Learning what's relevant in a largely irrelevant world:

The role of selective attention in learning

Yuan Chang Leong

Princeton University

This thesis was submitted to Princeton University in partial fulfillment of the requirements for the:

Degree of Bachelor of Arts in Psychology

&

Certificate in Quantitative and Computational Neuroscience

Adviser: Yael Niv

April 2013

HONOR PLEDGE

I pledge my honor that this paper represents my own work in accordance with University regulations

Yuan Chang Leong

Acknowledgements

It has been uncomfortable writing this thesis in the first person singular, as much of the current work has been a collaborative effort. I am immensely grateful to my adviser, Professor Yael Niv, for her excellent guidance and teaching. To reuse the words of Santiago Ramón y Cajal, I have "had the good fortune to receive an education in the laboratory of a distinguished scientist, under the beneficial influence of living rules embodied in a learned personality who is inspired by the noble vocation of science combined with teaching", and I am truly grateful for the experience. I would also like to thank everyone in the Niv Lab, especially Reka Daniel, Andra Geana and Robert Wilson, for their invaluable advice and help. I am especially grateful to Angela Radulescu, who besides providing help and advice, spent too many of her weekends at the fMRI scanner with me. This thesis has by no means been an individual endeavor - a big thank you to all those who have helped in one way of another!

Since the thesis marks the end of my time as an undergraduate at Princeton, I would also like to take this opportunity to acknowledge the faculty in the department whom I have had the privilege of working with. They include Professor Uri Hasson and Professor Susan Fiske. Special mention also goes to Professor Philip Johnson-Laird, who convinced me early on that research was a worthy endeavor and that one can make both a life and a living in science. I have been truly privileged to be part of the Princeton Department of Psychology.

Lastly, I would like to thank my parents and two lovely sisters, who have showered me with love, care and concern all these years. This thesis is dedicated to them.

Yuan Chang Leong

Abstract

This thesis presents a framework to study the interaction between attention and learning. The framework proposes that learning processes act on an attentionally-filtered representation of the environment and that the attention filter is dynamically modulated by the outcomes of ongoing learning. These assumptions were tested in a series of experiments in which participants performed a multi-dimensional decision-making task with probabilistic rewards. Choice behavior was analyzed using computational models. Some of these models incorporated information about participants' focus of attention, which was decoded on each trial by combining eye-tracking with pattern classification of functional magnetic resonance imaging (fMRI) data. Model-based analysis of behavior provided preliminary evidence that attention helps determine what we learn about, but we also learn what to attend to.

TABLE OF CONTENTS

Acknowledgments 3

Abstract 4

Chapter 1 Introduction: Making sense of a complex world 7

1.1. Learning..9

1.1.1. Bayesian Learning..9

1.1.2. Reinforcement Learning...10

1.1.3. Learning in Animals and Humans..13

1.1.4. The Curse of Dimensionality...16

1.2. Attention...18

1.2.1. Mechanisms of Attentional Selection...18

1.2.2. A Biased Competition Account of Attention......................................19

1.2.3. Attention as Cognitive Control..20

1.2.4. Fundamental Components of Attention...22

1.3. Learning to Selectively Attend...24

1.3.1. Cognitive Processes are Guided by Internal Representations................24

1.3.2. Structure Learning...25

1.3.3. Serial Hypothesis Testing...25

1.3.4. Value-Driven Attention...26

1.3.5. A Proposed Framework for Attention Learning...................................27

Chapter 2 General Methods 29

2.1. The Faces/Houses/Tools Task..29

2.2. Trial-by-Trial Modeling of Behavior...31

2.2.1. Notation...32

2.2.2. Bayesian Model……………………..…………………………...………..33

2.2.3. Function Approximation Model (Uniform Weighting) ……...………....34

2.2.4. Function Approximation Model with Decay……...……….....................35

2.2.5. Choice Probabilities……...………...36

2.2.6. Model Comparison……...………...37

2.3. Overview of the Current Series of Experiments……...……….................................40

Chapter 3 Experiment 1: Modeling the FHT Task 41

3.1. Method...……………..……………………………..…………………………….41

3.2. Analysis and Results...……………………………..…………………………...43

3.3. Discussion...…………………………..…………………………..................45

Chapter 4 Experiment 2: Decoding the Focus of Attention 48

4.1. Method...……………..……………………………..…………………………….50

4.2. Analysis and Results...……………………………..…………………………...55

4.3. Discussion...…………………………..…………………………..................58

Chapter 5 Experiment 3: Attention Processes in the FHT Task 61

5.1. Method...……………..……………………………..…………………………….62

5.2. Analysis and Results...……………………………..…………………………...64

5.3. Discussion...…………………………..…………………………..................68

Chapter 6 General Discussion 71

6.1. The Attention-Learning Framework ...……………..……………………………… 72

6.2. Lessons Learnt...……………..……………………………..…………………… 75

6.3. The Way Forward…………..…………………………..…..…………………… 76

References 78

Chapter 1

Introduction: Making sense of a complex world

We live in an incredibly stimulating environment. Every waking moment, our senses are bombarded by a plethora of sights, sounds and smells that constantly compete for our attention. Yet, at any instant, only a fraction of the available sensory information is behaviorally relevant. Given our limited cognitive capacity, processing all available information would be computationally expensive, if not humanly impossible (Lavie, 2005; Lennie, 2003). Furthermore, interference from irrelevant information is potentially distracting and could bias us towards inappropriate responses (MacLeod, 1991; Owen et al., 1993). To behave adaptively in this complex world, we need to filter out behaviorally irrelevant information and selectively attend to aspects of the environment that matter most.

The importance of selective attention in regulating cognitive processes is well established (Miller & Cohen, 2001). Individuals build an internal representation of the world that encodes, among other things, information about what is currently task-relevant. When confronted with rich sensory information and competing response possibilities, selective attention facilitates the selection of the appropriate response by filtering the information available in the environment and directing cognitive resources towards the processing of task-relevant information (Desimone & Duncan, 1995; Knudsen, 2007). In a complex multi-dimensional world, however, it may not always be immediately evident what is relevant and what is distraction. One has to build a representation of the world by learning contingencies between different stimuli and outcomes through trial-and-error. This learning interacts with selective attention mechanisms, and is modulated by ongoing feedback from the environment (Wilson & Niv, 2011). While much is known about the operations of selective attention when there is a clear internal representation of what is relevant, less work has been done to investigate how selective attention mechanisms interact with learning processes to build this internal representation in the first place.

In this thesis, I present a theoretical framework that integrates existing theories of attention and learning. The proposed framework is motivated by the perspective that attention and learning are intricately related – attention determines what we learn about, but we also learn what to attend to. Specifically, I make the claim that learning processes act on an *attentionally-filtered representation* of the environment. This representation of the world is not static, but is updated according to outcomes of ongoing decisions. As the model of the world changes, the attention filter is *dynamically adjusted* to direct cognitive processes towards what is currently deemed task-relevant. In other words, learning is constrained by attention, but attention is also modulated by the outcomes of learning. With experience, individuals learn what to attend to via the recurrent loop between attention and learning mechanisms. I refer to this process as *attention learning*. The goal of this thesis is to chart out the mechanisms involved in attention learning.

With that goal in mind, I begin by reviewing the two parallel threads of research that motivate the current work. I start with a discussion of learning, focusing on two different frameworks that have influenced research in the field – Bayesian learning and reinforcement learning. The computational principles described here lay the groundwork for building the learning component in the attention learning framework. Following which, I review the literature on the psychological and neural basis of attention. In particular, I address how attention processes relate to existing models of learning and decision-making. Finally, I review several studies that have begun to investigate the interaction between learning and attention processes.

1.1. Learning

1.1.1. Bayesian Learning

The Bayesian framework has been applied to explain various findings in psychology and neuroscience (for a review, see O'Reilly et al., 2012). Bayesian inference is a form of inductive reasoning in that it involves making rational inferences based on observed information. In this thesis, I am concerned with the application of Bayesian approaches to learning. From a Bayesian perspective, learning is formalized as the updating of existing beliefs given new observations. To be clear, I take beliefs to mean hypotheses about the world. For example, one might hold the belief that it rains frequently in London. One can then update this belief when confronted with data that either supports or contradicts the belief (e.g., a day in London with or without rain).

In a Bayesian framework, beliefs are represented as probability density functions (PDFs). Encoding beliefs as probability distributions rather than single values is advantageous, because it allows Bayesian systems to capture the observer's *uncertainty* about a particular hypothesis. Formally, given new data, a belief can be optimally updated according to Bayes' theorem, which computes the posterior probability of the hypothesis given the observation and the prior belief

$$\underbrace{p(\text{belief} \mid \text{sensory input})}_{\text{posterior}} \propto \underbrace{p(\text{sensory input} \mid \text{belief})}_{\text{likelihood}} \times \underbrace{p(\text{belief})}_{\text{prior}} \qquad \text{(Equation 1.1)}$$

For example, consider the case of a person predicting if it would rain on a particular day in London. The person has two pieces of information: 1) he observes dark clouds in the sky, and knows from experience that dark clouds are often observed on rainy days (i.e. *the likelihood of the observation given the belief*, or p(sensory input | belief)). 2) He also knows that it often rains in London (i.e. *the prior,* or p(belief)). Bayes theorem provides a statistically correct means of combining these two pieces of information to arrive at the probability that it will rain on that day (i.e. *the posterior*, or p(belief | sensory input)).

The beliefs represented by a Bayesian system are not static. Instead, they can be optimally updated based on experience. Let us consider again the previous example. Consider the case where this individual experienced dark cloudy skies with no rain for a month-long stay in London. From a rational perspective, it would make sense to update the belief that it often rains in London. A Bayesian system would do this optimally, according to Bayes theorem (Equation 1.1). The posterior distribution is not merely a prediction about whether it would rain, but can also be used as the new prior of whether it usually rains. This new prior can then be updated using another new observation. In other words, a Bayesian system performs sequential learning by integrating multiple observations over time and updating beliefs accordingly. With an increasing number of observations, the Bayesian system arrives at an increasingly accurate estimate of the belief.

In summary, the Bayesian approach provides a normative framework of learning. Bayesian systems are normative in the sense that they provide a statistically correct means to optimally combine different pieces of information. This allows the learning agent to arrive at beliefs about the world that are statistically correct given observed data.

1.1.2. Reinforcement Learning

Reinforcement learning (RL) provides an alternative framework to study and understand learning. RL methods were first developed within the operations research and artificial learning communities, but have since been applied to study human and animal learning (Balleine, Daw & O'Doherty, 2008). In the RL framework, learning is characterized as the process by which a decision-maker updates the values of choices based on reward or punishment received (Sutton & Barto, 1998). Similar to Bayesian systems, RL models formalize the learning process in precise mathematical terms.

In particular, RL models are concerned with learning the optimal *policy* that maximizes reward in a given task. Here, a policy refers to a mapping of possible *states* to possible *actions*. In other words, a policy determines what actions the agent will take in a given state. States can be thought of as different configurations of stimuli that describe the environment while actions are the different decisions the agent can make in a given state. Given a particular action in a particular state, the agent receives a reward with a certain probability. The action will also move the agent (possibly stochastically) from one state of the world to another. The reward received at each state is drawn from a distribution defined by the *reward function*, while the probability of moving from one state to another is defined by the *transition function*. The goal of the agent is to take a sequence of actions that maximizes the overall, long-term reward received in the task. The agent does this by learning an estimate of the *value function*, or the expected net long-term reward associated with each state or state-action pair. Policies, reward functions, transition functions and value functions are components shared by all RL models. RL models primarily differ on assumptions about how the value function is learned.

One influential RL model is the temporal-difference (TD) model proposed by Sutton and Barto (1990; 1998). The TD model is an extension of the Rescorla-Wagner model of Pavlovian conditioning (Rescorla & Wagner, 1972) that allows an agent to learn about long-term values. Like the Rescorla-Wagner model, the TD model builds on the assumption that learning happens when there is a discrepancy between what was expected and what actually happened. In particular, learning in the TD model is driven by a "surprise signal", or *prediction error*, that quantifies the difference between the expected value of a state and the observed value of a state at a given a time point. Importantly, the value of a state takes into account both immediate rewards as well as potential future rewards. The expected value of future rewards is determined by the expected value of the next observed state, scaled by a discount factor. The discount factor captures the notion that later rewards are worth less than earlier ones. Thus, the prediction error is calculated as follows:

$$\delta_{t+1} = r_{t+1} + \gamma V(S_{t+1}) - V(S_t) \qquad \text{(Equation 1.2)}$$

where δ_{t+1} is the prediction error at time t+1, r_{t+1} is the reward received, γ is the discount factor, $V(S_{t+1})$ is the expected value of the next observed state and $V(S_t)$ is the expected value of the previous state. The prediction error is then used to update the expected value of the current state according to the following equation:

$$V_{new}(S_t) = V_{old}(S_t) + \eta\delta_{t+1} \qquad \text{(Equation 1.3)}$$

where η is a learning rate parameter that determines how much the expected value of the current state is updated on each trial. In a given task, it is assumed that the agent has the opportunity to repeatedly sample the reward probabilities and transition probabilities associated with the different states. Each time the agent is in a particular state, it can incrementally update its expectations about that state based on the observed outcome. With sufficient experience (and suitable setting of learning rates), the estimate of an expectation eventually converges to its true value.

In summary, RL provides an alternative framework within which to analyze and study learning. A RL agent learns from trial and error to associate environment states with expected values that take into account predictions of long-term future consequences. Given these expected values, the agent can then formulate an optimal policy aimed at maximizing rewards and minimizing punishments.

1.1.3. Learning in Animals and Humans

Thus far, I have discussed learning in abstract mathematical terms. The Bayesian and RL frameworks provide normative computational principles for solving the learning problem, but can they be applied to learning in animals and humans? Furthermore, is there evidence that the underlying computations are implemented in the brain? In this section, I review the literature on the behavioral and neural correlates of learning, and discuss how they might relate to Bayesian and RL computations.

Reinforcement Learning and the Reward Prediction Error Hypothesis of Dopamine

In recent years, the study of learning has been heavily influenced by findings and theories concerning the computational role of neuromodulators such as dopamine, acetylcholine and serotonin (Daw & Doya, 2006; Doya, 2008). In particular, the dopaminergic system has generated much interest among researchers. While early theories had posited that dopamine encodes a reward signal (Ettenberg, 1989; Wise et al., 1978; Wise, 1982), later results were inconsistent with this original hypothesis. In one pioneering study, Schultz and colleagues (1993) conducted single-cell recordings of dopamine neurons in monkeys performing a spatial delayed response task. They showed that dopamine firing to primary rewards transferred to a reward-predicting cue over the course of learning. The authors interpret their results as suggesting that dopamine neurons were not responding to rewards per se, but were instead sensitive to the first available cue that predicted reward. Prior to learning, the first predictor of reward would be the delivery of reward. However, after the monkey learned the association between cue and reward, it could predict the delivery of reward based on the appearance of the cue. Hence, if dopamine encodes the expectation of reward rather than reward itself, dopamine neurons would respond at the onset of the reward-predicting cue and not during delivery of reward.

It was not long before researchers recognized the strong resemblance between dopamine firing patterns and learning signals in RL models (Montague et al., 1995; 1996; Schultz et al., 1997). According to the TD model, prediction errors are generated when actual outcomes deviate from expected outcomes. Prior to learning, the largest prediction error would be generated at time of the unexpected reward. However, if the agent can fully predict the reward based on an external cue, a prediction error would no longer be generated at the time of reward. This is reminiscent of the dopamine firing patterns reviewed earlier. Such similarities have prompted researchers to propose the *reward prediction error hypothesis of dopamine*, which suggests that dopamine encodes a learning signal akin to the prediction error term (Montague et al., 1996). Much evidence has since accumulated in favor of this hypothesis. Several studies have shown that the magnitude of the dopaminergic response to the reward-predicting cue scales with the reward probability associated with the cue, as would be predicted by a TD model (Fiorillo et al. 2003; Morris et al. 2004). Furthermore, when dopamine responses were analyzed at the trial-by-trial level, firing patterns on each trial coincided with the prediction error computed by a TD model (Bayer & Glimcher, 2005).

