{|\Delta Clicks|} = 0.25 \pm 0.03$, $p < 0.001$; $b_{RT} = -0.32 \pm 0.03$, $p < 0.001$; low clicks: $b_{|\Delta Clicks|} = 0.17 \pm 0.02$, $p < 0.001$; $b_{RT} = -0.39 \pm 0.04$, $p < 0.001$). Overall, despite the difference in the high clicks condition described above, we still found that frame modulated the relationship between evidence and confidence, with a higher impact of the low evidence option (the lack of auditory events) in the negative (low clicks) frame.

8.4.2 Simulations of asymmetric variance can generate evidence bias

The previous analysis demonstrates a robust positive evidence effect on confidence which depended on the goal of the task. To understand the algorithmic mechanisms behind this effect we developed a Bayesian graphical model based on signal detection theory (SDT). Here we assume an observer that has certain preliminary knowledge of the statistics of the evidence in the task and generates two relevant categories: high and low evidence. For example, in a value-based task, the observer should have internal beliefs that characterise the expected values for a high-preference alternative and another set of expected values of the low-preference items. In this work, we refer to these as the belief distributions (Figure 8.2A). Therefore, when the observer makes a binary decision, she contrasts the available options with her internal belief distributions and picks the most likely option (e.g., that the left option belongs to the low value distribution and the right option to the high value one vs left option is high value and right option is low value). In practical terms, we model this comparison as a decision variable using a log-likelihood ratio (L): if $L \geq 0$, the right-sided option was chosen, and if $L < 0$, the left-sided option was selected, assuming the observer reports the option with higher evidence. In the negative frame, the model made choices
Figure 8.2: Bayesian observer model – asymmetric variance model (AVM). (A) The model proposed in this work is grounded on SDT. AVM considers an observer with internal belief distributions that categorizes the alternatives as high evidence and low evidence options. These two belief distributions were modelled as Gaussians with parameters \((\mu_{\text{high}}, \sigma_{\text{high}})\) and \((\mu_{\text{low}}, \sigma_{\text{low}})\). This separation into two categories can be applied to the evidence for each one of our experiments. As stated in its name, AVM allows variances to vary independently which replicated the evidence effects on confidence observed in human experiments and its task-goal dependency. In high evidence frames (e.g., like, most, high motion, or high clicks frames) the model predicts \(\sigma_{\text{high}} > \sigma_{\text{low}}\). In the opposite low evidence frames, \(\sigma_{\text{high}} < \sigma_{\text{low}}\) can generate the behavioural effects found in humans. (B) Graphical Bayesian model used for simulations. For details see Methods- Model simulations. High Ev Opt (L/R): for simulations, the direction of high evidence was generated from random samples of a Bernoulli distribution; \(E_r\): observed evidence for the right-side alternative; \(E_l\): observed evidence for the left-side alternative; LLR: log-likelihood ratio term (L).

in the opposite direction: when \(L \geq 0\), the left-sided option was chosen, and if \(L < 0\), the right-sided option was selected (Figure 8.2B).

While the standard assumption in SDT is the equal variance for target and non-target evidence distributions \([13, 106, 368, 127, 386]\) our model allows the distributions to have asymmetric variances. It has been suggested that unequal variance, in particular, higher variance for the target option \(\sigma_{\text{target}}^2/\sigma_{\text{non-target}}^2 > 1\), should be expected in more ecological situations or when a target signal emerges from background noise \([387, 381]\). The novelty of our proposal is that the task we are modelling allows for a change of what constitutes a target distribution depending on the frame: in some cases, the target is the option with higher evidence, and in others, the target is the alternative with lower evidence. Therefore, goal-dependent asymmetries should be reflected in a dynamic change in the variance of these belief distributions: in the high frame we expect \(\sigma_{\text{high}}^2/\sigma_{\text{low}}^2 > 1\) and in the low frame \(\sigma_{\text{low}}^2/\sigma_{\text{high}}^2 > 1\). In the following sections, we show how simulations of the asymmetric variance model (AVM) captured both the evidence bias on confidence and the goal-dependent flip of that effect. We additionally reported simulations of an equal variance model (EVM) that represented the
predictions made by a model with standard SDT assumptions. Finally, we also included an additional model that predicted PEB based on a heuristic definition of confidence (heuristic model, HM, [368]). In this case, the evidence that characterises the chosen option was used in the computation of confidence, while evidence supporting the unchosen option was discarded. See Methods for more details on the simulation models. For each model, we simulated 5000 trials and we fitted logistic and linear models to characterise synthetic choice and confidence behaviour, respectively, in the high and low evidence frames.

8.4.3 Asymmetric variance model

The simulations for the AVM model generated goal-dependent choice behaviour given changes in the decision variable threshold (i.e., the selected options for \( L \geq 0 \) and \( L < 0 \) changed depending on the frame, see details for simulations in Methods – Model simulations). From a logistic regression predicting choice we showed that in the high frame, the option with higher evidence was selected in the simulated choices, while in the low frame the alternative with lower evidence was selected (Figure 8.3A). The step-like feature of the logistic curve in the simulations was due to the deterministic nature of the choice threshold (e.g., if \( L \geq 0 \) right option is chosen, otherwise left option is picked). Confidence was calculated as the magnitude of the likelihood ratio (\(|L|\)), and this step was identical for both simulated goals. In both frames, simulated confidence presented a U-shape relative to the difference in evidence between the alternatives, a standard signature of human confidence signal [67] (Figure 8.3A). From a linear regression predicting confidence from the evidence in support of the chosen and unchosen options, we found the simulations indicated that confidence was mostly affected by the chosen alternatives in both frames, with unchosen options having a lesser effect (high frame: \( \beta_{\text{chosen}} = 0.994 \pm 0.003 \), \( p < 0.001 \), \( \beta_{\text{unchosen}} = -0.597 \pm 0.003 \), \( p < 0.001 \); low frame: \( \beta_{\text{chosen}} = -0.986 \pm 0.004 \), \( p < 0.001 \); \( \beta_{\text{unchosen}} = 0.606 \pm 0.004 \), \( p < 0.001 \)). Note the model also predicted that in a high frame, the influence of chosen evidence on confidence was positive while in a low frame, the effect of the chosen evidence was negative. The opposite pattern was observed for the unchosen evidence (Figure 8.3B). This replicates the findings in humans, reported in the previous section. The AVM also predicts the modulation of confidence by \( \Sigma \text{Evidence} \) and its interaction with the frame (Figure 8.3C). While in the high frame \( \Sigma \text{Evidence} \) had a positive effect on confidence, in the low frame the effect was negative (high frame: \( \beta_{\Sigma \text{Evidence}} = 0.317 \pm 0.003 \), \( p < 0.001 \); low frame: \( \beta_{\Sigma \text{Evidence}} = -0.332 \pm 0.003 \), \( p < 0.001 \)). For an additional visualization of the effects, we plotted confidence over the evidence space, defined by the evidence for left and right options (Figure 8.3D). It was noticeable that lower confidence values were found at the bottom left quadrant in the high frame and at the top right quadrant in the low frame, coincident with a goal-dependent effect on confidence. All these results were obtained using asymmetrical belief variance in high (\( \sigma_{\text{high}} = 1.2 \), \( \sigma_{\text{low}} = 1 \)) and low evidence (\( \sigma_{\text{high}} = 1 \), \( \sigma_{\text{low}} = 1.2 \)) frames. We further corroborated similar effects were observed for various possible combinations of \( \sigma_{\text{high}} \) and \( \sigma_{\text{low}} \), with the constraint that the variance asymmetry was defined in a goal-dependent way (i.e., higher variance for the goal-relevant distribution in the task).
Chapter 8: Decision Goals and Their Impact in Confidence

Figure 8.3: Asymmetric Variance Model (AVM) simulations. Frames in which the model selects high evidence (blue) or low evidence (red) are presented. (A) Simulated choice and confidence. (B) Linear regressions predicting confidence using chosen and unchosen evidence as predictors indicated an overweighting of chosen evidence in both frames. (C) Simulated trials predicted the flip in $\Sigma$Evidence effect on confidence. (D) Confidence in the evidence space, defined by (binned) evidence for left and right options. Displacement of higher confidence levels towards the high evidence values is observed in the high evidence frame. High confidence towards lower values is visualized in the low evidence frame.