Supporting evidence has also emerged in the literature on human learning. Due to the invasiveness of single-cell recordings, researchers have primarily relied on functional magnetic resonance imaging (fMRI) to probe neural processes associated with reinforcement learning in the human brain. fMRI measures a blood-oxygen-level-dependent (BOLD) signal, which is thought to be a correlate of neural activity. Prediction errors computed by TD models have been found to correlate with the BOLD signal in dorsal and ventral striatal areas (McClure et al., 2003; O'Doherty et al., 2003). The striatum receives strong projections from midbrain dopamine neurons and it is reasonable to speculate that BOLD activity in the striatum might reflect dopaminergic activity. As the neural mechanisms underlying the BOLD signal remain debated, the fMRI studies are, at best, indirect evidence for the reward prediction error hypothesis of dopamine. It should be noted, however, that RL models not only predict neural activity, but also human choice behavior, providing additional support for the use of RL as a framework to study human learning and decision-making (Niv, 2009).

Bayesian Computations in Brain and Behavior

In general, the Bayesian framework has been more often applied to study decision-making rather than learning per se. Bayesian models are commonly used to account for both behavior and patterns of neural activity associated with making choices between different options (Beck et al., 2008; Yang & Shadlen, 2007). Less is known about how Bayesian systems might be involved in learning the values associated with the choices or the sequence of actions that maximize long term rewards. That said, there is evidence that under certain conditions, learning can be better described by Bayesian models than by other theories. For example, Behrens and colleagues (2007) demonstrated that a Bayesian learner predicted choice behavior better than various RL models in a task environment with changing levels of volatility. Volatility was operationalized as the likelihood that reward contingencies in the task would change over time. The Bayesian learner was able to keep track of the levels of volatility in the environment and adaptively modulate the rate at which estimates of reward probability change. This was similar to the strategy employed by the human participants, who weighed information received during volatile conditions more than information received during stable conditions. Furthermore, the authors found that activity in the anterior cingulate cortex (ACC) correlated with the levels of volatility computed by the Bayesian learner, suggesting that the brain might perform computations consistent with a Bayesian approach.

1.1.4. The Curse of Dimensionality

The evidence thus far indicates that existing behavioral and neural data fit nicely with normative computational principles. Without a doubt, the convergence of mathematically precise computational principles with experimental data has revolutionized the study of learning by providing researchers with a framework to infer the underlying mechanisms and computations from behavioral or neural data. However, it should be noted that most learning experiments have been conducted under simplistic laboratory controlled settings where only a few salient cues are associated with particular reward probabilities. While Bayesian learning and RL can account for learning under such controlled conditions, they do less well in explaining learning in complex multidimensional situations reminiscent of real-world situations.

It is a well-known problem in operations research and machine learning that the number of states of a task increases exponentially with increasing number of dimensions on which these states are defined. Bellman (1957) referred to this as the *curse of dimensionality*. A fully Bayesian learning model would maintain a probability distribution for each feature in each dimension. The computational resources required to store and update these representations would increase combinatorially with number of dimensions. As such, it may become computationally intractable to perform the necessary calculations in a high-dimensional space (Mathys et al., 2011). RL models also do not fare well in high-dimensional environments. RL algorithms assign values to states (or state-action pairs). As the number of possible states increases, the amount of experience required to arrive at an approximately correct value of each state also increases. Given this limitation, RL models would be highly inefficient in our multi-dimensional world.

Despite the limitations of these computational models, both animals and humans are able solve complicated learning problems given limited experience. As such, it is apparent that current computational models have grossly underestimated the learning capabilities of biological agents. One proposed solution has been to combine Bayesian and RL approaches. Specifically, a Bayesian system is thought to learn the task structure (i.e. the space of possible states) that will be acted upon by RL processes (Gershman, Niv & Cohen, 2010; Jacobs & Kruschke, 2010; Wunderlich et al. 2011). Such Bayesian-RL hybrids have been shown to lead to the optimal solution given a reasonable number of trials. While some studies have demonstrated that Bayesian-RL hybrids predict behavioral data better than fully RL or fully Bayesian models (Gershman et al., 2010; Wunderlich et al. 2011), a recent study suggests that participants' behavior can be better explained by a simpler strategy of serial hypothesis testing via selective attention to each feature in the dimension space (Wilson & Niv, 2012). Employing selective attention as a strategy to navigate the complex multidimensional world may represent a trade-off between optimal learning and computational demands. This perspective is consistent with contemporary theories of attention as a cognitive control mechanism that facilitates information processing by filtering behaviorally relevant information from behaviorally irrelevant information (Miller & Cohen, 2001). It is on that note that I turn to the second part of my literature review.

1.2. Attention

The functional role of attention in regulating cognitive processes has been known since the dawn of experimental psychology (James, 1890). Since then, much work has been done to uncover the psychological and neural underpinnings of attention. Notable findings include evidence of our limited attention capacity (Lavie & Tsal, 1994; Rees et al., 1997), that attention modulates behavioral (Posner, 1980; Treisman & Gelade, 1980) and neuronal (Moran & Desimone, 1985; McAdams & Maunsell, 1999) responses, and that attention mechanisms are modulated by both top-down and bottom-up processes (Cheal & Lyon, 1991; Kastner et al., 1998). In this section, I review the relevant literature on the psychological and neural basis of attention, with an emphasis on different models of attention. These models provide the necessary theoretical framework within which we can interpret existing findings. I also focus my discussion on the visual domain, as much more is known about visual selective attention relative to attention in other modalities.

1.2.1. Mechanisms of Attentional Selection

Attention allows for the selection of behaviorally relevant information while filtering out behaviorally irrelevant information. One central theme in attention research is to chart out the different levels of processing at which selection occurs. Attention researchers have identified four main types of attentional selection: space-based attentional selection (Kastner et al. 1998; McAdams & Maunsell, 1999; Moran & Desimone, 1985), feature-based attentional selection (Treue & Martinez-Trujillo, 1999; Maunsell & Treue, 2006), object-based attentional selection (Duncan, 1984; O'Craven et al., 1999) and object category-based attentional selection (Peelen et al., 2009; Seidl et al., 2012). Different experimental paradigms engage different types of attentional selection. For example, space-based attentional selection is often studied using spatial cueing paradigms while feature-based attentional selection is studied with feature-based search.

Regardless of the level at which attentional selection is operating, there are several commonalities in its modulatory effects on neural activity. Firstly, attention enhances the neuronal response in brain regions selective for the attended information (e.g., a feature, object, object category or location in space). Secondly, attention suppresses the neuronal response to unattended information. In addition, attention has been shown to increase baseline neural activity of neurons with receptive fields within the attended location, even in the absence of visual stimulation (Kastner et al., 1999; Luck et al., 1997). However, experiments by McMains and colleagues (2007) suggest that the attentional modulation of baseline activity might be specific to spatial attention, as attending to a particular feature did not increase the baseline activity of feature-selective neurons. Their results indicate that there might be inherent differences between the mechanisms underlying different levels of attentional selection.

It should be noted that while different levels of attentional selection are often studied separately, the distinctions between them are not necessarily behaviorally relevant. One can easily imagine the case where an observer attends to a particular feature of a specific object from a certain category located at a particular location in space. In real life, these attentional selection mechanisms often operate in parallel to modulate information processing.

1.2.2. A Biased Competition Account of Attention

According to the biased competition model proposed by Desimone and Duncan (1995), the modulatory effects of attention can be best understood in the context of competitive interactions between multiple stimuli in the visual field. The model assumes that processing capacity of visual information is limited and that different objects in the visual field compete with one another for neuronal representation. Attention operates by biasing competition in favor of relevant stimuli. The bias is driven by both bottom-up processes such as attention to salient stimulus properties (Reynolds & Desimone 2001; Kastner & Beck, 2005), and top-down cognitive processes such as directed attention (Kastner et al., 1998, Reynolds et al., 1999). While the biased competition model was first proposed as a theory of spatial attention, it has since been expanded to include other forms of attentional selection (Desimone, 1998).

Much evidence has accumulated in favor of the biased competition model. Both monkey electrophysiology and human fMRI studies find that neuronal response to a stimulus is suppressed by the presence of nearby distractors (Reynolds et al., 1999; Kastner et al., 1998; Kastner & Beck, 2005). This suppression has been interpreted as the result of competitive interactions between different objects. Attention processes resolve this competition by biasing processing resources in favor of attended stimuli. When attention is directed towards a stimulus, it counteracts the suppression induced by nearby distractors. The neuronal response to the stimulus is as strong as when the stimulus is presented alone. Interestingly, attention has no effect on neural activity when the stimulus is presented without distractors. That is, the response to a stimulus presented alone is not different when attention is directed towards it than when attention is directed away from it. Such results argue against earlier theories that viewed attention as a mental spotlight that enhances the processing of stimuli within the spotlight. Instead, it is consistent with the view that attention acts on competitive interactions and that modulatory effects would only be observed in the presence of competition.

1.2.3. Attention as Cognitive Control

Miller and Cohen (2001) incorporated a similar biased competition framework into their model of cognitive control. This work was highly important and influential because it bridged attention, which had until then been primarily studied as a perceptual process, with executive function. According to the model, the prefrontal cortex (PFC) is an important source of the top-down biasing signals driving attention mechanisms. In particular, the PFC actively maintains a representation of what is currently task relevant. This representation is maintained as a sustained pattern of neural activity that guides activity flow along task-relevant pathways in other parts of the brain. In the context of visual attention, PFC activity biases competition in favor of task-relevant visual input. Within this framework, attention is a cognitive control mechanism that adaptively modulates information processing based on current task demands.

This model is supported by the wealth of evidence implicating the PFC in encoding and actively maintaining internal representations. Depending on the task, these representations could be attentional templates, perceptual categories, or abstract behavioral rules (Duncan, 2001; Freedman et al., 2001; Miller et al., 2002; Wallis et al., 2001; Woolgar et al., 2011). When an animal or person performs a task, the PFC exhibits persistent activity specific to the task demands (Brass & von Cramon, 2004; Curtis & D'Esposito, 2003). This persistent activity then directs attention towards task-relevant information (Miller & Cohen, 2001). It should be noted that the PFC does not work in isolation in modulating cognitive control. Instead, it interacts with posterior parietal regions to allocate top-down attentional resources (Brass & von Cramon, 2004; Corbetta, 1998; Fassbender et al., 2006). In particular, it has been argued that since the PFC represents information in abstract terms, the information has to be transformed into a reference frame suitable for space-specific or response-specific bias signals. The posterior parietal cortex (PPC) is thought to be involved in such transformations (Anderson & Buneo, 2002).

1.2.4. Fundamental Components of Attention

Recently, Eric Knudsen (2007) proposed a conceptual framework of attention that draws heavily from the earlier models of Desimone & Duncan (1995) and Miller & Cohen (2001). The goal of this conceptual framework was to present a way of thinking about attention that facilitates the analysis of attention in terms of the underlying neurobiological mechanisms. Knudsen's framework is useful in two ways. Firstly, it breaks down attention into *fundamental components*, or distinct functional processes that make unique and essential contributions to attention. By identifying the key functional components of attention, we can study attention in terms of basic neural mechanisms that might be shared with other cognitive processes. Secondly, it frames the control of attention in the context of a *recurrent loop* between different component processes. According to this perspective, attention is not controlled by a one-off signal from a mental switch, but is continuously modulated by the processing of incoming information.

The framework proposes that four processes are fundamental to attention: bottom-up salience filters, competitive selection, working memory and top-down sensitivity control. Salience filters automatically enhance responses to stimuli with salient properties. For example, stimuli that "pop-out" from the visual scene because they are distinct from the surrounding context, or because they appear infrequently, are more likely to be attended to. Sensory information about the world passes through the salience filters and competes for access to cognitive resources. This competition is further biased by the agent's internal representation of the world, which among other things, encodes what is currently task relevant. Information that wins the competitive selection process then gains access to working memory. Here, working memory is analogous to the PFC's active maintenance of information as proposed by Miller and Cohen (2001), and is responsible for directing top-down biasing signals to further enhance the sensitivity of representations currently held in working memory. In other words, sensory information competes for representation in working memory. The information with the greatest signal strength controls working memory and also directs top-down bias signals that modulate the signal strength of ascending representations. As such, the voluntary control of attention can be thought of as a recurrent loop involving working memory, top-down sensitivity control and competitive selection. The idea of a recurrent loop is an important one, because it allows for feedback processes to act on each of the components. Interestingly, Knudsen did not incorporate learning processes in his framework of attention. However, evidence suggests that attention biases are modulated by learning and feedback (Duncan, 2001; Gottlieb, 2011). This interaction between attention and learning mechanisms is the focus of the final part of my literature review.

1.3. Learning to Selectively Attend

In the previous sections, I have reviewed learning and attention as independents lines of research. However, the two cognitive processes are intricately related and often operate in tandem in any given behavioral task. In this section, I discuss the relationship between learning and attention.

1.3.1. Cognitive Processes are Guided by Internal Representations

As reviewed earlier, the real world is incredibly complex and multidimensional. Learning on all possible dimensions is potentially intractable. For learning to be efficient, we should only learn on the dimensions of the environment that are currently relevant. Selective attention has been proposed as the mechanism by which learning is constrained (Gershman et al. 2010, Wilson & Niv, 2011). By attending to the relevant dimensions, only relevant information will gain access to working memory (Knudsen, 2007; Miller & Cohen, 2001). As working memory is responsible for the evaluation and manipulation of information, attending to the relevant information would constrain learning to representations that are currently active.

An accurate model of the task environment is crucial to the coordination between learning and attention. The agent needs to first know what aspects of the environment are important before it can direct cognitive resources towards the relevant information. In other words, the agent needs to build an internal representation of the world to guide its cognitive processes. While much is known about how learning and attention operate on this representation, less is known about how this representation is learned in the first place. Moreover, in real life situations, where the environment is constantly changing, what is relevant right now might not be relevant later. As such, it would be adaptive to be able to update one’s internal representation based on ongoing feedback.

1.3.2. Structure Learning

As briefly outlined earlier (see section 1.1.4.), structure learning by a Bayesian system has been proposed as one of the mechanisms involved in building and updating a representation of the world. Specifically, a Bayesian system can compute a probability distribution about what is currently relevant based on observed data. Attention and learning are then focused only on the relevant aspects of the environment. In a related line of work, several investigators have applied a Bayesian framework to explain attentional selection (Dayan, Kakade & Montague, 2000; Dayan, Yu & Cohen, 2009; Gershman et al., 2010). In particular, Dayan and colleagues (2000) demonstrated that selective attention to particular stimuli can arise naturally from statistical inference of the association between stimuli and reward. Gershman and colleagues (2010) developed a similar model to explain attention to dimensions. The important point here is that Bayesian learning provides a method to compute a statistical correct representation of the world that allows for optimal allocation of attention.

1.3.3. Serial Hypothesis Testing

Recently, Wilson and Niv (2011) demonstrated that human choice behavior on a multidimensional decision-making task was better described by a suboptimal strategy based on serial hypothesis testing of the relevance of each feature of the environment rather than by the normative strategy based on Bayesian inference over all features at once. They formalized this serial hypothesis testing strategy as a selective attention model that assumed participants attended to each feature in turn and tested if the feature predicted reward. When sufficient evidence accumulated indicating that the feature did not predict reward, the model switchded its focus of attention randomly to another feature. The authors interpreted their results to suggest that while Bayesian inference is optimal, it might be too computationally demanding for the brain. Instead, participants rely on an alternative strategy that is computationally efficient but statically suboptimal.

Wilson and Niv acknowledged that there are many other possible variants of the selective attention model that might better predict participants' choice behavior. In particular, they commented that participants are unlikely to randomly select the next focus of attention, as had been assumed by their current model. As such, it would make sense to extend the existing model by incorporating a decision rule to determine the next focus of attention.

1.3.4. Value-Driven Attention

One possible strategy is to preferentially allocate attention to stimuli with high learned value (i.e., stimuli that have acquired value by association with rewards). There is accumulating behavioral and neural evidence suggesting that animals and humans do indeed direct attention to stimuli that have been reliably paired with rewards (Anderson, Laurent & Yantis, 2011a; 2011b; Hickey et al. 2010; Peck et al., 2009; for a review see Gottlieb, 2012). For example, Anderson and colleagues (2011a) demonstrated that visual search for a target is slowed by the presence of a non-salient, task-irrelevant distractor item that has been previously associated with monetary reward in a separate training session. In a follow-up experiment, they showed that a distractor associated with a larger reward slowed visual search more than an equally salient distractor associated with a smaller reward (Anderson et al., 2011b). Their results suggest that attention is biased towards stimuli that have acquired value via associative learning. Furthermore, the magnitude of attention bias scales with the learned value of the stimulus.

While little is known about the neural basis of value-driven selective attention in humans, Peck et al. (2009) provided neural evidence of related mechanisms in the monkey brain. In their task, monkeys were trained to saccade to a target in response to a visual cue. Each trial began with 50% prior probability of reward. However, the visual cue signaled with certainty whether the trial will end with a reward or no reward. To progress to the next trial, monkeys had to saccade to a target that appeared after the disappearance of the visual cue. The target was located randomly either at the same or opposite location of the cue. To maximize reward in this task, monkeys had to saccade to the target as quickly as possible regardless of the type of cue. However, the authors found that behavioral performance was facilitated when the target appeared on the same side as a cue associated with reward, but impaired when the target appeared on the opposite side. Importantly, opposite results were observed when visual cue was associated with no reward. As the facilitation and interference effects were spatially specific, they were likely to be brought about by a spatial attentional bias rather than global changes in arousal or motivation. The authors interpreted their results as suggesting that attention is automatically biased towards stimuli that have been associated with rewards. They also found that the attentional effects correlated with sustained activity in the lateral interparietal area (LIP), an area commonly implicated in attention control. In her review of the experiment, Gottlieb (2012) speculated that the value-driven orienting of attention might arise through modulation of LIP activity by dopamine prediction error signals.

1.3.5. A Proposed Framework for Attention Learning

This thesis aims to present a theoretical framework for thinking about the relationship between attention and learning that integrates the threads of research reviewed thus far. The proposed *attention learning* framework builds on the findings of Wilson & Niv (2011), but incorporates a value-driven component to attention allocation. In this section, I lay out the two main assumptions of the framework and describe testable predictions that follow from these assumptions.

Firstly, the framework assumes that learning mechanisms act on an attentionally-filtered representation of the world (*Assumption 1*). Attention facilitates learning by biasing processing resources towards relevant information. Assumption 1 would predict that a choice model incorporating participants' focus of attention would account for learning behavior better than a choice model that does not incorporate participants' focus of attention (*Prediction 1*). Importantly, participants can attend to one or more dimensions at the same time. The model assumes that this attentional focus is dynamically adjusted according to the outcomes of ongoing decisions (*Assumption 2)*. Specifically, attention is preferentially allocated to stimuli that consistently predict reward and diverted away from stimuli that do not consistently predict reward. Assumption 2 would predict that the attended stimuli would also be stimuli with high learned value (*Prediction 2*).

In the following chapters, I present the results from a series of new experiments that tested these predictions. In these experiments, participants played a multi-dimensional probabilistic decision-making task where only one dimension was relevant for predicting reward. Their behavior was then analyzed using computational models. Each model made different assumptions about participants' strategy in solving the task. Based on how well each model fits behavioral data, I made inferences about how learning processes might interact with attention processes to guide adaptive behavior.

Chapter 2

General Methods

In this chapter, I describe the general experimental design and computational models common to all three experiments reported in this thesis. The current approach builds heavily on the earlier work of Gershman et al. (2010) and Wilson & Niv (2011). As was the case in those studies, participants performed a multidimensional decision-making task with probabilistic rewards. The behavioral data were then analyzed using different computational models that make different qualitative assumptions about behavior. Testing the quantitative predictions of the models against actual behavioral data provided an empirical measure of how well each qualitative assumption described participants' behavior.

2.1. The Faces/Houses/Tools (FHT) Task

The Faces/Houses/Tools (FHT) Task used in the current series of experiments was first developed by DeWoskin (2011). It is a variant of the task used in Gershman et al. (2010) and Wilson & Niv (2011), which was in turn based on the Wisconsin Card Sorting Task (Milner, 1963). The task uses compound stimuli, each defined along three dimensions (Figure 2.1). Here, dimensions are operationalized as object categories – faces, houses and tools. Each stimulus contains an exemplar from each of the three categories, vertically arranged into a column. The exemplars can be thought of as features of the dimensions. On each trial, the stimuli were generated by randomly assigning a feature on each dimension to each stimulus. That is, on every trial, each stimulus was a random combination of a face, a house and a tool. Faces, houses and tools were chosen as the dimensions for two reasons. Firstly, these are well-learned categories and participants would intuitively group the individual features along the object categories when presented with the stimuli (Mahon & Caramazza, 2009; Martin, 2007). Secondly, these categories are represented in partially distinct brain areas such that attention to the object categories can be decoded by applying pattern-classification algorithms to neuroimaging data (Norman et al., 2006; Peelen et al., 2009; see Experiment 2).

On each trial, participants were presented with three stimuli. They were tasked to choose one of the stimuli and were rewarded based on their choice. The trials were organized into games of 25 trials. In any one game, only one of the dimensions was relevant to determining reward and only one feature in that dimension was considered the "target" feature and was associated with a high probability of reward. If participants chose the column containing the most rewarding feature in the relevant dimension, they would receive a reward (1 point) with 0.75 probability and no reward (0 points) with 0.25 probability. If they chose a column that did not contain the most rewarding feature, they would receive a reward with 0.25 probability and no reward with 0.75 probability. Probabilistic rewards were used to better emulate real-world learning where outcomes are rarely deterministic. Probabilistic rewards also prolonged the learning process and afforded more power for detailed computational analyses on the dynamics of learning.

To maximize reward received, participants had to figure out the most rewarding feature in the relevant dimension and choose the column containing that feature as many times as possible. Participants were not told which feature was the target feature and had to figure it out through trial-and-error over the course of a game.

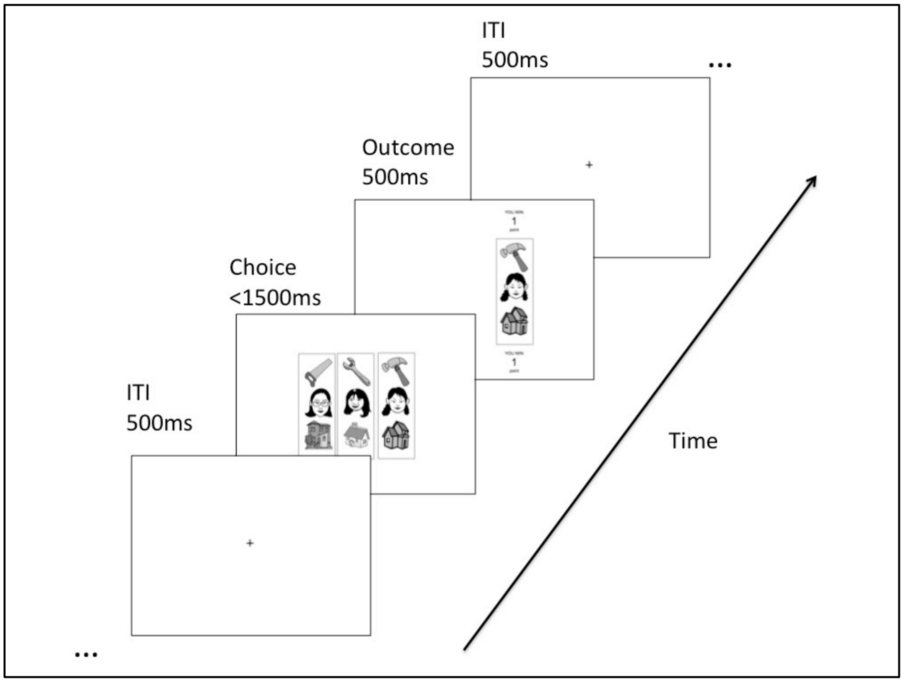

Figure 2.1. An example trial of the FHT task. Participants were presented with three options, each containing a face, a house and a tool. They had 1.5s to choose one of the three columns. Following which, the outcome was displayed for 0.5s. This was followed by 0.5s of ITI before the next trial begins.

2.2. Trial-by-Trial Modeling of Behavior

The attention learning framework proposes that the interaction between learning and attention is highly dynamic, and is modulated on a *trial-by-trial* basis in response to feedback. As such, trial-averaging or block-averaging methods commonly used to study attention would be ill-suited for the current set of studies as they do not keep track of the response and outcome histories from trial to trial. Instead, a more fine-grained approach is necessary to chart out the dynamic relationship between attention and learning. One such approach is to use computational models. In recent years, computational models have been used to analyze both behavioral and neural data (for a review and detailed tutorial of the approach see Daw, 2009). The appeal of computational models lies in their ability to formalize qualitative assumptions as quantitative hypotheses that make predictions about behavior on each trial based on the choices and outcomes on preceding trials. This allows experimenters to investigate how the specific history of events influences current behavior. In addition, computational models can be used to generate hidden variables that underlie the mechanisms being studied. For example, computational models have been used to track value estimates and prediction errors, which are in turn used as regressors to search for corresponding neural correlates (Kable & Glimscher, 2009; Niv, 2009). Given these properties, computational models are well-suited for the current set of analyses.

2.2.1. Notation

For clarity, I preface my description of the different models by defining the notation used in this thesis. The relevant dimension (faces, houses or tools) on trial t is denoted by $d \in \{1, 2, 3\}$. The target feature on that dimension (i.e., the specific face, house or tool) is denoted by $f \in \{1, 2, 3\}$. The three stimuli presented on trial t are denoted by the matrix $\mathbf{s_t} = [s_t(1), s_t(2), s_{t-}(3)]$, where each s_t is a vector that stores the features of the corresponding stimulus. For example, $s_t(1) = [2, 1, 3]$ implies that stimulus 1 on trial t contains the 2nd face, 1st house and 3rd tool. The participant's choice is denoted by c_t where $c_t \in \{1, 2, 3\}$ while the reward received is r_t where $r_t \in \{1, 0\}$. $c_{1:t}$ and $r_{1:t}$ denote the set of choices and rewards from trial 1 to trial t. For compactness, we define $D_{1:t} = (c_{1:t}, r_{1:t})$ as the past history of choices and rewards up to trial t. The reward probability of the most rewarding feature is ρ_h while the reward probability of other features is ρ_l.

2.2.2 Bayesian Model

Given choices and outcomes from all previous trials, the Bayesian model computes the posterior probability distribution over the identity of relevant dimension and target feature (i.e., $p(d, f \mid D_{1:t})$). Using this posterior probability distribution, the model then computes the expected reward of each stimulus (Gershman et al., 2010). The expected reward of each stimulus can be taken as the expected value associated with that stimulus. Since reward is either 1 or 0, the value of a stimulus is equivalent to its reward probability given past choices and outcomes. If the agent had explicit knowledge about the identity of the most rewarding feature and the relevant dimension, the computation of value would be straightforward: The value of a stimulus i would be given by

$$p(r_{t+1}|d, f, s_{t+1}(i)) \begin{cases} \rho_h & \text{if } s_{t+1}(i) \text{ contains } f \text{ in } d \\ \rho_l & \text{otherwise} \end{cases} \qquad \text{(Equation 2.1)}$$

However, as the identity of d and f are unknown, the model has to marginalize out uncertainty over d and f. Hence, the value of stimulus i is given by:

$$p(r_{t+1}|D_{1:t}, s_{t+1}(i)) = \sum_{d} \sum_{f} p(r_{t+1}|d, f, s_{t+1}(i)) p(d, f|D_{1:t}) \qquad \text{(Equation 2.2)}$$

In other words, the value of a stimulus is the expectation of reward weighted by the probability that each dimension-feature pair is the most rewarding one, summed over all dimensions and features for that stimulus.

On each trial, the model makes a choice and receives reward feedback on that choice. This feedback is then used to update the probability distribution over all dimensions and features:

$$p(d, f|D_{1:t+1}) \propto p(r_{t+1}|d, f, c_{t+1}) p(d, f|D_{1:t}) \qquad \text{(Equation 2.3)}$$

where the likelihood $p(r_{t+1} \mid d, f, c_{t+1})$ equals to ρ_h if a reward was received and c_{t+1} contained f in d, 1-ρ_h if a reward was not received and c_{t+1} contained f in d, ρ_l if a reward was received and c_{t+1} did not contain f in d, 1-ρ_l if a reward was not received and c_{t+1} did not contain f in d. The new posterior distribution, $p(d, f \mid \mathrm{D}_{1:t+1})$, is then normalized and used as the prior distribution in computing values for the next trial (Equation 2.2).

The Bayesian model is a model of optimal learning. On every trial, it gains information about all dimension-feature pairs and computes a statistically correct value estimate to guide action.

2.2.3 Function Approximation Model (Uniform Weighting)

This model treats each compound stimulus as a weighted sum of its features. In this instantiation of the model, each feature is weighted uniformly. Essentially, the model generalizes across different possible configurations of stimuli and stores a weight for each dimension-feature pair. The value of choosing a stimulus is then the average of its feature weights.

$$V_t(c_t) = \frac{1}{3}\sum_{d=1}^{3} w_t(d, s_t(c_t)) \qquad \text{(Equation 2.4)}$$

where w_t is the weight matrix and $w_t(d, s_t(c_t))$ is the weight of the feature on dimension d of the chosen stimulus. On each trial, w_t is updated using the TD learning rule

$$\delta_{t+1} = r_{t+1} - V_t(c_t) \qquad \text{(Equation 2.5)}$$

$$w_{t+1}(d, s_t(c_t)) = w_t(d, s_t(c_t)) + \frac{1}{3}\eta\delta_{t+1} \qquad \text{(Equation 2.6)}$$

In particular, the prediction error, δ_{t+1}, is distributed equally across the three features of the chosen stimulus. In other words, the function approximation (FA) model assumes that participants weight each dimension equally and learn on all dimensions of the chosen stimulus. Unlike the Bayesian model, the FA model does not maintain a posterior distribution over the different features. Instead, it keeps track of a point estimate of the weights associated with each feature.

2.2.4. Function Approximation Model with Decay

This model computes stimulus value (Equation 2.4) and updates feature weights (Equations 2.5 and 2.6) using the same equations as the FA model just described. To avoid confusion with the FA model, I refer to this model as the Decay model. It is a decay model in the sense that weights of unchosen features are "decayed" to zero at a particular rate:

$$w_{t+1}(d, s_t(c'_t)) = (1 - \eta_k) w_t(d, s_t(c'_t)) \qquad \text{(Equation 2.7)}$$

where c'_t refers to the unchosen stimuli and η_k is a free parameter that determines the rate of decay. The higher the value of η_k, the faster the decay. If a feature is consistently not chosen, the weight associated with the feature eventually decays to zero. The decay can be thought of as gradual forgetting of the weights of unchosen features over the course of a game. It also allows the model to effectively behave as a selective attention model. The assumption here is that participants attend only to the chosen stimulus and not the unchosen stimuli. In addition, participants might attend to a specific feature and make choices based on the attended feature. If participants choose the same feature over the course of a few trials, the weight associated with that feature would increase while all other weights are decayed. As each stimulus is defined on three dimensions, features from the other two dimensions would be randomly combined with the attended feature. However, as the combinations of features are random, features on the non-attended dimension would not be chosen reliably, and the weights associated with them would not increase significantly. In the case where participants are attending to only one feature at a time, the Decay model behaves similarly as the serial hypothesis testing model proposed by Wilson & Niv (2011). However, unlike the serial hypothesis testing model, the Decay model allows participants to focus on more than one feature at the same time. For example, there might be trials in which two features (on different dimensions) are consistently paired together for several trials. It is possible that participants are attending to both features and testing two hypotheses simultaneously. If the two features are chosen together, the decay model can emulate such diffused attention.