8.4.4 Equal variance model

As a reference, we also ran simulations for the model with symmetric variance. The only difference between EVM and AVM is that the variance values for high and low evidence distributions were constrained to be equal in EVM; all the other elements of the model were identical (e.g., $L$ as the decision variable, $|L|$ as confidence). We found this model also generated choices depending on task frames and simulated confidence presented a U-shape (Figure 8.4A). In EVM confidence was equally affected by chosen and unchosen evidence, in both frames (high frame: $\hat{\beta}_{\text{chosen}} = 0.856$, $p < 0.001$, $\hat{\beta}_{\text{unchosen}} = -0.859$, $p < 0.001$; low frame: $\hat{\beta}_{\text{chosen}} = -0.853 \pm 0.004$, $p < 0.001$; $\hat{\beta}_{\text{unchosen}} = -0.865 \pm 0.004$, $p < 0.001$) (Figure 8.4B). Consequently, the model did not predict a $\Sigma$Evidence modulation of confidence in either of the frames (high frame: $\hat{\beta}_{\Sigma\text{Evidence}} = -5.378 \times 10^{-17} \approx 0$, $p = 0.585$, ns; low frame: $\hat{\beta}_{\Sigma\text{Evidence}} = -3.643 \times 10^{-17} \approx 0$, $p = 0.108$, ns) (Figure 8.4C). The bidimensional evidence plot did not present any specific displacement of confidence depending on frames (Figure 8.4D). These results were obtained using a single symmetrical variance ($\sigma_{\text{high}} = 1$, $\sigma_{\text{low}} = 1$), but further simulations for various values of equal variance ($\sigma_{\text{high}} = \sigma_{\text{low}}$) presented similar effects. The EVM model predicts a behaviour in line with a confidence signal generated by the *balance of evidence* rule [368, 386].

8.4.5 Heuristic model

Finally, we simulated trials from a model that considers a metacognitive heuristic in which only evidence supporting the chosen option is used to generate the confidence signal [368, 388, 386]. In our modelling, the heuristic model (HM) shares a similar architecture with the EVM at the choice level...
Figure 8.4: Equal Variance Model (EVM) simulations. Frames in which the model selects high evidence (blue) or low evidence (red) are presented. (A) Simulated choice and confidence. (B) Linear regressions predicting confidence using chosen and unchosen evidence as predictors indicated an equal weight of both streams of evidence. (C) Simulated trials predicted no $\Sigma$Evidence effect on confidence. (D) Confidence in the evidence space, defined by (binned) evidence for left and right options.

which allows replicating the frame-dependent flip in choice behaviour (Figure 8.5A). However, since in this model confidence is calculated differently, generated directly from the chosen evidence, HM predicted an inverted U-shape for confidence in the low frame (Figure 8.5A). This means confidence in the negative frame was lower when one of the alternatives had much lower evidence than the other, which occurs on the extremes of the $\Delta$Evidence axis. In other words, when evidence for the chosen option was very low in the low evidence frame, this model predicts low confidence. Participants did not show this pattern, since they presented higher confidence when lower evidence was selected in the negative frames, as shown in our previous analysis. Note that in EVM and AVM, confidence is extracted from the likelihood ratio, i.e., how likely a decision was correct considering evidence plus statistical information of the belief distributions. The confidence in HM depends exclusively on the chosen evidence (high frame: $\beta_{\text{chosen}} = 1$, $p<0.001$, $\beta_{\text{unchosen}} = -1.926 \times 10^{-16} \approx 0$, $p<0.001$; low frame: $\beta_{\text{chosen}} = 1$, $p<0.001$; $\beta_{\text{unchosen}} = 1.743 \times 10^{-17} \approx 0$, $p<0.001$). Unlike human data, the effect of chosen evidence on confidence was always positive, disregarding the change in the frame (Figure 8.5B). In line with these results, we found the simulations with HM could replicate the modulation of confidence by $\Sigma$Evidence, but not the flip of the effect in the low frame (high frame: $\beta_{\Sigma\text{Evidence}} = 0.813$, $p<0.001$; low frame: $\beta_{\Sigma\text{Evidence}} = 0.811$, $p<0.001$) (Figure 8.5C). The bidimensional evidence plot shows how the heuristic definition is not flexible enough to account for the changes in confidence when the goal of the task shifts (Figure 8.5D). It could be argued that we are presenting an unfair representation of the heuristic model, which could be easily adjusted to capture the confidence effects in the low evidence frames. For example, adapting evidence at the input so the low evidence cases become relevant in the negative frames [186]. This is certainly the case, however, implementing this type of change
means that the main rationale behind the heuristic model would not hold, i.e., stimulus strength is driving confidence computation.

![Figure 8.5: Heuristic Model (HM) simulations. Frames in which the model selects high evidence (blue) or low evidence (red) are presented. (A) Simulated choice and confidence. HM fails to simulate a standard U-shape for confidence in the low evidence frame. (B) Linear regressions predicting confidence using chosen and unchosen evidence as predictors indicated that only chosen option contributes to confidence. (C) Simulated trials predicted modulation of confidence by $\Sigma$Evidence, but not the change (flip) of the effect in the low evidence frame. (D) Confidence in the evidence space, defined by (binned) evidence for left and right options.](image)

8.4.6 Asymmetric variance model fit to human data

In the previous section, we showed that AVM can generate a goal-dependent evidence bias on confidence as observed in human behaviour. In this model, the asymmetry of belief variance is crucial to generate this effect. In the following analysis, we established whether the model fitted to experimental data is sufficient to generate variance asymmetries that follow the same patterns predicted by the simulations. To account for the stochasticity of human responses we included additional parameters in the model (e.g., sampling variance), but the core architecture used in the simulations was maintained for the analysis of all the experiments (see Methods – Model fitting for details). Additionally, the generation of confidence reports in this version of the model also incorporated reaction time (RT) information, such that the fitting process controls for its influence. Model fitting was performed in a Bayesian framework pooling together all the even-numbered trials across participants. Samples from the posterior distribution of fitted parameters were used to generate synthetic choices and confidence (Simulations). Analysis of the participant’s behaviour (Human) were presented as a reference (Figure 8.6). Note that the main objective of this step was to replicate the patterns of results we observed in human participants (e.g., the sign of the effect of $\Sigma$Evidence on confidence), however, we are not making any claim on the magnitude of the effects in simulations compared with the magnitude of human regressions parameters.
8.4 RESULTS

8.4.6.1 Experiment 1

The AVM model fitted to the value-based experiment’s data replicated goal-dependent changes in behaviour. From a logistic model predicting simulated choices we found that in the like frame the high value item was preferentially selected, while in the dislike frame the low value item was picked (logistic function slope, $\hat{\beta}_{\text{like}} = 0.534$, $\hat{\beta}_{\text{dislike}} = -0.92$). Splitting the synthetic trials by confidence levels we found that high confidence choices resulted in better discrimination between item’s value, which is observed from a steeper slope of the logistic curve (low confidence: $\hat{\beta}_{\text{like}} = 0.224$, $\hat{\beta}_{\text{dislike}} = -0.427$; high confidence: $\hat{\beta}_{\text{like}} = 0.735$, $\hat{\beta}_{\text{dislike}} = -1.574$). From a linear regression model predicting confidence, we found the simulated trials also displayed a goal-dependent bias effect (Figure 8.6A, left panel), with a rise in $\Sigma_{\text{Value}}$ predicting an increase in confidence during like frame, and a decrease of confidence in the dislike frame (like: $\hat{\beta}_{\Sigma_{\text{Value}}} = 0.088 \pm 0.003$, $p<0.001$; dislike: $\hat{\beta}_{\Sigma_{\text{Value}}} = -0.277 \pm 0.002$, $p<0.001$) (Figure 8.6A, left panel). Crucially, the model predicted asymmetry in the latent variables $\sigma_{\text{high}}$ and $\sigma_{\text{low}}$ (Figuretest your understanding of the text. What is the main finding of Experiment 1 in the AVM model fit? How was this finding replicated in the model simulations? What does this imply about the relationship between confidence and task performance?
while in the fewest frame the low evidence distribution was found to have the highest variance (mean $\sigma_{\text{low}} = 4.65$; mean $\sigma_{\text{high}} = 4.97$; mean $\Delta \sigma_{\text{bel}} = 0.32$, $p<0.001$) and the low value distribution presented higher variance in the dislike frame (mean $\sigma_{\text{low}} = 3.24$; mean $\sigma_{\text{high}} = 2.43$; mean $\Delta \sigma_{\text{bel}} = -0.81$, $p<0.001$). Note we present in the figure the distribution of $\Delta \sigma_{\text{bel}}$ which indicates that in each one of the simulated trials (1000) a goal-dependent $\Delta \sigma_{\text{bel}}$ asymmetry was observed.