2.2.5. Choice Probabilities

The computational models were applied to generate predictions of participants' choice behavior on each trial. Specifically, choice probabilities were computed according to the commonly used softmax action selection rule:

$$\pi_t(c) = \frac{e^{\beta V_t(c)}}{\sum_a e^{\beta V_t(a)}} \qquad \text{(Equation 2.8)}$$

where $\pi(c)$ is the probability of choosing choice c in the current trial, $V_t(c)$ is the value of this stimulus configuration in the current trial, a represents each action that could be taken in the current trial (i.e., each available stimulus configuration on that trial). β is an inverse-temperature parameter that determines the balance between "exploration" and "exploitation". When β is high, the agent is more likely to exploit known information and choose the option with the highest value. When β is low, the agent will explore less-valued options more frequently.

According to the softmax rule, choice probabilities are weighted as a function of their expected value. The higher the value of a choice, the more likely it will be chosen. Including β as a free parameter then allows for a degree of stochasticity in participants' choice behavior.

2.2.6. Model Comparison

Each of the models makes different assumptions about participants' behavior *(Figure 2.2)*. The Bayesian model assumes optimal learning and uses Bayes' Theorem to compute a statistically correct probability distribution of the identity of the most rewarding dimension-feature pair over all features and dimensions. It also assumes the most diffused focus of attention in that all available information is used to update the PDFs. The FA model is less optimal than the Bayesian model. Instead of maintaining and updating probability distributions, it learns weights associated with each feature. These weights are point estimates that do not take uncertainty into account. As such, the FA model does not make optimal use of available information. The FA model also assumes a stronger focus of attention. Instead of learning on all dimensions and features, the FA model only learns from the chosen features. The Decay model is the least optimal model of the three. Since the target feature does not change within a game, the optimal estimate of value should take into account all past outcomes. The Decay model, however, "forgets" information associated with unchosen features over the course of the game. In addition, the Decay model assumes the strongest attention focus. As described earlier, the weight decay allows the Decay model to emulate a selective attention model. If participants consistently attend and choose the same features, those features will have the strongest weights and would dictate choice.

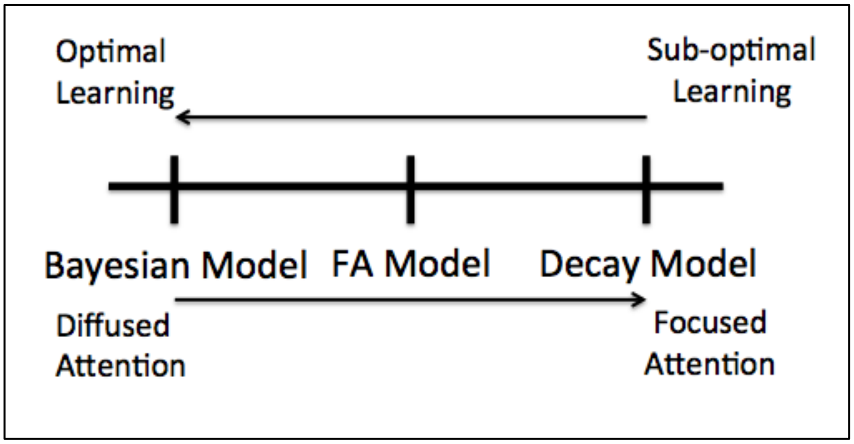

Figure 2.2. Different models make different assumptions about attention and learning.

These different assumptions can be thought of as hypotheses about the attention and learning processes giving rise to the behavioral data. Importantly, the hypotheses can be tested against one another by comparing how well the models explain empirical data. Several metrics have been proposed for the purpose of model comparison (see Daw, 2009 for a discussion of the merits and pitfalls of each). In this thesis, I evaluated model fits using the Bayesian Information Criterion (BIC; Schwarz, 1978), which approximates the posterior probability of a model given behavioral data. Specifically, I first optimized the model parameters by minimizing the log posterior of the data using the Matlab function `fmincon`. The best fitting parameters were then used to compute the BIC approximation to the Bayesian evidence E_m for each model:

$$E_m \approx log(p(D|M, \hat{\theta}_M) - \frac{m}{2} log\ N \qquad \text{(Equation 2.9)}$$

where $p(D|M,\ \hat{\theta}_M)$ is the probability of data D given model M and optimal parameters $\hat{\theta}_M$, while m and N refers to the number of free parameters and data points respectively. The BIC approximation accounts for the number of free parameters with a penalty term (second term of Equation 2.9). To provide a more intuitive measure of model evidence, I divided the total Bayesian evidence of each model by the number of trials for which a participant provided a response, and exponentiated it to yield a corrected average likelihood per trial. The corrected average likelihood per trial varies between 0 and 1, and can be interpreted as the average likelihood of each choice given a particular model. Hence, an average likelihood per trial of 1/3 would indicate that the model is at chance at predicting choices, and that the data provides no support for the model.

For convenience, the list of free parameters for each model, along with accompany priors and constraints, are summarized in Table 2.1.

Model	Parameter	Prior	Constraints
Bayesian	β	Gamma(2,3)	$0 \le \beta \le \infty$
FA	β η	Gamma(2,3) none	$0 \le \beta \le \infty$ $0 \le \eta \le 2$*
Decay	β η η_k	Gamma(2,3) none none	$0 \le \beta \le \infty$ $0 \le \eta \le 2$* $0 \le \eta_k \le 1$

Table 2.1. List of parameters with accompanying priors and constraints used in the model-based analysis. * In the function approximation models, value was computed as the average of three feature weights (Equation 2.4) and the prediction error was distributed evenly across three different features (Equation 2.5). Following this formulation, learning rate η is 9 times larger than that of conventional reinforcement learning update functions. Hence, learning rate η was allowed to vary from 0 to 2, rather than the conventional 0 to 1, so that value estimates would be on the same scale as reward.

2.3. Overview of Current Series of Experiments

In this thesis, I ran a series of experiments aimed at testing the assumptions of the attention learning framework. Experiment 1 investigated participants' learning strategies by applying computational models to analyze choice behavior in a multidimensional decision-making task (Chapter 3). In Experiment 2, I combined fMRI and eye-tracking methods to develop a technique for decoding the focus of attention at a trial-by-trial level (Chapter 4). I then applied this decoding technique in Experiment 3 to investigate the trial-by-trial dynamics of the interaction between learning and attention (Chapter 5).

Chapter 3

Experiment 1: Modeling Behavior in the FHT Task

In Experiment 1, I tested the three models on behavioral data from participants performing the FHT task. Following Gershman et al. (2010) and Wilson & Niv (2011), I hypothesize that participants do not attend to all dimensions of the stimuli when they learn and make choices. Given that working memory constraints limit the maximum amount of cognitive resources one has available for the processing and storage of information, the maintenance of previously processed information would become increasingly difficult as new information is encountered (Baddeley, 2003). As such, I also hypothesize that individuals would gradually forget information of features they have not been attending to. Hence, I predict that the Decay model would account for behavioral data better than the FA and Bayesian models.

3.1. Methods

Participants

Eighteen participants were recruited from the Princeton community (4 male, 15 female, ages 18-30, mean age = 20.8). All participants reported normal or corrected-to-normal vision. The study was approximately 60 minutes in length. Participants received $12 in compensation for their time. Participants also received a cash bonus (up to $6) depending on their performance on the task. Informed consent was obtained from each participant at the start of the session. The study was approved by the Princeton University Institutional Review Board.

Materials

Stimuli consisted of black and white cartoon images of faces, houses and tools (Figure 2.1). Stimuli were presented using Psychtoolbox (Kleiner, Brainard & Pelli, 2007) in Matlab. Participants responded using a keyboard connected to the presentation computer.

Procedures

Participants were first provided with instructions that explained the basic design of the task. They were shown the stimulus display of an example trial. The display consisted of three columns of images, each containing a face, a house and a tool. The vertical ordering of image categories changed from game to game. Participants were told that they had to select one of the three columns and that they would be rewarded based on their choice. They were informed that in any one game, only one category of images (faces, houses or tools) would be relevant to predicting reward and only one image in that category is associated with a high probability of reward. Participants were told that they should try to get as much reward possible and that they would receive a cash bonus based on their performance.

On each trial, participants had 1.5 seconds to choose a stimulus before the trial was aborted. If participants failed to choose a stimulus after 1.5 seconds, the stimuli were removed from screen and a “Too Slow” message was delivered on the center of the screen for 0.5 seconds. If a stimulus was selected, the two unselected stimuli disappeared from the screen and the outcome (“You win 1 point” or “Sorry, 0 points”) was displayed on the top and on the bottom of the chosen stimulus for 0.5 seconds. Following which, there was a 0.5 seconds inter-trial-interval before the next trial began (Figure 2.1). At the end of each game, participants were told the number of points they won in that game.

Participants first played three instructed games in which they were told the relevant object category. These instructed games were included for participants to familiarize themselves with the structure of the game. Following which, the experiment began. The experiment ended after participants played 56 games (1400 trials).

Statistical Analyses

All statistical tests were carried out in MATLAB. Unless otherwise stated, t-tests were two-tailed paired sample t-tests, with alpha = 0.05.

3.2. Analysis and Results

Behavioral Performance

Task performance was evaluated by calculating the percentage of correct choices as a function of the number of trials in a game. This was then averaged across games and subjects to generate the learning curve in Figure 3.1. To be clear, correct choices were defined as choices where participants chose the column containing the most rewarding feature. Chance performance was defined as 33% (1 in 3 chance of choosing the correct column). More detailed analyses indicated that performance on the first 3 trials was not significantly different from chance ($t(17) = 0.86, p = 0.40$) while performance on the last 3 trials was significantly better than chance ($t(17) = 21.38, p < 0.001$), demonstrating that participants were able to solve the task over the course of a game.

Model-Based Analysis

Behavioral data were analyzed using the three models described in section 2.2 .(i.e., the Bayesian model, the FA model and the Decay model). The models were fit for each subject separately and the best-fitting values for each parameter were used for subsequent analyses. Table 3.1 summarizes the mean best-fit values for each free parameter in the different models.

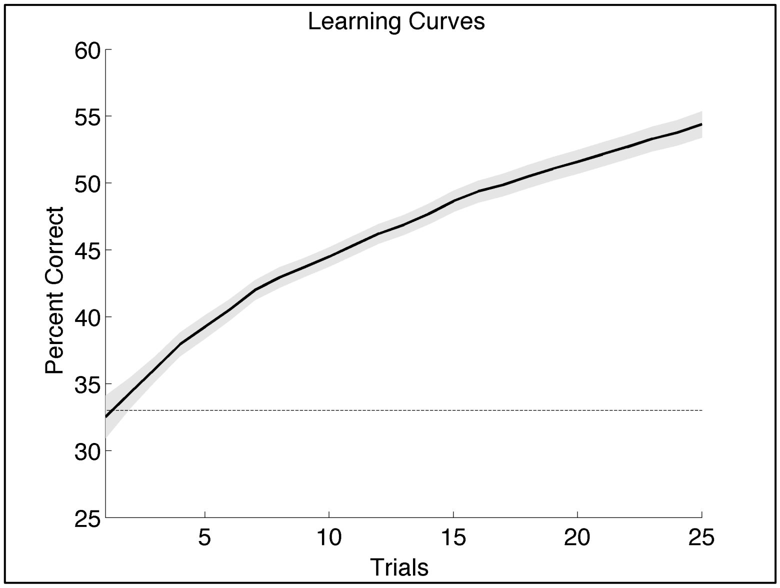

Figure 3.1. Performance on the FHT task as a function of number of trials after the start of the game. Performance was averaged across games and subjects (solid black line). Shaded area indicates S.E.M. Dashed line indicates chance level (33%)

Model	Parameter	Mean (±SEM)
Bayesian	β	4.07 ± 0.18
FA	β η	23.75 ± 2.66 0.21 ± 0.019
Decay	β η η_k	11.79 ± 0.63 1.01 ± 0.04 0.55 ± 0.025

Table 3.1. Summary of best-fit estimates of model parameters (Mean and SEM)

Figure 3.2. shows the average likelihood per trial of each model. All models performed better than chance (Bayesian model: t(17) = 11.4, $p < 0.001$; FA model: t(17) = 15.9, $p < 0.001$; Decay model: t(17) = 19.4, $p < 0.001$). The Decay model provided the best fit to behavioral data with an average likelihood per trial of 0.57 (*SEM* = 0.012). The Bayesian model performed the worst, with an average likelihood per trial of 0.41 (*SEM* = 0.007). Average likelihood per trial of

the FA model was in between at 0.50 (*SEM* = 0.010). A paired t-test indicated that the Decay model performed significantly better than both the Bayesian model ($t(17) = 13.89, p < 0.001$) and the FA model ($t(17) = 21.7, p < 0.001$). The FA model also performed significantly better than the Bayesian model ($t(17) = 8.89, p < 0.05$).

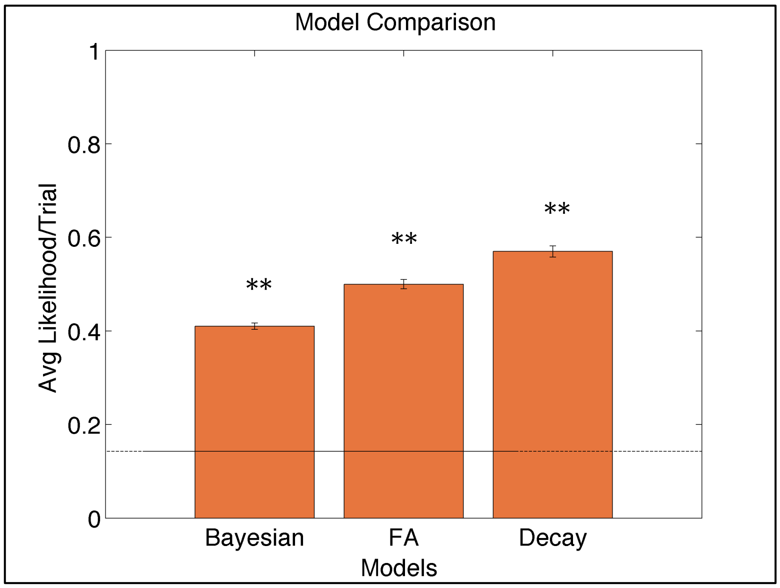

Figure 3.2. Corrected average likelihood per trial of each model. All models performed significantly better than chance. The Decay model performed significantly better than the FA model and the Bayesian Model ($p < 0.001$). The FA model performed significantly better than the Bayesian Model ($p < 0.05$). ** $p < 0.001$)

3.3. Discussion

In this experiment, participants performed a multi-dimensional decision-making task with probabilistic rewards. In any given game, only one dimension was relevant in determining reward and only one feature (the target feature) was associated with a high reward probability. Participants had to figure out the target feature in the relevant dimension through trial-and-error feedback. The learning curve indicates that participants were able to solve the task. Over the course of a game, they were able to figure out the target feature and choose it consistently to obtain maximum reward.

To study the strategy participants used to solve the task, I compared the fits of three different models to participants' trial-by-trial behavior. The Bayesian model performed worst in accounting for participants' choices, suggesting that participants did not employ an optimal Bayesian strategy that learns on all features and dimensions at the same time. This is consistent with the results of other experiments using similar paradigms (Gershman et al., 2010; Wilson & Niv, 2011). Between the remaining two models, the Decay model accounted for participants' behavior better than the FA model. Both of these models update the weights of chosen features based on reward feedback. The key difference between the Decay model and the FA model is that the Decay model decays the weights of unchosen features at a certain rate every trial. There are two possible reasons why decaying unchosen weights would allow the Decay model to perform better than the FA model. Firstly, participants might be using working memory to maintain a representation of all the feature weights. As working memory capacity is limited, information about earlier trials might be forgotten as participants progress through a game. This forgetting is captured by the Decay model but not by the FA model. Secondly, participants might be employing selective attention in solving the task. That is, participants might be learning and making their decisions based on only a subset of features at each point in time. The weight decay allows the Decay model to approximate a selective attention model because only features being attended to and consistently chosen can acquire a significant weight. Choice behavior is then dominated by features that have been consistently attended to.

The results of Experiment 1 suggest that participants utilize a suboptimal strategy of learning only about chosen features and that they gradually forget the weights of unchosen features. They also provide preliminary evidence that participants selectively attend to a subset of features to make choices and learn from outcomes. However, because there was no measure of participants' focus of attention in the experiment, it would be hasty to conclude that participants employ selective attention to solve this task. In the next experiment, I develop and test a method of decoding attention that can be used to directly study the effects of attention on learning at the trial-by-trial level.

Chapter 4

Experiment 2: Decoding the Focus of Attention

One central challenge in studying attention-related processes is that attention is a hidden psychological process that is difficult to measure concretely. In a multidimensional decision-making task like the FHT task, where each option is defined by multiple features, it is not possible to directly infer participants' focus of attention from choice data alone. In other words, because participants choose a compound stimulus consisting of face, a house and a tool on each trial, it is not possible to know whether they are choosing the stimulus because of the face, the house, the tool, or some combination of these features. This is particularly problematic if we are interested in studying trial-to-trial changes in the focus of attention. The Decay model circumvents this problem by implicitly tracking the sequence of choices, which can help make inferences about attention. That is, because the model preserves the weights of chosen features while decaying those of unchosen features, we can read out participants' focus of attention from the feature weights. However, the Decay model can at best provide an indirect measure of attention. To better study the dynamic interaction between learning and attention, a more precise method is needed. In Experiment 2, I designed and tested a novel technique that decodes the focus of attention at the trial-by-trial level by combining eye-tracking with fMRI methods.

Eye-tracking involves measuring participants' point of gaze. The method has been used to localize attention in various experimental paradigms (Grant & Spivey, 2003; Rehder & Hoffman, 2005). The assumption here is that the point of gaze reflects the focus of attention. As such, eye fixations to a particular location in space can be interpreted as attention to that spatial location. This assumption is supported by a wealth of evidence suggesting that eye movements and attention processes are tightly linked (Deubel & Schneider, 1996; Kowler et al., 1995). As the stimulus dimensions are spatially separated in the FHT task, eye fixations to the different dimensions can be measured and taken as a proxy measure of attention to those dimensions.

Attention, however, can operate covertly in the absence of overt motor behavior (Juan, Shorter-Jacobi & Schall, 2004; Posner, 1980). Eye-tracking captures only overt attention and does not provide a measure of covert attention. To measure covert attention, I turned to pattern classification of fMRI data (Norman et al., 2006). The approach was motivated by findings indicating that different object categories are represented in the brain by partially distinct patterns of neural activity (Mahon & Caramazza, 2009; Martin, 2007). A pattern classifier can then be trained to decode the object category being processed given a particular pattern of activity (Haxby et al., 2001). As attention biases processing in favor of attended stimuli, the pattern of neural activity would also reflect the focus of attention (Peelen et al., 2009; Peelen & Kastner, 2011).

In this experiment, to verify and test the utility of both measures of attention, participants performed a variant of the FHT task in which the relevant dimension was revealed to them at the start of each game. Assuming that participants are rational and optimal, they would attend only to the relevant dimension throughout the game. Participants performed the task in an MRI scanner while their brain activity was recorded and their eye movements were tracked. A linear support vector machine (SVM) was used to decode participants' focus of attention from fMRI data. The output of the SVM on each trial was then combined with eye-tracking measures to obtain a final prediction of the focus attention. The accuracy of the technique was evaluated by comparing the predictions with participants' focus of attention as determined by the instructions in each game.

4.1. Methods

Participants

Five participants were recruited from the Princeton community (3 males, 2 females, ages 18-20, mean age = 19). All participants were right-handed and reported normal or corrected-to-normal vision. The study was approximately 120 minutes in length. Participants received $40 in compensation for their time. They also received a cash bonus (up to $6) based on performance. Informed consent was obtained from each participant at the start of the session. The study was approved by the Princeton University Institutional Review Board.

Materials

Stimuli consisted of grey-scaled photographs of famous faces (Albert Einstein, Abraham Lincoln, George Clooney), famous landmarks (Big Ben, Notre Dame, Taj Mahal) and common tools (hammer, screwdriver, spanner) (Figure 4.2). This change was implemented following results from a pilot study indicating that the SVM was better at classifying neural activity evoked by this set of stimuli than that evoked by cartoon images of faces, houses and tools (results not shown).

Procedures

Participants performed three different tasks that were run within the same scanning session: (1) one-back detection task (henceforward, 1BDT) (2) three-dimensional one-back detection task (3D1B) and (3) instructed FHT task (iFHT). Participants were provided with instructions and practiced each task before the start of the scanning session.

One-back detection task (1BDT). On each trial, participants were presented with three images from one object category. In any one game, the same three images from the same object category were used. Participants' task was to respond with a button press when the order of three images on the current trial matches that from the previous trial (Figure 4.1). The order of images was pseudorandomized such that there was a 1/3 probability that the order on each trial was the same as that on the previous trial. The images were presented for 1.4s during which participants could make their response with a button press. Following which, the outcome of the trial was presented for 0.5s. Correct hits (responding when a response was required) were rewarded with a gain of 3 points while misses (failing to respond when they should) and false alarms (responding when a response was not required) were punished with a loss of 3 points. Correct rejections did not incur a reward. Each trial was followed by a 0.1s ITI. Participants performed 2 runs of the one-back detection task, each containing 21 games of 10 trials each (210 trials per run). One third of the games were Face games that showed images of Einstein, Lincoln and Clooney in different orders. One third of the games were Landmark games that showed images of Big Ben, Taj Mahal and Notre Dame in different orders. The remaining one third of the games were Tool games that showed images of a hammer, a screwdriver and a spanner in different orders. Within each run, the game types were presented in counterbalanced order using a Latin square design to minimize possible order effects.

Three dimensional one-back detection task (3D1B). On each trial, participants were presented with all 9 images from the 3 object categories. However, only one of the categories was behaviorally relevant. Participants were instructed regarding the relevant category, and told to attend only to this category and ignore the other two categories. They were instructed to perform the one-back detection task on the relevant category (Figure 4.2). The relevant category changed every few trials, always in an instructed way. Each game consisted of 45 trials and each category was relevant 5 times in a game. A change in the relevant category was signaled by a red rectangle around the new relevant category. The order of relevant categories was counterbalanced using a Latin square design to minimize possible order effects. On each trial, the stimulus display would be presented for 1.4 seconds, during which participants can make their response. Following which, the outcome of the trial was presented for 0.5s. Outcomes were determined as in the one-back detection task. Each trial was followed by a variable ITI (2-6s, Mean = 3.51s). Participants performed 2 runs of the three dimensional one-back detection task, each consisting of 3 games (135 trials per run).

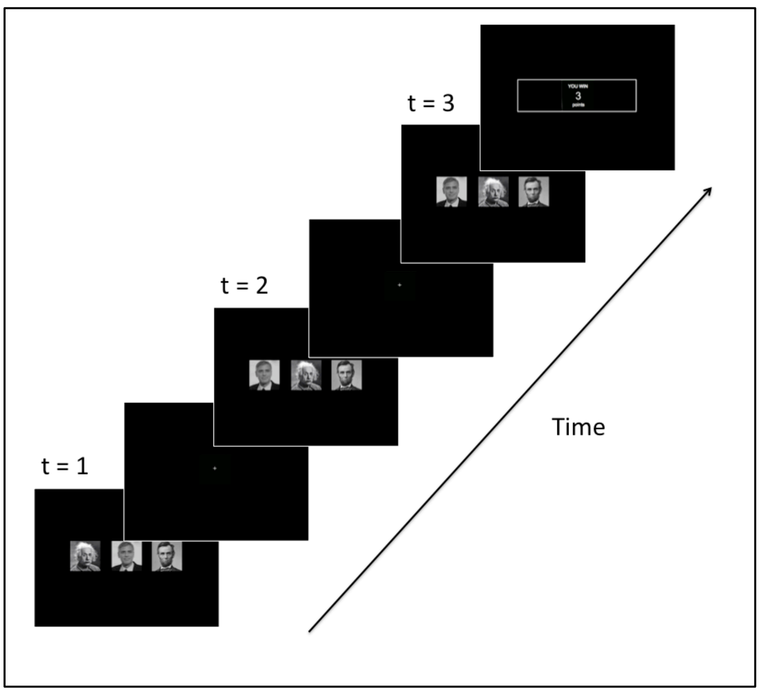

Figure 4.1. Three example trials of the one-back detection task (1BDT). The order of faces on trial 3 is the same as the order of faces on trial 2. Participant responds correctly and is rewarded with three points.

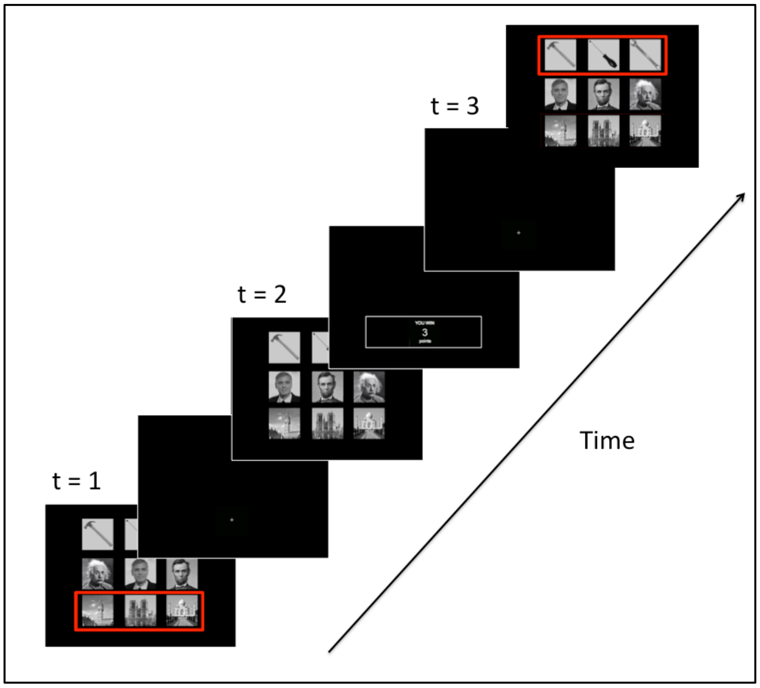

Figure 4.2. Three example trials of three dimensional one-back detection task (3D1B). On trial 1, the relevant category is landmarks. On trial 2, the order of landmarks is the same as that on trial 1. Participant responds correctly and is rewarded with three points. On trial 3, the relevant category is changed to tools, as indicated by the red rectangle. Participant should now play the one-back detection task on tools, and as such not respond even if the order of the landmarks is the same as that on the previous trial.

Instructed FHT task (iFHT). This was a variant of the FHT task in which participants were told the relevant dimension at the start of each game. To maximize reward, participants had to figure out the most rewarding feature in that dimension over the course of the game. On each trial, participants had 1.5 seconds to choose a stimulus before the trial was aborted. Outcome was determined as in the original FHT task, and was displayed for 0.5 seconds. Each trial was followed by a variable ITI (2-6s, Mean = 3.51s). At the end of each game, participants were asked to indicate the correct dimension of that game. This was to ensure that participants understood the instructions and were attending to the correct dimension. Data were excluded for games in which participants answered this question incorrectly. Participants performed two runs of the iFHT task. Each run consisted of 18 games of 10 trials each (180 trials per run). In one-third of the games, face was the relevant dimension, in another one-third, landmark was the relevant dimension and in the remaining one-third, tool was the relevant dimension. The different game types were presented in counterbalanced order using a Latin square design to minimize possible order effects.

Eye-tracking methods. Eye-tracking data were acquired using the iView X MRI-LR system (SMI Sensomotoric Instruments) with a sampling rate of 60 Hz. The system output files were then analyzed using in-house MATLAB code. Data from before 200ms after each stimulus offset were discarded to account for saccade latency.

fMRI methods. Imaging data were collected using a 3-Tesla MRI scanner (Siemens Skyra; Siemens). At the start of each session, a high-resolution T1-weighted structural image of the participant's brain was obtained (magnetization-prepared rapid acquisition gradient echo sequence; TR = 2300 ms; TE = 3.1 ms; flip angle = 9°; voxel size = $1 \times 1 \times 1$ mm^3). Participants performed 6 functional runs in the following order: 1 run of 1BDT, 2 runs of iFHT, 2 runs of 3D1B and 1 run of 1BDT. Stimulus onset on each trial was time-locked to the start of a TR. For all functional runs, 34 transverse slices were acquired in interleaved order (echo planar sequence, TR = 2000 ms; TE = 30 ms; flip angle = 71°; voxel size = $3 \times 3 \times 3$ mm^3). Slices were tilted at 30° to the AC-PC line to minimize distortion and signal loss in orbitofrontal cortex and the medial temporal lobes. Image volumes were preprocessed using FEAT v.5.98 that is available as part of the FMRIB software library (FMRIB, Oxford, UK). Volumes were motion corrected using FSL's MCFLIRT with the first volume of the first run as the reference volume. The optimization uses trilinear interpolation. Each functional scan was also linearly detrended and transformed into z-scores using the PyMVPA software package (Hanke et al., 2009).

Linear SVMs with a fixed regularization parameter of $C = 1$ were trained on data from either the 1BDT or 3D1B task to distinguish between the three different object categories given patterns of neural activity. Analysis was restricted to voxels in a ventral visual stream mask consisting of the bilateral occipital lobe and ventral temporal cortex (Figure 4.3). The mask was created in MNI space and transformed into each participant's native space using FSL's FLIRT implementation. Classification was implemented within PyMVPA using the LIBSVM software.

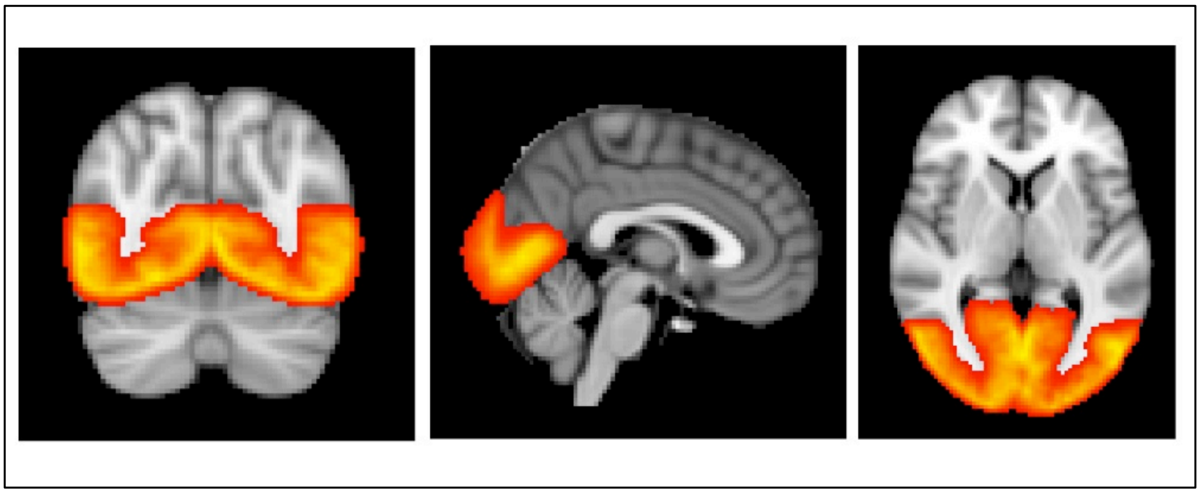

Figure 4.3. Ventral visual stream mask (orange) overlaid on brain template in MNI space. From left, coronal, sagittal, and axial slices.