8.4.6.2 Experiment 2

In the perceptual experiment on dot numerosity, the AVM was also successful in capturing participant’s behaviour. The model predicted frame-dependent choice behaviour (logistic function slope, $\beta_{\text{most}} = 1.219$, $\beta_{\text{few}} = -1.644$), including an increase in choice discrimination sensitivity for high confidence trials (low confidence: $\beta_{\text{most}} = 0.49$, $\beta_{\text{few}} = -0.856$; high confidence: $\beta_{\text{most}} = 1.829$, $\beta_{\text{few}} = -2.766$). The simulations for this experiment also captured a goal-dependent evidence bias on confidence (Figure 8.6B, left panel), with a boost of confidence when $\Sigma \text{Dots}$ was higher in the most frame (most: $\beta_{\Sigma \text{Dots}} = 0.124 \pm 0.003$, $p<0.001$) and an increase in confidence when $\Sigma \text{Dots}$ was lower in fewest frame (fewest: $\beta_{\Sigma \text{Dots}} = -0.329 \pm 0.003$, $p<0.001$). In this experiment the latent variables of the AVM, $\sigma_{\text{high}}$ and $\sigma_{\text{low}}$, also presented asymmetries depending on the frame (Figure 8.6B, right panel): in the most frame the distribution characterising higher evidence presented higher variance (mean $\sigma_{\text{low}} = 4.217$; mean $\sigma_{\text{high}} = 4.990$; mean $\Delta \sigma_{\text{bel}} = 0.773$, $p<0.001$), while in the fewest frame the low evidence distribution was found to have the higher variance (mean $\sigma_{\text{low}} = 4.998$; mean $\sigma_{\text{high}} = 3.499$; mean $\Delta \sigma_{\text{bel}} = -1.499$, $p<0.001$). Similarly, to experiment 1, the simulations reported asymmetries of $\Delta \sigma_{\text{bel}}$ distributions in every instance.

8.4.6.3 Experiment 3

We again found a similar pattern of results for the simulated trials using the AVM model fitted to our dot motion experiment. From the logistic regression analysis, we found that synthetic choices selected high motion and low motion directions depending on the trial frame (logistic function slope, $\hat{\beta}_{\text{highMotion}} = 0.733$, $\hat{\beta}_{\text{lowMotion}} = -0.629$) and confidence modulated the precision of the choices (low confidence: $\hat{\beta}_{\text{highMotion}} = 0.376$, $\hat{\beta}_{\text{lowMotion}} = -0.354$; high confidence: $\hat{\beta}_{\text{highMotion}} = 0.930$, $\hat{\beta}_{\text{lowMotion}} = -0.741$). We also found the simulations of confidence displayed the flip of $\Sigma \text{MotionStr}$ depending on the frame (Figure 8.6C, left panel), which indicated that the model captured the goal-dependent evidence bias (high motion: $\hat{\beta}_{\Sigma \text{MotionStr}} = 0.067 \pm 0.002$, $p<0.001$; low motion: $\hat{\beta}_{\Sigma \text{MotionStr}} = -0.113 \pm 0.001$, $p<0.001$). The asymmetry of belief variance was again observed in this experiment (Figure 8.6C, right panel): the high evidence distribution had a higher variance in the high motion frame (mean $\sigma_{\text{low}} = 1.074$; mean $\sigma_{\text{high}} = 1.185$; mean $\Delta \sigma_{\text{bel}} = 0.111$, $p<0.001$) and the low evidence distribution had higher variance in the low motion frame (mean $\sigma_{\text{low}} = 1.946$; mean $\sigma_{\text{high}} = 1.859$; mean $\Delta \sigma_{\text{bel}} = -0.0865$, $p<0.001$).

8.4.6.4 Experiment 4

Finally, we found that the AVM model could capture the behavioural signatures we observed in the sound discrimination experiment. From the logistic
8.4 Results

8.4.7 Attention as the cognitive mechanism

Our AVM model is agnostic of the cognitive process that drives the asymmetry in variance and triggers the effect on confidence. Here we test the hypothesis that selective attention might underpin this process generating an imbalance in information processing. In previous work [186], we reported that visual attention is preferentially allocated towards the alternatives that have higher relevance for the goal of the task. This was found in the original analysis of the data reported here in experiments 1 and 2, e.g., in value-based and perceptual tasks. Furthermore, in that same work, we reported that gaze-weighted accumulator models [35] could capture the overall evidence effect estimating confidence from a balance of evidence [124]. Though, this behaviour was captured only if the attentional information was used by the model. These results, in line with our current SDT-inspired model, hint that goal-directed asymmetries in the integration of evidence for the decision are key for confidence.

In this section, we further investigate this hypothesis, considering the result from our experiment 3: in the dot motion experiment with a single aperture, the goal-relevant evidence bias in confidence was present even when the spatial displacement of attention was not required. This finding hints at a more general attentional process. Pupil variation has been employed extendedly in neuroscience and psychology as a method to study the unfolding of cognitive processes over time [391, 374]. The pupils of the eye are constricted in response to light and dilated when exposed to darkness, however, these variations can also be due to autonomic arousal. The dilation of pupils is controlled by the noradrenergic locus coeruleus (LC), a small brainstem nucleus with an important role in task-related processes [392]. Phasic LC activation has been proposed to facilitate behaviours that...
help to optimize task performance, alerting, orienting and other higher-level processes [374]. Pupillometry in decision-making has revealed increases in pupil size in response to effort, low confidence choices, and higher levels of uncertainty and surprise [375, 376, 377, 378, 394].

In our fourth experiment (n = 32), participants completed an auditory binary choice task in two frames, following a similar design to the ones used in experiments 1-3. In this perceptual experiment, participants were presented with auditory click sequences, separately in the left and right-side speakers, while pupil size variation was tracked [376]. Left and right clicks were presented simultaneously. In the High Clicks frame, participants were instructed to select the side where more clicks were presented; in the Low Clicks frame, the alternative with the lower number of clicks was to be chosen (Figure 8.1G).

In a hierarchical regression (Figure 8.1H), we analysed the effect of evidence on confidence reports in the audio task. We found a negative effect of \( \Sigma \)Clicks on confidence in the High Clicks (\( \beta_{\Sigma \text{Clicks}} = -0.06 \pm 0.03, p<0.05 \)) and Low Clicks frames (\( \beta_{\Sigma \text{Clicks}} = -0.16 \pm 0.02, p<0.001 \)). Although the magnitude of the effect in the High Clicks frame was closer to zero, it was still a significant negative effect. This lack of a positive effect in High Clicks, in contrast to the other experiments, may be due to some particularity in the processing of auditory stimuli, given that experiments 1-3 were performed

---

Figure 8.7: Bayesian graphical model indicating the formalization of the asymmetric variance model (AVM) for model fitting in the experiments. See details in Methods – Model fitting.
on visual information. We performed a separate hierarchical regression analysis using chosen and unchosen evidence as predictors of confidence. In this case, chosen evidence did have a positive effect on confidence in the High Clicks frame, while a negative in the Low Clicks condition (High Clicks frame: $\hat{\beta}_{\text{ChosenClicks}} = 0.18 \pm 0.02$, $t = 6.75$, $p<0.001$; Low Clicks frame: $\hat{\beta}_{\text{ChosenClicks}} = -0.22 \pm 0.03$, $t = 6.58$, $p<0.001$). Despite the discrepancy in the overall evidence effect relative to previous experiments, we still could see a significant difference in the $\Sigma$Evidence effect on confidence, depending on frames ($t(31)=5.364$, $p<0.001$). These findings reassure us that the frame still has an impact on the processing of evidence in the auditory dimension. For completeness, in this regression analysis, we also found that an increase in the clicks difference ($|\Delta\text{Clicks}|$, controlling for the difficulty of the trials) had a positive effect on confidence in both frames (High Clicks: $\hat{\beta}_{|\Delta\text{Clicks}|} = 0.26 \pm 0.02$, $p<0.001$; Low Clicks: $\hat{\beta}_{|\Delta\text{Clicks}|} = 0.18 \pm 0.02$, $p<0.001$), while increases in reaction time (RT) reduced the level of confidence (High Clicks: $\hat{\beta}_{\text{RT}} = -0.32 \pm 0.03$, $p<0.001$; Low Clicks: $\hat{\beta}_{\text{RT}} = -0.39 \pm 0.03$, $p<0.001$).

After these behavioural checks, we analysed the variation of relative pupil traces in the task (please check Methods for the details). We found that after the presentation of the sound stimuli, pupil size increased reaching a peak around choice time, following a similar pattern of results in other decision experiments [376, 377]. Overall, trials reported with low confidence had on average higher pupil size at decision time (average relative pupil area, high confidence = $0.79 \pm 0.07$; low confidence = $0.88 \pm 0.08$, $t(31) = -3.08$, $p<0.01$). Trials of higher difficulty (i.e., lower differences in the number of
clicks between the options) also presented higher pupil sizes relative to easier trials at decision time (average relative pupil area, high $|\Delta \text{Clicks}| = 0.81 \pm 0.07$; low $|\Delta \text{Clicks}| = 0.86 \pm 0.07$, $t(31) = -2.34$, $p<0.05$). Correct trials also had on average lower pupil size during the decision period compared with error trials (average relative pupil area, correct $= 0.82 \pm 0.07$; incorrect $= 0.91 \pm 0.09$, $t(31) = -3.09$, $p<0.01$).

Given that we found difficulty and confidence effects at decision time, we centred our pupil analysis on this period, separately for each frame. Additionally, pre-decision time is important since it should characterise the deliberation stage, where we expect to find asymmetries in the evidence integration. We focused our analysis on the effect of chosen evidence on pupil variation since the results above highlighted the relevance of this factor on confidence. We separated trials using a median split of the number of clicks in the chosen option (Figure 8.9B). We found that when participants’ goal was to choose the higher evidence alternative (High Clicks frame), their pupil area was smaller in trials with a higher number of clicks. Conversely, in the Low Clicks frame, their pupil size was reduced in trials with a lower number of clicks. In other words, we found that the frame affected the way the pupil responded to the selected evidence. To corroborate this finding, we fitted a linear regression model predicting relative pupil size at each time point at the individual participant level. We included chosen and unchosen evidence as predictors, in addition to reaction time and pupil position in the X-Y coordinates [377]. We found that during the pre-decision period, chosen evidence had a negative effect on pupil size in the High Clicks frame, while this effect was positive in the Low Clicks frame (Figure 8.9C). The participants’ regression coefficient was significantly different between frames: $\hat{\beta}_{\text{High Clicks}} = -0.052\pm 0.005$, $\hat{\beta}_{\text{Low Clicks}} = 0.033\pm 0.009$, $\Delta \hat{\beta}$ FDR-corrected permutation test $p<0.01$, cluster size $= 6$, indicated with black line in Figure 8.9C).

This confirms that in the positive frame, an increase in chosen evidence reduces the pupil size, while in the negative frame, the increase in chosen evidence boosts pupil size. We repeated this analysis using $\Sigma \text{Evidence}$ instead of chosen evidence, however, we could not capture a significant difference between the frames.

This result is in line with our AVM model, where goals affect internal beliefs and the assessment of the evidence. Previous studies have shown that increases in pupil size are related to encountering surprising events [393, 394, 395]. In these cognitive experiments, surprise has been operationalized as the difference between the agent’s expectations and the actual observations. In other words, the more certain one had been about an expectation that turned out to be wrong, the higher the surprise. In our experiment, we can consider that the different goals change the expectations that participants have on the objective of their choices. For example, in the positive frame, participants expect to find a High Number of clicks, this will satisfy the objective of the search. Therefore, when they indeed choose an option with a relatively high number of clicks (compared to all the possible trials and clicks possibilities) they are less surprised than when picking the option with a low number of clicks (even when the answer could have been correct). The frame reverses this expectation for the negative frame, which is reflected in the flip in pupil behaviour. Overall, we show that an increase in pupil size reflects the trials where the frame expectations were unmet, given that participants chose alternatives with evidence that was relatively further away
from the goal of the task. The frames cue attentional states in participants that are adequate to gather the relevant information for the task, changing participants’ expectations, which reflects the goal-dependent change of priors in our model.

Figure 8.9: Pupil variations depend on the task goal. A) Auditory decision experimental design. In each trial, sequences of click sounds were presented to left and right-side speakers and participants were instructed to select the option with a higher or lower number of clicks, depending on the frame. Response time was not constrained. They reported their confidence in the decision at the end of each trial. Pupilometry was tracked throughout the whole process. B) Pupil traces by frame. Trials were separated using the median split of the number of clicks in the chosen alternative: the trials with more clicks selected (Top) and less clicks selected (Bottom). For the pre-decision time, while in the High Clicks frame, participants showed lower pupil size in the trials with a higher number of clicks, in the Low Clicks frame the participants had a lower pupil size when they chose the most appropriate evidence for that trial, the option with lower clicks. Lines represent average pupil trace across all participants’ trials. The shaded area indicates standard error. C) We corroborated the pupil results with a linear regression analysis predicting relative pupil area at each time point. The effect of chosen clicks on the pupil is presented in the figure. During the pre-decision time, a significant difference in the effect of chosen clicks was found between High and Low Clicks frames. Lines represent the average effect of the chosen evidence parameter estimated at a participant level. The shaded area indicates the standard error, estimated from the participant’s $\beta$ group sample. The black line indicates a significant difference between frames, estimated from an FDR-corrected permutation test.
Confidence in humans has been traditionally characterised as capturing the accuracy of people’s choices. However, multiple reports have found that the actual confidence signal displays various patterns that deviate from this straightforward interpretation. The sources of variability in confidence have been attributed to random noise or systematic influences, which can impact the entire decision process or just the confidence computation [396]. In this study, we focus on the phenomenon named “positive evidence bias”, or the fact that confidence is affected mainly by the amount of evidence in favour of a choice, unlike objective choice [63, 367, 368, 388]. We show in a series of experiments that this evidence bias is a sophisticated effect that depends not only on the perceptual features of the alternatives but could be driven by an adaptation to integrate evidence that is useful for the internal goals of the observer.

One explanation for this apparently suboptimal strategy in confidence is that in discrimination tasks, where the choice requires comparing two or more options, the decision system operates in a detection-like fashion, i.e., identifying the “physical” presence of the stimulus [368, 372, 373]. The idea behind this is that in natural environments the number of stimulus alternatives is not always certain, meaning that the observer needs to compare against the possibility that the stimulus is absent. This causes a bias in favour of the options with more evidence, which has been captured by heuristic models that discard the information of the unchosen alternatives to compute confidence. However, in most of these studies, the goal of the tasks is to choose the alternatives with higher strength of evidence, e.g., more dots, motion, or light intensity. Thanks to the use of decisions in multiple frames, where participants must choose the option with the most or least evidence, we can disentangle evidence from choice. Our results show that the confidence “bias” is driven by goals, which is particularly intriguing in the case of the “negative” frames, where the absence of evidence in the chosen alternative boosted confidence. In this way, we hint that the evidence bias in confidence is more general, driven by the internal objectives that the agents want to fulfil.