4.2. Analysis and Results

Predicting Attention using Eye-Tracking. A rectangular area of interest (AOI) was defined around each object category (Figure 4.4). The proportion of eye-fixations within an AOI between 200ms after stimulus onset to outcome offset was taken as a measure of attention to the corresponding category. The object category with the highest proportion of fixations was taken as the prediction of the focus of attention on that trial. When the predictions were compared to participant's actual focus of attention in iFHT games, mean accuracy was 73.4% and significantly above chance ($t(4) = 3.6$, $p < 0.05$) (Figure 4.5).

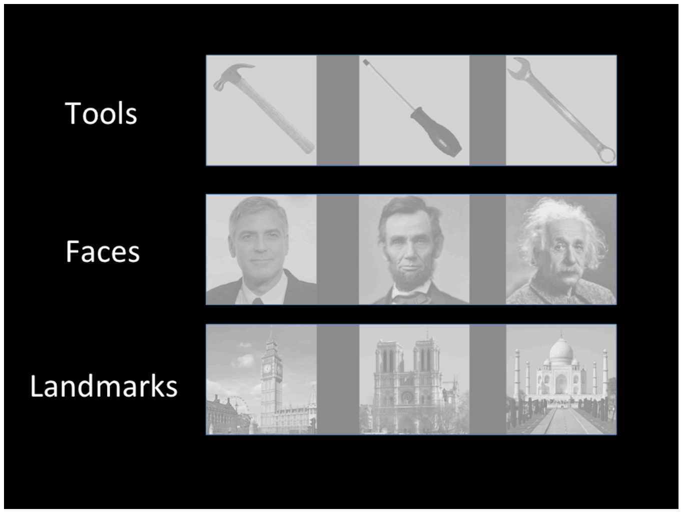

Figure 4.4. Areas of interest (shaded area) for each object category.

Predicting Attention from fMRI Data. When the SVM was trained on data from the 1BDT, mean accuracy in predicting the focus of attention in the iFHT was 65.7% and significantly better than chance (t(4) = 6.3, $p < 0.05$) while accuracy in predicting the focus of attention in the 3D1B task was 72.6% on average and also significantly better than chance (t(4) = 5.6, $p = 0.005$). When the SVM was trained on data from the 3D1B task, mean accuracy in predicting the focus of attention in the iFHT task was 81.3%, which was significantly better than chance (t(4) = 10.5, $p < 0.001$) (Figure 4.5). The difference in accuracy when the SVM was trained on the 3D1B and when the SVM was trained on the 1BDT was marginally significant (t(4) = 2.3, $p = 0.08$). In contrast, the SVM trained on the 3D1B task was not significantly different than that from eye-tracking (t(4) = 1.2, $p = 0.29$),

Composite Measure of Attention. To obtain a composite measure of attention, the proportion of eye-fixations to a category on each trial was multiplied by the corresponding output from the classifier. The classifier was trained on the 3D1B task. This choice was motivated by two factors. Firstly, there was a marginally significant increase in classifier performance when training on data from the 3D1B task as compared to the 1BDT. Secondly, the

3D1B task was more similar to the actual FHT task in that all three object categories were presented to the participant on each trial and the focus of attention could switch every few trials. The combined measure of attention was normalized to a value that ranged from 0 to 1. The object category with the highest value was taken as the prediction of the participants' focus of attention on that trial. The mean accuracy of the combined measure was 89.4% and was significantly better than chance ($t(4) = 15.4$, $p < 0.001$). This accuracy was higher than that obtained from eye-tracking or pattern-classification alone, but the differences were not statistically significant (compared to eye-tracking: $t(4) = 1.99$, $p = 0.11$; compared to pattern classification: $t(4) = 1.8$, $p = 0.14$).

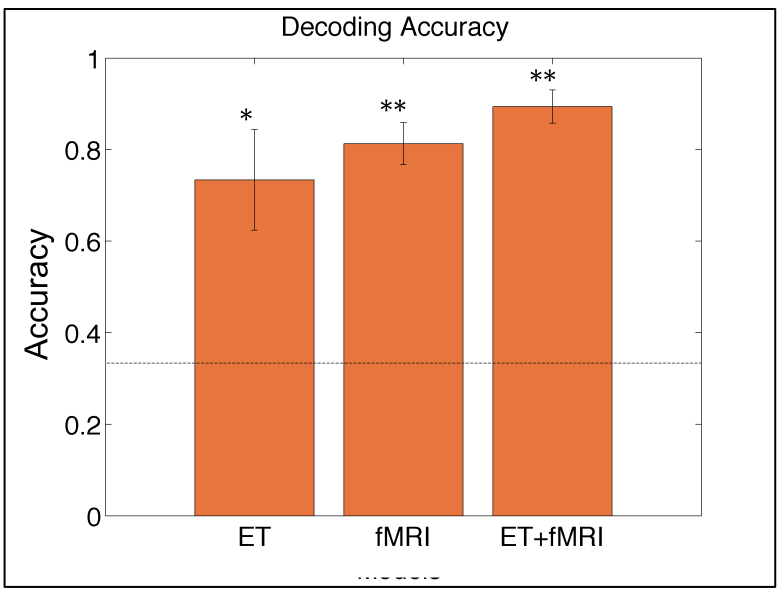

Figure 4.5. Decoding accuracies of different methods. ET: Eye-tracking; fMRI: Pattern classifier trained on data from 3D1B task; ET+fMRI: Composite measure of attention combining ET and fMRI. All accuracies were significantly above chance (dashed line) but not significantly different from one another. *$p < 0.05$; ** $p < 0.001$)

4.3. Discussion

The aim of this experiment was to develop and test a technique that decodes the trial-by-trial focus of attention in the FHT task. Participants played a variant of the FHT task in which they were told the relevant dimension at the start of each game. Assuming that participants attended only to the relevant dimension throughout the game, this provided the ground truth needed to evaluate the accuracy of different decoding methods. Both eye-tracking and pattern classification of fMRI data were highly accurate in predicting participants' focus of attention. Surprisingly, the pattern classifier trained on data from 3D1B task performed better than the pattern classifier trained on the 1BDT. In the 3D1B task, participants were shown all 9 images from all three categories. In contrast, participants were presented with one object category at a time in the 1BDT. One would expect the 1BDT to provide "cleaner" training data to the classifier and thus lead to high classification accuracies. However, one possible reason why training on data from the 1BDT would lead to lower classification accuracies is that the 1BDT was organized into games in which the same object category was shown for 10 consecutive trials. This would result in strong temporal autocorrelations between training patterns for the same object category. As such, noise would be strongly correlated between training patterns and would hurt classification performance (Pereira, Mitchell & Botvinick, 2009). In contrast, the focus of attention switched every few trials of the 3D1B task and the training patterns were spread further apart. Hence, even though the 3D1B task had fewer trials overall, it provided a better set of training data than the 1BDT.

One might worry that accuracy in decoding the focus of attention in the iFHT task would not translate to accuracy in decoding the focus of attention in the regular FHT task. This is particularly so since the iFHT task was organized into games of 10 consecutive trials in which participants attended to the same object category. In contrast, the focus of attention might change every few trials in the regular FHT task as participants cannot be sure of the relevant category on each game. This might be problematic for the decoding of attention by pattern classification of fMRI data due to the temporal smoothing caused by the hemodynamic lag. Specifically, the BOLD signal associated with one category on a particular trial might be modulated by the signal associated with other categories on previous trials, thus confusing the classifier. To test the temporal resolution of our decoding methods, I trained a classifier on data from the 1BDT and tested it on data from the 3D1B task. The 3D1B task is similar to the regular FHT task in that the focus of attention switches every few trials. Accuracy in predicting attention during the 3D1B task was high at 72.6%, suggesting that the method is able to decode the focus of attention with reasonable temporal resolution.

I then combined the eye-tracking results with the pattern classification results to obtain a composite measure of attention. The composite measure predicted the focus of attention better than eye-tracking or pattern classification alone, but the differences were not statistically significant. Given that there were only 5 participants and that the difference was trending towards significance, more data should be collected before conclusions are drawn about performance difference between the composite measure and its individual components. Nevertheless, it can be argued that there is value in using the composite measure as a predictor of attention. In the current experiment, participants were always instructed regarding the relevant category to attend to. The optimum strategy would be to fixate only on the relevant category. As such, both overt and covert attention would be directed towards the same category. However, in the regular FHT task, participants do not know what category is relevant. In that case, there might be a stronger incentive to rely on covert processes to attend to multiple categories at the same time. Since the composite measure of attention performs better (or at least as well) as the individual measures alone, I used it in Experiment 3 to track the focus of attention as participants played the regular FHT task.

Chapter 5

Experiment 3: Attention Processes in the FHT Task

The aim of this experiment was to investigate if participants employ selective attention to solve the FHT task, and to characterize how attention processes are affected by choice behavior and reward feedback. In Experiment 2, I developed and tested a technique for decoding the trial-by-trial focus of attention in a variant of the FHT task. Experiment 2 demonstrated that by combining eye-tracking with pattern classification of fMRI data, a relatively accurate measure of attention can be computed. I now apply this technique to decode the focus of attention in the regular FHT task.

Participants' choice behavior was modeled using the FA model and the Decay model. The Bayesian model was excluded since Experiment 1 had demonstrated that it was a poor model of choice behavior. I also built modified versions of the FA model and the Decay model that incorporated an explicit selective attention component in determine choice and learning. If participants used selective attention to learn and make decisions, models with a selective attention component would predict choice behavior better than models without a selective attention component. I was also interested in modeling how the focus of attention was modulated. Based on accumulating evidence suggesting that attention is drawn to stimuli that have acquired high value (Anderson, Laurent & Yantis, 2011a; 2011b; Hickey et al. 2010; Peck et al., 2009), I predicted that attention would be directed to features that are associated with high value.

5.1 Methods

Participants

Eight participants were recruited from the Princeton community (3 males, 5 females, ages 18-20, mean age = 18.9). All participants were right-handed and reported normal or corrected-to-normal vision. The study was approximately 120 minutes in length. Participants received $40 in compensation for their time. They also received a cash bonus (up to $6) based on performance. Informed consent was obtained from each participant at the start of the session. The study was approved by the Princeton University Institutional Review Board. Data from one participant was discarded because the participant fell asleep during the experiment.

Materials

Experiment 3 used the same stimuli as Experiment 2 (Figure 4.2).

Procedures

Participants performed four runs of the FHT task and two runs of the 3D1B task, in that order, within a single scanning session. Prior to the experimental runs, a structural image of the participant's brain was obtained. During this structural scan, participants played seven practice games. The first four practice games were "fast" games with ITIs of 0.5 seconds while the remaining three practice games were "slow" games with variable ITIs between 2-6 seconds (Mean = 3.5s). The "fast" games were included to facilitate the learning of the reward structure of the game. During experimental runs, participants only played "slow" games. Each run of the FHT task consisted of 6 games of 25 trials each (150 trials per run). Aside from including variable ITIs and using real photographs instead of cartoon images, the task was identical to the one used in Experiment 1. Each run of the 3D1B task consisted of 3 games of 45 trials each (135 trials per run). The 3D1B task was identical to the one used in Experiment 2.

Data Acquisition and Pre-processing. Eye-tracking data and fMRI data were acquired and pre-processed using the same procedure as Experiment 2, except that functional volumes were collected with a 10% gap to allow for better coverage of the whole brain. Eye-tracking data were not acquired for one participant due to technical failure.

Modifications to Models. The FA model and Decay model were modified by incorporating a selective attention component to value computation and update. In particular, stimulus value was calculated as a weighted average of the feature weights using dimensional attention as weights for each feature (Figure 5.1):

$$V_t(c_t) = \sum_{d=1}^{3} w_t(d, s_t(c_t))\phi_d \qquad \text{(Equation 5.1)}$$

where ϕ_d is the composite measure of attention to dimension d computed by normalizing the product of the proportion of eye fixations to d and the corresponding output of pattern classifier.

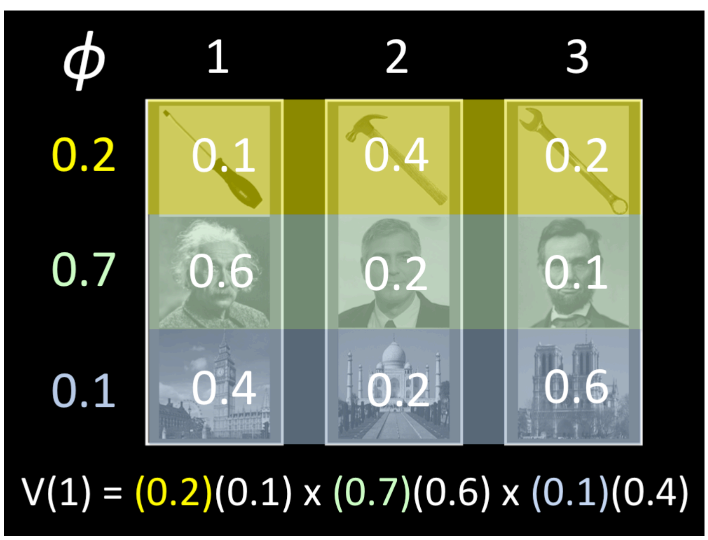

Figure 5.1. Value computation in selective attention models. The leftmost column shows the attentional focus to each object category (yellow – tools; green – faces; blue – landmarks). Feature weights are shown on top of the corresponding feature. Value of stimulus is calculated as the average weight of the individual features weighted by attentional focus to the corresponding category. For example, the value of the leftmost stimulus (stimulus 1) is calculated by adding the product of the attentional focus to Tools and the weight of the screwdriver, the product of the attentional focus to Faces and the weight of Einstein, and the product of the attentional focus to Landmarks and the weight of Big Ben.

The weight updates of the chosen feature were also scaled by ϕ_d:

$$w_{t+1}(d, s_t(c_t)) = w_t(d, s_t(c_t)) + \phi_d \eta \delta_{t+1} \qquad \text{(Equation 5.2)}$$

5.2 Analysis and Results

Behavioral Performance

Figure 5.2 shows the average learning curve for all participants. Average performance on the first 3 trials was not significantly different from chance (t(6) = 0.47, p = 0.66) while average performance on the last 3 trials was significantly better than chance (t(6) = 7.1, $p < 0.001$), demonstrating that participants were able to solve the task.

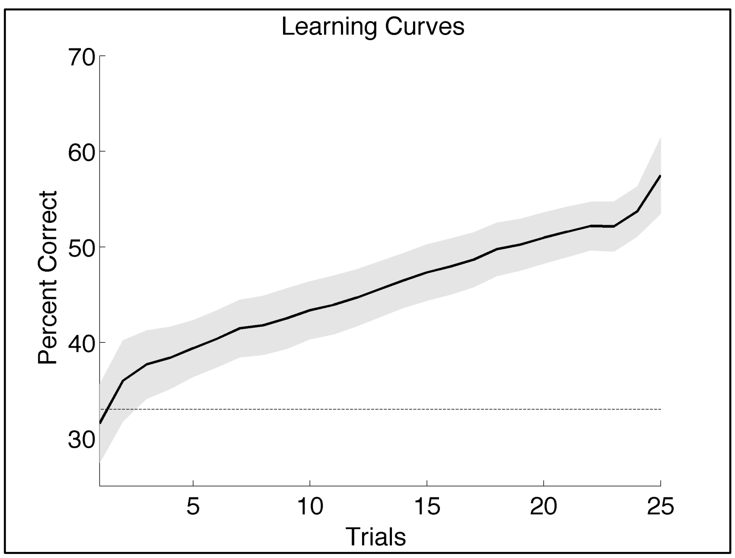

Figure 5.2. Performance on the FHT task as a function of number of trials after the start of the game. Performance was averaged across games and subjects (solid black line). Shaded area indicates S.E.M. Dashed line indicates chance level (33%).