To our knowledge, this work is the first one in expanding the evidence biases in confidence, commonly studied in perceptual decisions, to value-based choices. In value decision it has been suggested the evidence used to make choices is extracted from representations of higher-level concepts, not necessarily relying on low-level sensory features as the perceptual decision does [397, 51, 383, 382]. This further supports the view that the impact of evidence in confidence is not defined purely on the concrete perceptual features or visibility of the stimuli. The observation of similar patterns in the generation of confidence across different decision domains demonstrates the robustness of these effects and it may also hint at general mechanisms underlying the metacognitive assessments, if not to decisions as a whole [398, 399, 400].

In our computational framework, we proposed a Bayesian graphical model grounded on SDT to characterise the confidence behaviour. These types of models, sometimes referred to as an ideal observer [381], consider the agent knows the underlying statistics of the experiment, with internal distributions that characterise expected target signal and noise. SDT models commonly assume equal variance for the target and non-target (Gaussian) distributions [368, 127, 3]. We characterised the target/ non-target distribu-
tions as high and low evidence distributions to capture the binary choices. This means that to make a choice the observer estimates the probability that one of the sampled options “belongs” to the high evidence distribution and the other to the low evidence distribution or vice versa. The model used the magnitude of these probabilities to estimate a confidence signal. We found that including asymmetric variance of the belief distributions in the model, specifically higher variance for the goal-relevant distribution, allowed the prediction of the goal-relevant behaviour in confidence, including the effect of $\Sigma$Evidence and the predominance of chosen evidence. A higher variability of the target beliefs distributions has been observed in ecological decisions [387, 381]. For example, it is expected that the target options contain the desired signal plus noise while the non-target alternative contains only noise. This intuition is clear in standard perceptual experiments where “signal” can be properly quantified (e.g., luminosity, contrast, number of dots) over background noise, but it surprising something similar could be observed in higher-level value decisions and the negative frames of the tasks. That is why our interpretation is that goals change the prior beliefs, inserting asymmetries in the assessment of the evidence. In this case, the increase in the belief variance could be understood as the agent being more lenient in categorizing the evidence in the goal-relevant category. The results of our pupillometry experiment are in line with this interpretation, showing that frame-dependent modulations of the autonomic surprise response hint at a change in participants’ expectations towards the relevant evidence for the goal. It is not surprising that humans adapt their internal representations to become more receptive to the evidence relevant to task demands [401, 275, 99, 230]. The presence of asymmetries in the beliefs, which eventually influence how information is integrated into the decision variable, seems crucial in the determination of confidence.

As previously reported, confidence generated from models akin to a balance of evidence with standard assumptions (i.e., equal belief variance or symmetric integration evidence) cannot capture the biases observed in human confidence, since they represent decision accuracy only [63, 368, 386, 186]. To characterise these biases in confidence reports, models in which agents implement heuristic strategies that assume blindness to disconfirming evidence have been proposed [368, 372, 373]. These models can predict evidence bias towards chosen evidence and, as we show in our simulations, the modulation of confidence by $\Sigma$Evidence. However, we found that the heuristic model, in its native specification, cannot account for confidence behaviour with changes in task frame, fundamentally given that confidence is directly calculated from the strength of evidence. Since the heuristic model was proposed based on observations from perceptual experiments with a single task goal, the connection of evidence and confidence, without any intermediate stage, did not generate any issues. However, adding “opposing” goals, as in our case, hinders the model’s capacity to account for the behaviour, unless it is assumed that evidence itself can vary depending on the context (e.g., agents integrate the lack of dots in the fewest frame decision). It is certainly possible to adapt this heuristic model to incorporate evidence in the negative frames, however, this would be outside the central assumption of the model: the decision system is confusing a discrimination task with a detection task which in turn generates the bias. The other important assumption of the heuristic model is that unchosen evidence is ignored by the confidence computation. Our analysis of value-based and perceptual experiments showed that although confidence indeed overweighted the
evidence of the chosen option, the evidence of the unchosen alternative was not fully discarded. A novel model allows a flexible weighting of unchosen evidence still giving privilege to the chosen evidence [402]. However, must be seen if this new version can explain the experimental evidence produced by shifting in goals.

Our previous findings [186] have shown that asymmetries in evidence accumulation generated by visual attention are key to generating goal-dependent biases in confidence. Along the same lines, other studies have proposed that visual attention could be affecting the evidence accumulation reliability of the visually fixated options in sequential sampling models [349, 350]. Therefore, we propose that attention plays a critical role in the presence of evidence biases in confidence, given that the sampling or assessment of evidence is already oriented towards gathering information for options and features that serve the goal. This attentional focus does not rely entirely on a (visual) spatial displacement of attention, as shown by our dot-motion and auditory tasks. We consider that the increased expectation of goal-relevant evidence indeed could be characterised as an attentional effect, that influences how decision evidence is integrated to make a choice. This strategy makes sense from the perspective of naturalistic decisions where limited cognitive resources and time are available to assess the best option, usually among a multitude of alternatives [186]. Note that according to this view, the evidence bias is born from the decision process itself [396], which is in line with recent reports that confidence is not affected by the noisy encoding of the options [125] and that positive evidence bias does not depend on post-decisional processing [371].

Our pupil study further supports that goals shape internal beliefs and attentional states. We show that lower pupil sizes appeared in the trials where participants chose the options containing evidence aligned with the goal of the task. Previous reports on pupil dilation have linked it to surprise [393, 394, 395], the violation of prior beliefs and conflictual decisions [378, 403]. Considering that participants set their expectations to find the evidence relevant for the frame, these results show that congruent choices elicited less surprise, measured by the pupil size in the time leading to the response. Previous studies have shown differences in participants’ pupil responses at the pre-decision time (e.g., differences between error and correct trials) [169, 376, 377, 378], potentially a reflection of the deliberation process [404]. The relationship between attention and pupil variations has been widely explored [288, 405, 374]. Pupil measures have also been connected to the operation of the noradrenergic system, with implications on how agents engage in exploration or exploitation behaviours [406, 392]. These results confirm the effect of goals in the way evidence is integrated into the decision, in line with the findings on attentional gain, not exclusively linked to visibility or spatial location of the decision options. However, it must be noted that with the current experimental design (i.e., displaying both audio options simultaneously) it is not possible to probe some of the predictions of our model, such as the specific variance increment linked to the goal-relevant alternative. Future studies that disentangle the processing of the alternatives could be useful to explore how these changes in specific priors operate during the deliberation. For example, presenting the alternatives sequentially would allow exploring the reaction to individual options, tracking whether the “surprise” measured from pupil dilation is lower for the goal-relevant alternatives, indicating facilitated processing of this specific information.
We found that all the experiments displayed a similar signature in confidence, with the effect $\Sigma$Evidence modulated by the frame. While $\Sigma$Evidence had a positive influence on confidence in the high evidence frame in all the visual experiments, this effect was negative in the auditory task. We think this variation emerges from the approach we used for the presentation of the evidence: as in previous work [376] participants could listen to the auditory stimuli for a limited time, unlike all our visual tasks where the available time for sampling the alternatives was not constrained. This means that participants may have not adjusted their sampling time for trials with more clicks and the tight presentation of the sequential sound cues may have complicated their identification as single clicks. All these factors could have added extra difficulty to these trials, which translated into a negative bias on the perceived confidence for decisions with an overall high number of clicks. Previous studies have shown that confidence modulation can be affected by time limitations on decisions [68]. For example, while faster response times are associated with higher confidence in free response tasks, there is a monotonically increasing relationship between confidence and response time when the decision time is manipulated [407]. Despite this confound, the goal-relevant evidence bias was robust enough so we could still find a modulation of $\Sigma$Evidence by frame, with the low evidence option having a higher influence on confidence in the low clicks frame.

One of the implications of the evidence overweighting of chosen alternatives is the possible decoupling between task performance and metacognitive sensitivity [367, 368, 372]. For example, it is possible to have two decisions with similar difficulty (e.g., the same difference in dot numbers between two patches) but one of them displaying higher overall evidence (e.g., a higher number of dots in total). In this case, participants report higher confidence in the higher evidence case even when the performance at the decision level may be identical. This means variation of confidence is not aligned with objective changes in decision accuracy, making confidence less informative on the actual performance, i.e., there is a drop in metacognitive sensitivity [368]. In our experiments, we studied confidence reports, but we did not account for further metacognitive measures. Standard studies to measure metacognition maintain a constant performance ($d'$) to avoid the confounding influence of difficulty on confidence. In our design, we used multiple levels of difficulty to cover a wide evidence space. It has been reported [381] that metacognitive sensitivity (meta-$d'$) could increase in observers that assume unequal bidimensional variance (e.g., $\sigma^2_{\text{Target}} / \sigma^2_{\text{Nontarget}} > 1$), however future work needs to probe in more detail the impact in metacognition, especially in scenarios with shifting goals. Our frame design could be advantageous for this type of study, since it does not require changing the presentation of evidence, allowing to keep identical choices and performance, even when shifting the goal of the task.

In conclusion, our work expands the understanding of the origin of the confidence signal, across value and perceptual domains. Thanks to the manipulation of goals in the tasks we have shown that confidence is influenced mainly by the evidence that is most relevant or useful for the behavioural objective of the participants. This is not necessarily dependent on the visibility or perceptual features of the options, since choices for low or absent evidence can also display a boost in confidence. Even more, we show that decisions involving higher-level constructs such as value-based choices also present a goal-relevant bias on confidence. We proposed a simple Bayesian model that includes goal-relevant asymmetries of the
internal beliefs, which captures the idiosyncratic patterns observed in human confidence reports. Our final pupillometry experiment adds further evidence to these asymmetries in the assessment of information by showing that internal expectations could be affected by behavioural goals. In real-life environments there are no choices made in isolation, they are always oriented to fulfil specific behavioural demands. Our results suggest that in this same vein, confidence is more than just tracking the accuracy of the single choice, it could check to which degree the choice is satisfying the demands imposed by the goal. This means that how we integrate information is already tinged by our inner motivations, akin to a confirmation bias occurring before every choice, i.e., we assess information from the angle that helps us to confirm the success of our goal. Given that this evidence effect appears as a consequence of adaptive adjustments to complex goal demands, it may be questionable to refer to it as a “bias”. What seems useless in our simple and artificial experimental scenarios, could be implemented to obtain (from the agent’s perspective) an advantageous position in real-world decisions.
GENERAL DISCUSSION

9.1 SUMMARY

In this work, I have shown how goals and contextual information impact the decision-making process of healthy humans. By integrating the results of multiple experiments across value-based and perceptual modalities, I have described how cognitive processing, attention allocation, brain representations, and ultimately choices and metacognitive reports flexibly adapt to the demands of the task. In this section, I present the results of my studies from the perspective of the wider literature and discuss the implications and directions that this line of research may offer in the future.

9.2 DECISIONS AND THE GOAL PERSPECTIVE

A central question in value-based decision-making is how the brain computes and compares the value of alternatives. To answer this, we need to understand what the relevant representations are and how the brain encodes them. Neuroscience research has uncovered fundamental aspects of this value variable, pushing forward the idea that the brain constructs a common currency in which options of different natures can be mapped onto a single scale. For example, the value of an apple can be compared to that of the Mona Lisa. The standard view of value in neuroeconomics has been associated with the concept of reward, the pleasure that a specific alternative can deliver to an agent. Dissimilar options can be expressed in a common hedonic scale in a process that transforms the item’s multiple features and joins them into a single value measure. The brain seems to construct the value of an object from the combination of many of its fundamental features. For example, the preference value of food items can be traced back to their nutritious components, and the aesthetic preference of paintings and photographs can be extracted from visual features such as contrast, saturation, or concreteness. While information about features has been found at various levels in the brain (e.g., the occipital cortex representing low-level visual features or nutritious features of food in the lateral OFC), specific brain representations for value have been found in prefrontal cortex areas such as the OFC and vmPFC. Indeed, representations of the alternatives involved in value-based decisions have been found in frontal regions, with stronger patterns for the most desirable option. The construction of value can reflect disorders affecting the decision process, e.g., food choices in patients with eating disorders. Food features such as taste, and health can be decoded from the prefrontal regions such as OFC. While in healthy populations food valuation is mainly influenced by taste, health-related features are mostly correlated with subjective choices in eating disorders patients. Even more, a characteristic of the maladaptive decisions...
in these patients is an aversion towards fatty and caloric food. This result is a good example that additional factors besides pleasure and reward are relevant for choices based on subjective preferences.

The findings presented in this work add an extra layer to our understanding of human decisions, pointing out that value also incorporates a dimension indicating how coherent or useful the options are to fulfil the current goals of the agent [84, 275]. This means that there is an online process that is deployed every time a decision is made to adjust the representation, silencing irrelevant and amplifying relevant information. This involves a process of dimensionality reduction to simplify the representations and reduce computational demands on the brain, facilitating the separation of core aspects useful for generalization to other scenarios [401]. The research on cognitive control has historically focused on how task demands impact the processing of information, however with a focus on perceptual decisions [275]. My results across the value-based and perceptual domains (for example in Chapter 3) support the idea that common mechanisms operate to favour the selection of relevant information as a domain general feature of the decision process.

Throughout my thesis, I researched numerous ways in which attention influences information selection in order to create efficient and useful representations that allow a decision to be made with the goal in mind. The interaction of value and attention has been a fruitful line of research showing that in simple choices, people tend to look at the most valuable items in the time leading to the decision [33, 34, 179]. In Chapter 3, I worked on that knowledge to show that a frame modification, even when apparently trivial, displaces attention so that the agent spends more time on the valuation of alternatives that are relevant to the task demand, and not purely focusing on the item with higher hedonic value. This finding is in line with other studies that have demonstrated that information sampling is affected by this type of contextual manipulation [180, 98]. The study presented in Chapter 6, on confirmation bias and attention, also reflects the impact of contextual changes. However, in this study, context change was not a product of an external manipulation (i.e., change in like/dislike cue) but an evolution of participants’ internal state as a consequence of their earlier choices. Confirmation bias was manifested in the exploration of alternatives, favouring the sampling of previously selected options. Indeed, the fact that the bias varies as a function of participants’ confidence in their first choice supports that internal states (context) drive the allocation of attention and the bias during the confirmation decision. This strategy may be advantageous in naturalistic scenarios where we must be sure that the selected option is consistently better, and innumerable alternative options cannot be fully explored, making it relevant to assess the selection already committed [342]. Attention is also known to play a fundamental role in learning, tracking the reliability of the features in the prediction of the desired outcomes [188]. In the study presented in Chapter 4, I showed how consistent reward guides the extraction of the features associated with it, helping the construction of abstractions that can be used to facilitate learning. I showed that activity in the sensory cortices of the brain was modulated by learning, potentially mediated by attention informed by abstractions. It is well known that attention strengthens brain representation and processing of the attended stimuli, e.g., reflected in controlling the adaptive gain of neuronal populations [392, 288] or hemodynamic activity in the brain [413]. The findings in Chapter 5 show how the attention to auditory or visual evidence reorganise representations in
both sensory cortices, in this case showing higher similarity (more consistent) patterns in the area of the brain relevant for the processing of the specific stream of information (i.e., higher similarity in visual cortex patterns when visual information is relevant for choice).

While attention is important to select information, especially when perceptual evidence needs to be identified, the construction of value memory is possibly the main source of information. It has been suggested that in a similar way to the sampling of visual evidence, memories are sampled to construct the value of options [51], a process that could rely on some type of internal attention [295, 276]. The hippocampus is a fundamental area involved with memory and has recently been related to the circuits operating during the value-decision process [52, 293, 53, 294]. My findings in Chapter 4 support the involvement of the hippocampus during value learning representing relevant task features, especially in the case that participants utilize the acquired abstractions. Other reports have found that brain representations, in the hippocampus but also other higher-order areas such as OFC, could be organised by the context and behavioural demands [86, 95, 98, 99]. In Chapter 5, I further highlighted that the hippocampus and OFC patterns seem to be modulated by context. Options that elicit frame-relevant memories seem to generate more stable hippocampal representations, possibly an indication of attentional influence over the internal sampling of information for the valuation process. In Chapter 4, I showed that vmPFC seems to influence the activity of the visual cortex in a top-down manner to organize the representation of goal-relevant features. Overall, the shift of brain representations by the influence of goals could reflect the internal sampling process of relevant information.