Validation of Eye-Tracking Performance

Participants were cued to attend to specific object categories during the 3D1B task. As such, a measure of eye-tracking accuracy can be obtained by comparing the attentional focus predicted from the eye-tracking results with the categories participants were instructed to attend.

The average accuracy of the eye-tracking predictions was 83%, which was significantly above chance ($t(5) = 5.3$, $p < 0.001$).

Leave-One-Out Cross Validation of Classifier Performance

To obtain a measure of classifier performance, I first ran a leave-one-out cross-validation analysis on the data from the 3D1B runs. As there were two 3D1B runs, cross-validation involved two iterations. On the first iteration, one of the runs was used as the training set while the other was used as the testing set. On the second iteration, the training and testing sets were swapped. Classification accuracies for the two iterations were averaged for each participant. The average classification accuracy across participants was 84.3%, which was significantly above chance ($t(6) = 43$, $p < 0.001$).

Model-based Analysis of Choice Behavior

Participants' choice behavior was first fitted using the FA model and the Decay model. Both models predicted choices significantly better than chance (FA: $t(6) = 6.3$, $p < 0.001$; Decay: $t(6) = 11.3$, $p < 0.001$). The Decay model predicted choices significantly better than the FA model ($t(6) = 9.1$, $p < 0.001$). Participants' choice behavior was then fitted with the selective attention version of the models (FA_SA and Decay_SA). Both models also predicted choice significantly better than chance (FA_SA: $t(6) = 6.3$, $p < 0.001$; Decay_SA: $t(6) = 8.9$, $p < 0.001$). The FA_SA model predicted choice behavior better than the FA model, but the difference was not significant ($t(6) = 1.2$, $p = 0.27$). The Decay_SA model predicted choice behavior worse than the Decay model ($t(6) = 2.6$, $p = 0.042$). The mean best-fit values for each parameter are summarized in Table 5.1. Model-based results are summarized in Figure 5.3.

A post-hoc analysis comparing the fits of FA model and the FA_SA model at the individual participant level revealed individual differences in which model was favored for each participant (Figure 5.4). For the majority of the participants (5 out of 7), the FA_SA model explained the data significantly better than the FA model.

Model	Parameter	Mean (±SEM)
FA	β η	21.3 ± 3.38 0.27 ± 0.08
FA_SA	β η	15.6 ± 2.91 0.18 ± 0.02
Decay	β η η_k	12.48 ± 1.42 1.02 ± 0.12 0.57 ± 0.09
Decay_SA	β η η_k	14.9 ± 1.73 0.21 ± 0.03 0.49 ± 0.09

Table 5.1. Summary of best-fit estimates of model parameters (Mean and SEM)

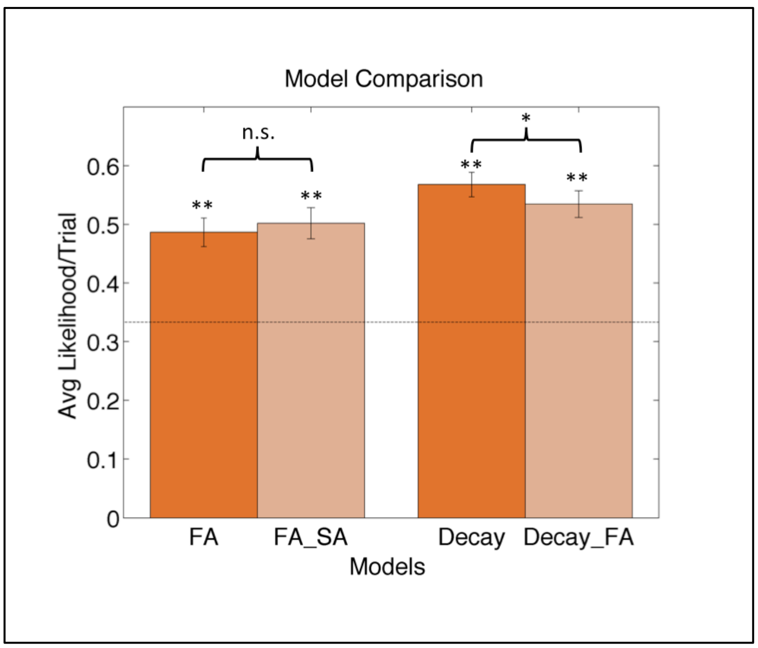

Figure 5.3. Corrected average likelihood per trial of each model averaged across all participants. All models performed significantly better than chance. Performance between FA and FA_SA models did not differ significantly. The Decay model performed significantly better than the Decay_FA model. *$p < 0.05$, **$p < 0.001$

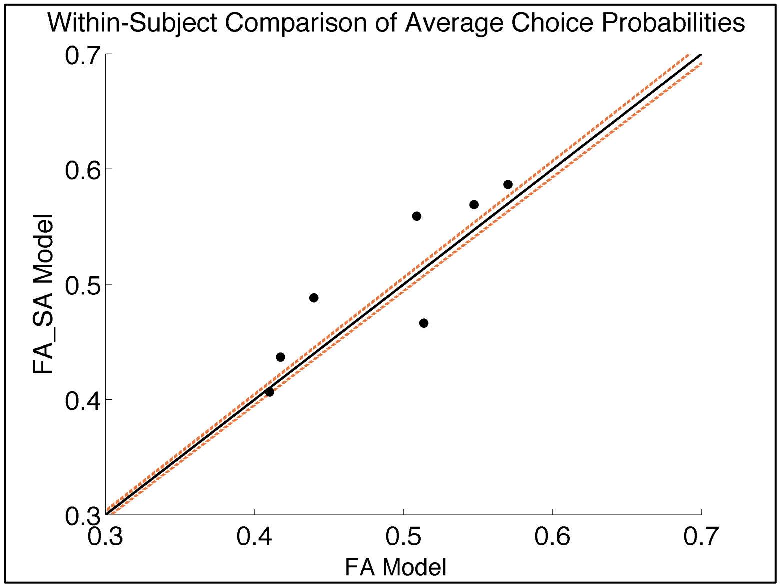

Figure 5.4. Within-participant comparison of average likelihood per trial for FA and FA_SA models. Points above the diagonal are favored by the FA_SA model. The orange dashed line indicates the confidence interval outside of which one model is more likely than the other with $p < 0.001$.

Decay Model as a Selective Attention Model

To test the claim that the Decay Model behaves as a selective attention model, I compared the trial-by-trial feature weights generated by the Decay model (with best-fitting parameters) to the composite measure of attention. The results indicate that on 58% (SEM = 0.03) of the trials, the dimension with the strongest attention focus was also the dimension with the feature of the highest weight. This proportion was significantly above chance ($t(6) = 8.6$, $p < 0.001$). Trials on which there was more than one feature with the highest weight were excluded from this analysis.

Value and Attention

In a separate analysis, trial-by-trial feature weights were computed using the FA model with the best-fitting parameters. To investigate the relationship between learned value and attention, I compared the feature weights on each trial to the composite measure of attention.

Results indicate that on 55% (SEM = 0.04) of the trials, the dimension with the strongest attention focus was also the dimension with the feature of the highest weight. This proportion was significantly above chance ($t(6) = 5.6$, $p < 0.05$). As in the analysis with the Decay model, trials on which there was more than one feature with the highest weight were excluded.

5.3. Discussion

Experiment 3 explored the interaction between attention and learning. Specifically, I applied the technique developed in Experiment 2 to decode the trial-by-trial focus of attention as participants performed the FHT task. This focus of attention was incorporated into computational models of behavior to investigate the role of attention in guiding choice and learning. If participants employed selective attention during choice-selection and learning, and if my decoding of their selective attention was accurate, models that incorporate a selective attention component should predict behavior better than those that do not. Furthermore, if correctly decoded, attention should be directed towards features that have acquired high value. The current results provide mixed support for these hypotheses.

The learning curve indicated that participants were able to solve the task in the MRI scanner. Model-based analysis of behavioral data indicated that the Decay model accounted for behavior better than the FA model, replicating the findings of Experiment 1. To study the role of attention, the models were then modified to include a selective attention component (FA_SA and Decay_SA models). Specifically, the modified models assumed that value computation and update is weighted by attentional focus. The attentional focus was computed by combining eye-tracking results with pattern classification of fMRI data as described in Experiment 2.

Numerically, the FA model with selective attention (FA_SA) predicted choice behavior better than the regular FA model, but the difference was not significant. Gershman et al. (2010)

had earlier demonstrated that there might be individual differences in strategies. These differences can be masked when model fits are averaged across participants. As such, I ran a within-subject analysis that tested the FA_SA model against the FA model for each participant. The analysis revealed individual differences in which model best fit the data. For the majority of participants, the FA_SA model accounted for behavior significantly better than the FA model. However, for one participant, the FA model accounted for behavior significantly better than the FA_SA model. Given this small sample size, the experiment is underpowered to make conclusive remarks about the extent to which each model is preferred in the general population. Nevertheless, my results do suggest that a reasonable number of participants use selective attention when making choices and learning.

Model comparison between the Decay model and the Decay_SA model was more conclusive. The Decay model provided a significantly better account of behavioral data than the Decay_SA model for all participants. This was a surprising result, as it seemed to suggest that participants did not employ selective attention in solving the task. However, such a conclusion would be premature. As described earlier, the Decay model implicitly implements an attentional focus by tracking the sequence of choices. Thus, adding the composite measure to the Decay model might have impaired model performance by assuming an overly narrow focus of attention. In addition, it is possible that multiplying eye-tracking results with the outputs of the SVM classifier is not the right way to combine the two measures. Another important factor is that the model assumes that attention is constant throughout the whole trial. However, it has been suggested that attention during choice and attention during learning are dissociable (Gottlieb, 2012). While attention during choice tends to be preferentially directed towards valuable stimuli, attention during learning tends to be preferentially directed towards stimuli with uncertain value, with the goal of reducing that uncertainty. As such, participants may be attending to one dimension when they make a choice and attending to another dimension when they receive reward feedback. The Decay_SA model conflates both types of attention, which might have hurt its performance in predicting choice behavior.

To directly test the claim that the Decay model emulates a selective attention model, I compared the feature weights generated from the Decay model to the focus of attention predicted by the composite measure of attention. Indeed, the dimension of the highest-weight feature was, on most trials, the same dimension that the composite measure of attention predicted participants were most attending to. This result suggests that the feature weights of the Decay model tracked participants' focus of attention. Interestingly, the same result was observed when the analysis was repeated with feature weights computed using the FA model, even though it did not include a selective attention component. One explanation is that attention was preferentially directed towards features that had been consistently associated with reward. This explanation would be consistent with previous work (Anderson, Laurent & Yantis, 2011a; 2011b; Hickey et al. 2010; Peck et al., 2009) and with the hypothesis that attention is modulated by learned outcomes. From an evolutionary perspective, such an attentional bias would be advantageous as stimuli that are predictive of reward tend also to be the stimuli that are relevant to behavior.

Due to the intricate relationship between attention and learning (i.e. attention determines what we learn about, but we also learn what to attend to), it is difficult to deconfound the influence of attention process on learning from the influence of learning processes on attention with the current experimental design. Nevertheless, I hope the current results have provided evidence that attention and learning work together. Other experimental paradigms are needed to characterize their individual roles.

Chapter 6

General Discussion

This thesis opened with the premise that the incredible complexity of the real world presents serious computational challenges for learning processes. Selective attention has been proposed as a mechanism that facilitates learning by constraining the amount of information processed at any given time (Gershman et al., 2010; Wilson & Niv, 2011). Existing theories of learning, however, have largely ignored or simplified the role of attention. Similarly, existing theories of attention have rarely considered the influence of learning mechanisms on attention processes. In this thesis, I attempted to integrate these two largely independent bodies of research by presenting an attention-learning framework describing how attention and learning processes work hand in hand to facilitate adaptive behavior. I then conducted a series of experiments aimed at testing the assumptions of the framework. In Experiment 1, I conducted model-based analysis of participants' choice behavior to investigate their strategy in solving a multi-dimensional probabilistic decision-making task. The results showed that participants adopted strategies favoring computational efficiency over optimality. In Experiment 2, I developed and tested a technique for decoding participants' focus of attention in the same task at the trial level. In Experiment 3, I applied this decoding method to study how attention interacts with learning. In the subsequent sections of this chapter, I briefly recapitulate the ideas behind the attention learning framework and describe how the current results fit within the framework. Finally, I summarize the main lessons learnt from this work and chart out future directions.

6.1 The Attention-Learning framework

The central idea behind the attention-learning framework can be summarized as follows: attention determines what we learn about, but we also need to learn what to attend to. This idea can be broken down into two straightforward assumptions – 1) learning mechanisms act on an attentionally-filtered representation of the world and 2) the attention filter is dynamically adjusted according to the outcomes of ongoing decisions. Both assumptions have found support in previous empirical work. I will now discuss how the current findings are consistent with or deviate from previous work.

Learning acts on an attentionally-filtered representation of the world

Bayesian inference provides a statistically correct strategy to integrate information across dimensions and across time. If participants are optimal, they ought to rely on a Bayesian strategy to learn and make choices. However, a fully Bayesian approach is likely to be computationally intractable, especially in light of limited cognitive capacity (Daw & Courville, 2008; Kruschke, 2006). Instead, selective attention is thought to be necessary to reduce the number of dimensions to learn about. While this reduces computational demand, it does so at the expense of statistical correctness. By applying computational models to behavioral data, both Gershman et al. (2010) and Wilson & Niv (2011) demonstrated that participants do indeed trade statistical optimality for computational efficiency when solving a multi-dimensional probabilistic decision-making task. The results of Experiment 1 replicated this finding. The Bayesian model provided the worst fit to behavioral data, suggesting that participants do not employ a Bayesian strategy in solving the task. Instead, participants' behavior was best described by function approximation models, which assume that participants incrementally update weights of chosen features (as opposed to maintaining and simultaneously updating a probability distribution over all features on each trial). The model that best described participants' behavior was a variant of the function approximation model that also decays weights on unchosen features. Since the reward probability of each feature does not change during a game, the optimal estimate of the reward probability should take into account all past outcomes. However, this may not be possible given memory constraints. To work within these constraints, participants might have to forget older information. As such, the decay is another way in which participants favor computational efficiency over optimality.

As discussed earlier, the decay rule also implements an implicit selective attention component to the model. If participants are attending and choosing the same feature for multiple trials, that feature would acquire a high weight while the other weights decay to zero. Hence, choice would be determined by the feature that is being attended to. To test this hypothesis, I developed a method to decode attention by combining eye-tracking with pattern classification of fMRI data (Experiment 2). I then used this method to decode the trial-by-trial focus of attention while participants played the FHT task (Experiment 3). I found that on most trials, the decoded focus of attention was on the feature with the highest weight as computed by the Decay model, suggesting that the Decay model was implicitly tracking participants' focus of attention, which might then account for the Decay model's superior performance in predicting choice behavior relative to the regular FA model. As a more direct test that participants were employing selective attention in their strategies, I incorporated a selective attention component to the FA model, which in its original form assumed that participants attended equally to all dimensions. The FA model with selective attention predicted choice better for the majority of the participants than the regular FA model, providing direct evidence that, at least for some participants, selective attention is an important part of their strategy in solving the task.

The above results provide preliminary evidence that learning processes do not act on all aspects of the environment. Instead, selective attention filters the information that is processed. While this is suboptimal, it is computationally efficient and might be a preferable strategy given limited cognitive capacity.