Recent work has shown that hierarchically structured models can capture the learning of various rules and abstractions, i.e., organize learned associations according to contexts [414, 415]. In Chapter 4 I showed that reinforcement learning (RL) based on goal-relevant features as states captured participant’s behaviour across learning. This resembles the proposal of hierarchical reinforcement learning [414] where sensory events (e.g., visual cues) are clustered into single latent rule/task sets, driven by common behavioural goals (actions). For example, according to this type of model a train station in Varanasi in India and the underground in Russell Square in London, despite their evident sensory differences, could be represented as similar latent contexts sets since the actions associated are similar, i.e., wait for the train and board when the doors open. A similar principle of context organisation seems to extend to motor learning [416], e.g., a novice cook may learn to chop a potato which creates a latent context, and the same set of skills can be employed to also chop tomatoes, yams, onions, and cheese. Overall, learned latent contexts can be used to rapidly respond to known contexts or to estimate responses in novel scenarios, a fundamental advantage of structuring information this way in the brain. Learning the latent context was part of the idea behind the study on abstractions in Chapter 4 (learning the task rules) but further studies analysing more variable and complex contexts are needed that deeply probe the adaptation process.

These findings of goal modulation of the decision process could be also interpreted as reflecting changes in agents’ expectations, e.g., hoping to find the item with a higher or lower value. These goal expectations are reflected in asymmetries in the way attention is deployed and how evidence is integrated (i.e., the agent is more “receptive” to goal-coherent evidence), as described in Chapter 3. In Chapter 6, the internal expectations of the agents after
having chosen an option modified future assessment of evidence, which was translated into confirmation bias.

In Chapters 7 and 8 I further presented how goals influence the post-decision stage, during the evaluation of confidence. Confidence and metacognition play a fundamental role in the monitoring, and assessment of choices and changes of mind [66, 57, 127, 234]. In Chapter 7, I showed that against standard models, confidence captured more than the direct accuracy of a choice or difficulty of the tasks, but also how close the alternatives are to the objective of the decision (i.e., I showed that overall evidence of the presented options displayed a goal-dependent effect on confidence). In Chapter 8, I expanded this finding, indicating that the so-called positive evidence bias in confidence depends on the agent’s goals and not merely on the perceptual features of the options. This effect means that goal-relevant information tends to be overweighted by confidence, and I show through complementary experiments and computational modelling that the change in goal expectations from the beginning of the decision process (i.e., changes in decision priors) could drive this phenomenon. It may be further suggested that evidence bias on confidence is akin to an “instantaneous” confirmation bias: setting a goal favours the integration of evidence that aligns with the participant’s goals. My results do not support the perspective of confidence ignoring part of the choice information to construct confidence in a post-decision stage [396, 371], but presents confidence being led by evidence that is biased quite early on the pre-decision stage. Overlapping signals encoding value and confidence have been found in various subregions of the prefrontal and cingulate cortices, including vmPFC [56, 66]. This commonality could be associated with a deeper algorithmic role of confidence in adaptive behaviour and the control of internal representations. My research in this thesis and other work from my research group suggest that confidence, usefulness, reliability, or value might be partially overlapping aspects of the same algorithm dedicated to implementing goals and deploying actions that are effective in specific contexts [99, 84].

Monitoring the quality of the representations for each goal is one of the aspects of confidence. Another important role of this signal is to be capable of changing between goals when one of them stops being appropriate. I did not cover this aspect in this thesis, however, it is an important avenue for future research. Confidence could also operate as a way to prioritize the selection between multiple tasks [417, 418]. An insightful experiment on this topic was recently presented by [419]: here monkeys made choices in a perceptual task that considered two rules or goals. These rules were hidden from the animals throughout the experiment and were continuously shifting. The sensory evidence used to make the decisions was also noisy and monkeys received positive or negative feedback after their choices. Therefore, monkeys’ errors had two possible origins: they could be mistaken because their selected rule was incorrect or because their sensory evidence was confusing. This work showed that monkeys (as humans) relied on their confidence to attribute the mistake to themselves or a hidden goal switch, with confidence controlling the accumulation of evidence in favour of a behaviour switch to the alternative strategy. Areas in the dorsomedial prefrontal cortex were suggested as accumulating this switch evidence. This is only one of the approaches to tackle the problem of task shift, an area of research of great interest not only for neuroscience but also for artificial intelligence [8] and psychiatry (please see below for clinical examples).
My findings are important to understand biases in human decisions, from the perspective of psychology and behavioural economics [420, 101]. For example, in the framing bias, the experimenter assumes that decisions can be made in a rational and balanced way, not affected by the way that the information is presented. What I found across these multiple studies was that frames actually change how the brain samples, represent and assign confidence in the alternatives. This means that, for example, the mental representation of a cat when we think of the best pet for our children may not be identical to perhaps the same cat when we think of a companion for our deadliest enemy. My studies attempted to build a platform to further explore the mechanisms behind this possibility.

9.3 Contexts and Goals

Goals and context are two concepts intimately related and sometimes used interchangeably. In general lines, context refers to specific configurations of the (internal or external) environment that influence an agent’s actions. For example, the context of cooking may involve internal (e.g., hunger) and external (e.g., being in a kitchen) states, with specific actions and results (e.g., mixing ingredients, getting the dough, turning on the oven, baking the pizza). Goals are the final objective of the agent’s effort, requiring some level of intentionality from their side (e.g., to eat freshly baked pizza). However, the interplay between context and goals is quite fluid. If the agent decides to pursue a different goal the context immediately changes, even if the external environment is identical (e.g., while you are cooking, suddenly a robber breaks into the kitchen) [99]. It is still an open question what the specific differences between these two categories are, or if this separation of context and goals maps onto different psychological constructs at all [84].

In this thesis, contexts were defined by simple changes in frame explicitly presented as cues and also through implicit changes, such as the variation of the rules to select Pacman characters in a learning task (Chapter 4). In both cases, contexts influenced the states in which agents made choices and the deployment of the strategies and actions. Modification of an agent’s internal state and choice history can also be considered as changes in context, as discussed in Chapter 6, where confirming a previous decision impacted the way people sampled and accumulated evidence in future instances. Context has been extensively studied in Pavlovian conditioning where it plays a key role in learning [421]. In these tasks, context usually is operationalized as the specific experimental chambers or the presentation of specific cues. Multiple functions for context have been proposed such as summation (context enhances the strength of conditioned responses, adding associate value to stimulus) or occasion setting (context-distinctive features gate specific associations between cues and unconditioned stimuli, US, i.e., rewards or punishments). Early in training, context, cues and US may be processed as separate elements but later in training, these three elements tend to be integrated and represented jointly [421].

Context has also been important in the study of memory and the construction of schemas. [95] have proposed 3 pillars of context research: context is 1) stable in time, 2) associatively complex and evolving, and 3) behaviourally relevant. These principles mainly apply to the learning process, such as the study presented in Chapter 4. In this experiment, participants after experiencing the same context (decision rule) steadily during some trials could generate a specific set of representations that efficiently evolved to
satisfy the behavioural demands. Longer studies tracking the learning in more complex scenarios could be informative of the gradual evolution of the contexts, with a focus on the factors driving the construction of contexts. In line with Stark proposal and others [414], my experiments hint at a central role of the behavioural demands on brain processing, which could transfer to context learning.

The perspective of context as described above is much related to cognitive maps for spatial and non-spatial environments, which relies on the multiple functions supported by hippocampal neurons [88, 89]. In this model, the context determines specific relationships between various elements in the task-relevant spaces. A recent study has proposed that mesolimbic dopaminergic firing could also be a mechanism that reflects causal relationships between cues and relevant events in the environments, such as rewards and punishments [422]. In this model, dopamine is not just tracking reward prediction error, as standardly described [142, 143, 133], but it is accounting for causal contingencies of reward in a “retrospective” fashion. In the standard temporal difference reinforcement learning (TDRL) interpretation, dopamine firing encodes cues’ prospective predictions of reward and updates of these expectations occur if outcomes deviate from the expectations. In this new model, each reward or relevant event is primarily identified as meaningful causal targets (MCTs), which triggers a search for potential preliminary cues that could be “causing” the reward, i.e., estimate the contingency between cues and reward. In this way, dopaminergic firing indicates that cues with enough “causal power” become MCTs themselves. Eventually, a reward may have many causal targets that predict its appearance forming a causal cognitive map between cues and reward. It is possible, that for different rewards each one of them will generate alternative mappings with relationships controlled by DA activity (e.g., a causal map to cook a pizza, a causal map to defend ourselves in a robbery). In other words, we could interpret these sets of cues and rewards/punishments as what we have defined as contexts and goals. These maps could operate in various ways, for example, if a particular goal gains preponderance in the agent’s mind (e.g., quench thirst), the search for elements that could help to cause the goal completion will also become targets in the mind (e.g., water). This favours other causal targets in the cognitive map to be highlighted by the agent’s cognitive systems via dopaminergic firing (e.g., finding signs of water such as the sound of a river, finding items that can be used to contain water). Further exploration of this model can be extremely relevant to understand the fluid relationship between contexts, goals and decision-making in animals.

In most of my experiments, as in general in the literature, goals were unambiguous and communicated through simple cues, such as the background colour or arrows on the screen. The process of recognizing contexts and goals in real life is far from simple since we need to extract them from a multisensory and continuous world, many times with scarce or inexistent feedback. Yet, humans and other animals are extremely proficient in performing this context identification operation in a short time [416, 423]. For example, when we leave home in the morning, in a second we can notice that it has been raining combining information from the sight of puddles, the sound of rain and the touch of drops on the skin. At the same time, whenever a novel situation appears, e.g., leaving the airport in a foreign country for the first time, we can immediately initiate the construction of new contexts, e.g., identifying the words to greet in the local language, the important landmarks in the city and the favourite dishes of the new cuisine. Implementing a
goal or learning a context involves a process of dimensionality reduction, where only the relevant aspects of the environment are used to construct the contextual mapping. In the real world, given that perception is generally a multisensory process means that those various streams of information are reorganized into distinct features useful for context identification. Previous studies have shown that multisensory stimuli foster an enhancement in neuronal firing which consequently improves detection, relative to unisensory integration [424]. An interesting avenue for future research is to assess how the construction of contexts and the identification of goals facilitates multisensory integration or vice versa, and how sensory integration predicts the appearance of specific context representations in the brain. As hinted above, the study of the relevant factors for context creation and inference is still scarce and it is certainly a fundamental avenue for future research, e.g., understanding the origin of addiction.

9.4 Psychiatric implications

Many psychiatric conditions have their root in decision-making limitations that could be originated from deficits in context inference, hindering the correct deployment of behavioural strategies, coping mechanisms, and moving away from detrimental goals [425, 353]. For example, patients with anxiety have difficulties when dealing with uncertain scenarios, staying in suboptimal states even when more rewarding alternatives are offered [426]. A similar principle may drive post-traumatic stress disorder (PTSD), where new scenarios can trigger old contexts where terrifying events have been experienced [427]. A signature of this condition is that the mere removal of the threatening factor does not generate the disappearance of the stress-related responses, in a similar way to fear extinction experiments where after the removal of the aversive stimuli (e.g. electric shock) animals still spontaneously display defensive behaviours (e.g., freezing). Proposals from neuroscience research have pointed out that in the case of extinction, it could take place a process of context creation while the old traumatic context remains intact, readily available to be reawakened [135]. It has been suggested that the gradual removal of the aversive stimuli can keep the old context states active, providing non-confirming information to update the predictions contingent on that state [428]. Further study of these mechanisms can be important to understanding memory-modification methods in cognitive behavioural therapy for PTSD.

In substance use disorders, after periods of abstinence, the reappearance of intense craving and eventual relapse can be triggered by cues associated with the context of addiction, e.g., passing by the street where the pub is located [429]. This craving can increase patient motivation to fulfil goals related to the object of addiction, to the detriment of other objectives, such as social responsibilities and even own survival. Indeed, the misidentification of context can have fatal consequences, as observed in a high proportion of drug overdose events reported in cases where drugs were consumed in atypical environments or circumstances, e.g., someone that usually consumes heroin alone at home does it at a party [430, 431, 432]. In this case, it is speculated that anticipatory signals in the usual context allow the body to express conditioned tolerance to handle the high doses of the drug. However, when the drug is consumed in unfamiliar contexts these responses are not deployed, risking intoxication. Further experiments in mice have corroborated the importance of these contextual factors in drug response [433, 434, 435].
As hinted by my studies (Chapters 7 and 8), confidence is central to guide behaviour, including the shift between goals and context identification [419]. Deficits in higher-order processes such as metacognition and introspective assessment of the implementation of strategies have been associated with various psychiatric conditions [358, 73, 362]. This could be another avenue for further insight into the underlying mechanisms of psychiatric conditions. For example, people with obsessive-compulsive disorder (OCD) were found to be capable of tracking the volatility of environments using internal confidence, useful to identify the moment to change the target behaviour. However, their actions did not reflect this information, meaning a disassociation between actions and beliefs, with OCD patients failing to use their models of the environment [358].

9.5 CONCLUSION

The capacity of our brain to instantly adapt to the unlimited contexts and goals we encounter in our life is possibly one of its most unique features. The brain is so proficient that makes it look like a simple and seamless process when there is actually a multitude of steps operating backstage, from the first gaze to the latest confidence. Comprehending the mechanisms that the biological machinery uses to achieve this tremendous task is one of the main targets of neuroscience. Progress in this question can have ramifications that range from understanding and finding solutions to brain disorders that impact the mental health of millions to the potential development of science-fiction-like systems capable of expressing general artificial intelligence. My work in this thesis is a tiny but hopefully constructive contribution to this momentous challenge.


[203] Kenji Kobayashi and Ming Hsu. Common neural code for reward and
information value. *Proceedings of the National Academy of Sciences of the

[204] Shari Liu, Tomer Ullman, Joshua B. Tenenbaum, and Elizabeth S.
Spelke. Ten-month-old infants infer the value of goals from the costs

[205] Daniel McNamee, Antonio Rangel, and John P. O’Doherty. Category-
dependent and category-independent goal-value codes in human ven-

[206] Sabina Gherman and Marios G. Philiastides. Human vmpfc encodes
early signatures of confidence in perceptual decisions. *eLife*, 7(NA):NA–NA,
2018.

[207] Allison D. Shapiro and Scott T. Grafton. Subjective value then

[208] Asaf Gilboa and Hannah Marlatte. Neurobiology of schemas and
2017.

[209] Dharshan Kumaran, Jennifer J. Summerfield, Demis Hassabis, and
Eleanor A. Maguire. Tracking the emergence of conceptual knowledge

[210] Michael L Mack, Bradley C Love, and Alison R Preston. Dynamic up-
dating of hippocampal object representations reflects new conceptual
knowledge. *Proceedings of the National Academy of Sciences*, 113(46):13203–
13208, 2016.

[211] Dorothy Tse, Rosamund F. Langston, Masaki Kakeyama, Ingrid Bethus,
Patrick A. Spooner, Emma R. Wood, Menno P. Witter, and Richard
G. M. Morris. Schemas and memory consolidation. *Science (New York,

[212] Caitlin R. Bowman and Dagmar Zeithamova. Abstract memory rep-
resentations in the ventromedial prefrontal cortex and hippocampus
support concept generalization. *The Journal of neuroscience : the official

[213] Roland G. Benoit, Karl K. Szpunar, and Daniel L. Schacter. Ventro-
medial prefrontal cortex supports affective future simulation by
integrating distributed knowledge. *Proceedings of the National Academy

Rushworth. 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 of the United States

[215] Nicolas W Schuck, Ming Bo Cai, Robert C Wilson, and Yael Niv. Human orbitofrontal cortex represents a cognitive map of state space.