Attention is dynamically modulated according to the outcomes of decisions

In the previous section, I discussed how attention facilitates learning by constraining the information to learn about. However, how does one decide what or where to attend to? As there was an explicit measure of attentional focus in Experiment 3, it was possible to test the factors that influence the focus of attention. In particular, I hypothesized that attention is preferentially directed towards stimuli that have been associated with rewards. In our models, reward association is captured by the weights of each feature. As such, I investigated if the focus of attention is related to the feature weights. The FA model was used in this analysis, as it does not make a priori assumptions about selective attention. Specifically, I counted the proportion of trials on which attention is directed towards the dimension with feature of the highest weight. This proportion was significantly above chance, suggesting that attention was indeed preferentially directed towards features that have been frequently associated with reward. This result suggests that the focus of attention is modulated by reward feedback. However, more detailed analysis needs to be conducted to investigate the dynamics of this modulation.

6.2 Lessons Learnt

Efficiency-Optimality Trade Off

There is a trade-off between computational efficiency and statistical optimality. In the current series of experiments, I showed that participants tended to favor computational efficiency over statistical optimality. The statistically optimal strategy might prove too computationally demanding for our cognitive resources. This is especially so in our rich and multidimensional world. Recognizing this trade off can help explain why people are sometimes suboptimal in their choices.

An intricate relationship between learning and attention

The current results suggest that there is an intricate relationship between learning and attention. Attention determines what we learn about, but we also learn what to attend to. However, the specific nature of this relationship remains to be explored. This calls for an integrated approach that brings together existing methods in the study of attention and learning. As more researchers recognize the importance of this relationship, we can hope to see a more concerted effort in charting out the mechanisms by which learning and attention interact in the near future.

A promising experimental approach

Perhaps the most significant contribution of this thesis is the development of an experimental approach that is suited to study the interaction between attention and learning. A multidimensional decision-making task like the FHT task creates a simplified task environment which nevertheless mimics the cluttered multidimensional state of the real world. As is the case in the real world, only a subset of the dimensions is relevant for the task at hand. We can then analyze participants' strategies in solving this task and make specific inferences about how people deal with multidimensional information in the real world. The FHT task is particularly appealing because it operationalizes dimensions as spatially separated object categories. As such, the focus of attention can be decoded using eye-tracking and pattern classification of fMRI data. In Experiment 2, I showed that this is a viable method of tracking attention on a trial-by-trial basis. This focus of attention can then be used as input to different computational models that are applied to explain participants' behavior. The models can also make predictions about how attention is modulated by ongoing choices. Importantly, these computational models operate at the trial-by-trial level, and would have sufficient temporal resolution to investigate how attention and learning interact on each trial. While this approach is still in its infancy, it holds much promise in charting out the mechanisms underlying the interactions between attention and learning.

6.3. The Way Forward

Collecting More Data

Both Experiment 2 and 3 suffer from small sample sizes. Perhaps the most immediate course of action moving forward would be to collect data from more participants for those experiments. It would be especially interesting to investigate if there is indeed a split between the strategies employed by participants such that the behavior of some participants are better described by a model with selective attention (FA model) while that of others are better described by a model without selective attention (FA_SA model).

Refining the Computational Models of Choice Behavior

The models can be refined to better match participants' behavior. One promising approach is to investigate different methods of combining the eye-tracking results with pattern classifier outputs to obtain the composite measure of attention. For example, instead of taking the normalized product of the two, it might make sense to take a weighted average. Each measure can be weighted by its relative accuracy in predicting attention during the 3D1B task. This would allow the composite measure to capture individual differences in how much each participant relies on covert vs. overt attention processes. Another possible modification is to allow for the dissociation between attention during choice and attention during learning. While the temporal resolution of fMRI might not allow for such a distinction, eye-tracking might provide us with a measure precise enough to differentiate between the two.

Building a Computational Model of Attention Modulation

To better study the dynamics of attention modulation, it would be helpful to build a computational model of how attentional focus changes at a trial-by-trial level given feature weights. It would be particularly interesting to see if a switch in the focus of attention can be predicted by changes in feature weights. Such a result would provide stronger evidence that attentional focus is modulated by the outcomes of ongoing learning.

Searching for the Neural Correlates of the Interaction between Learning and Attention

Given a satisfactory model of behavior, we can proceed to search for the neural correlates of the interaction between learning and attention. Specifically, the model can be used to generate hidden variables that track the dynamics of internal attention or learning processes. Regressing these variables against neural data would then reveal the brain areas that take part in computing or signaling these quantities. Some interesting internal variables include the prediction-error term posited by reinforcement learning algorithms or the broadness of attentional focus. It would be particularly interesting to investigate how the neural structures traditionally implicated in learning interact with the fronto-parietal attention network. As I have fMRI data of participants performing the FHT task, I would be able to run this analysis in the near future.

References

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual review of neuroscience*, *25*, 189–220. doi:10.1146/annurev.neuro. 25.112701.142922

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011a). Value-driven attentional capture. *Proceedings of the National Academy of Sciences,* 108: 10367–10371.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Learned value magnifies salience-based attentional capture. *PLoS ONE, 6*(11), e27926.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature reviews. Neuroscience*, *4*(10), 829–39. doi:10.1038/nrn1201

Balleine, B. W., Daw, N. D., & O'Doherty, J. (2009). Multiple forms of value learning and the function of dopamine. In P. W. Glimcher, C. F. Camerer, C. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision Making and the Brain* (pp. 367–387). London: Academic Press.

Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature neuroscience*, *8*(8), 1110–6. doi:10.1038/nn1501

Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., et al. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, *60*(6), 1142–52. doi:10.1016/j.neuron.2008.09.021

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature neuroscience*, *10*(9), 1214–21. doi:10.1038/nn1954

Bellman, R. (1957). *Dynamic Programming*. Princeton, NJ: Princeton University Press

Brass, M., & Von Cramon, D. Y. (2004). Selection for cognitive control: a functional magnetic resonance imaging study on the selection of task-relevant information. *The Journal of* *neuroscience : the official journal of the Society for Neuroscience*, *24*(40), 8847–52. doi:10.1523/JNEUROSCI.2513-04.2004

Cheal, M., & Lyon, D. R. (1991). Central and peripheral precuing of forced-choice discrimination. *The Quarterly Journal of Experimental Psychology Section A*, *43*(4), 859–880. doi:10.1080/14640749108400960

Daw, N. (2009). Trial-by-trial data analysis using computational models. In *Affect, Learning and Decision Making, Attention and Performance XXIII*, E.A. Phelps, T.W. Robbins, & M. Delgado (Eds.). New York, NY:Oxford University Press

Daw, N., & Courville, A. (2008). The pigeon as particle filter. *Advances in neural information processing systems, 20*, 369–376.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current opinion in neurobiology*, *16*(2), 199–204. doi:10.1016/j.conb.2006.03.006

Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention. *Nature neuroscience*, *3,* 1218-1223. doi:10.1038/81504

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *353*(1373), 1245–55. doi:10.1098/rstb.1998.0280

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual review of neuroscience*, *18*, 193–222. doi:10.1146/annurev.ne.18.030195. 001205

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827–1837. doi:10.1016/0042-6989(95)00294-4

Doya, K. (2008). Modulators of decision making. *Nature neuroscience*, *11*(4), 410–6. doi:10.1038/nn2077

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of experimental psychology. General*, *113*(4), 501–17.

Ettenberg, A. (1989). Dopamine, neuroleptics and reinforced behavior. *Neuroscience & Biobehavioral Reviews*, *13*(2-3), 105–111. doi:10.1016/S0149-7634(89)80018-1

Fassbender, C., Simoes-Franklin, C., Murphy, K., Hester, R., Meaney, J., Robertson, I. H., & Garavan, H. (2006). The role of a right fronto-parietal network in cognitive control. *Journal of Psychophysiology, 20*(4), 286-296. doi: 10.1027/0269-8803.20.4.286

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science (New York, N.Y.)*, *299*(5614), 1898–902. doi:10.1126/science.1077349

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science (New York, N.Y.)*, *291*(5502), 312–6. doi:10.1126/science.291.5502.312

Gershman, S.J., Cohen, J.D., and Niv, Y. (2010). Learning to selectively attend. In *32nd Annual Conference of the Cognitive Science Society,* Portland

Gottlieb, J. (2012). Attention, learning, and the value of information. *Neuron*, *76*(2), 281–295. doi:10.1016/j.neuron.2012.09.034

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, *14*(5), 462–466. doi:10.1111/1467-9280.02454

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V, & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*, *7*(1), 37–53. doi:10.1007/s12021-008-9041-y

Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science (New York, N.Y.)*, *293*(5539), 2425–30. doi:10.1126/science.1063736

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *The Journal of Neuroscience, 30*(33), 11096-11103.

Jacobs, R. A., & Kruschke, J. K. (2011). Bayesian learning theory applied to human cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(1), 8–21. doi:10.1002/wcs.80

James, W. (1890). *The principles of psychology, vol I*. New York, NY, US: Henry Holt and Co.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: consensus and controversy. *Neuron*, *63*(6), 733–45. doi:10.1016/j.neuron.2009.09.003

Kastner, S., DeWeerd, P., Desimone, R., & Ungerleider, L.G. (1998). Mechanisms of directed attention in ventral extrastriate cortex as revealed by functional MRI. *Science*, 282: 108-111.

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*(4), 751–761. doi:10.1016/S0896-6273(00)80734-5

Kleiner M., Brainard D., Pelli D. (2007) What's new in Psychtoolbox-3? *Perception, 36*. ECVP Abstract Supplement

Knudsen, E. I. (2007). Fundamental components of attention. *Annual review of neuroscience, 30*, 57-58. doi:10.1146/annurev.neuro.30.051606.094256

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897–1916. doi:10.1016/0042-6989(94)00279-U

Kruschke, J. K. (2006). Locally bayesian learning with applications to retrospective revaluation and highlighting. *Psychological review*, *113*(4), 677–99. doi:10.1037/0033-295X.113.4.677

Lavie, N. (2005). Distracted and confused? Selective attention under load. *Trends in cognitive sciences*, *9*(2), 75–82. doi:10.1016/j.tics.2004.12.004

Lavie, N., Tsai, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Percept. Psychophys.*, 56: 183-197.

Lennie, P. (2003). The cost of cortical computation. *Current Biology*, *13*(6), 493–497. doi:10.1016/S0960-9822(03)00135-0

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J Neurophysiol*, *77*(1), 24–42.

MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203.

Mahon, B. Z., & Caramazza, A. (2008). Concepts and categories: A cognitive neuropsychological perspective. *Annual review of psychology, 60*, 27-51. doi:10.1146/annurev.psych.60.110707.163532

Martin, A. (2007). The representation of object concepts in the brain. *Annual review of psychology*, *58*, 25–45. doi:10.1146/annurev.psych.57.102904.190143

Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A bayesian foundation for individual learning under uncertainty. *Frontiers in human neuroscience*, *5*, 39. doi:10.3389/fnhum.2011.00039

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in neurosciences*, *29*(6), 317–22. doi:10.1016/j.tins.2006.04.001

McAdams, C. J., & Maunsell, J. H. R. (1999). Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. *J. Neurosci.*, *19*(1), 431–441.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal Prediction Errors in a Passive Learning Task Activate Human Striatum. *Neuron*, *38*(2), 339–346. doi:10.1016/S0896-6273(03)00154-5

McMains, S. A., Fehd, H. M., Emmanouil, T.-A., & Kastner, S. (2007). Mechanisms of feature- and space-based attention: response modulation and baseline increases. *Journal of neurophysiology*, *98*(4), 2110–21. doi:10.1152/jn.00538.2007

Miller, E K, & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, *24*, 167–202. doi:10.1146/annurev.neuro.24.1.167

Miller, Earl K, Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: categories, concepts and cognition. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *357*(1424), 1123–36. doi:10.1098/rstb.2002.1099

Milner, B. (1963). Effects of Different Brain Lesions on Card Sorting: The Role of the Frontal Lobes. *Archives of Neurology*, *9*(1), 90–100. doi:10.1001/archneur.1963.00460070100010

Montague, P., Dayan, P., & Sejnowski, T. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.*, *16*(5), 1936–1947.

Montague, P. R., Dayan, P., Person, C., & Sejnowski, T. J. (1995). Bee foraging in uncertain environments using predictive hebbian learning. *Nature*, *377*(6551), 725–8. doi:10.1038/377725a0

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782–784. doi:10.1126/science.4023713

Morris, G., Arkadir, D., Nevet, A., Vaadia, E., & Bergman, H. (2004). Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron*, *43*(1), 133–43. doi:10.1016/j.neuron.2004.06.012

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

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006a). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, *10*(9), 424–30. doi:10.1016/j.tics.2006.07.005

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006b). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, *10*(9), 424–30. doi:10.1016/j.tics.2006.07.005

Owen, A. M., Roberts, A. C., Hodges, J. R., & Robbins, T. W. (1993). Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain*, *116*(5), 1159–1175. doi:10.1093/brain/116.5.1159

O'Reilly, J. X., Jbabdi, S., & Behrens, T. E. J. (2012). How can a Bayesian approach inform neuroscience? *The European journal of neuroscience*, *35*(7), 1169–79. doi:10.1111/j.1460-9568.2012.08010.x

Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R., & Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. *The Journal of Neuroscience, 29*(36), 11182-11191.

Peelen, M. V, Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, *460*(7251), 94–7. doi:10.1038/nature08103

Peelen, M. V, & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(29), 12125–30. doi:10.1073/pnas.1101042108

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage*, *45*(1 Suppl), S199–209. doi:10.1016/j.neuroimage.2008.11.007

Posner, M.I. (1980). Orienting of attention. *Quart. J. Psychol.*, 32: 3-25.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64). New York, NY: Appleton-Century-Crofts.

Rees, G., Frith, C.D., Lavie, N. (1997) Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*, 278: 1616-1619.

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive psychology*, *51*(1), 1–41. doi:10.1016/j.cogpsych.2004.11.001

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive Mechanisms Subserve Attention in Macaque Areas V2 and V4. *J. Neurosci.*, *19*(5), 1736–1753.

Reynolds, J. H., & Desimone, R. (2003). Interacting Roles of Attention and Visual Salience in V4. *Neuron*, *37*(5), 853–863. doi:10.1016/S0896-6273(03)00097-7

Sanes, J. R., & Lichtman, J. W. (2001). Induction, assembly, maturation and maintenance of a postsynaptic apparatus. *Nature reviews. Neuroscience*, *2*(11), 791–805. doi:10.1038/35097557

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current opinion in neurobiology*, *20*(2), 172–6. doi:10.1016/j.conb.2010.02.010

Schultz, W. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593–1599. doi:10.1126/science.275.5306.1593

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.*, *13*(3), 900–913.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464.

Seidl, K. N., Peelen, M. V, & Kastner, S. (2012). Neural evidence for distracter suppression during visual search in real-world scenes. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *32*(34), 11812–9. doi:10.1523/JNEUROSCI.1693-12.2012

Sutton, R.S. & Barto, A.G. (1990) Time derivative models of Pavlovian reinforcement. In M. Gabriel & J. Moore (Eds.) *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 539-602). Cambridge, MA: MIT Press.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Treue, S., & Martínez Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*(6736), 575–9. doi:10.1038/21176

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

Wilson, R. C., & Niv, Y. (2011). Inferring relevance in a changing world. *Frontiers in human neuroscience*, *5*, 189. doi:10.3389/fnhum.2011.00189

Wise, R. A. (1982). Neuroleptics and operant behavior: the anhedonia hypothesis. *Behav Brain Sci*, *5*(3), 87.

Wise, R., Spindler, J., deWit, H., & Gerberg, G. (1978). Neuroleptic-induced "anhedonia" in rats: pimozide blocks reward quality of food. *Science*, *201*(4352), 262–264. doi:10.1126/science.566469

Woolgar, A., Thompson, R., Bor, D., & Duncan, J. (2011). Multi-voxel coding of stimuli, rules, and responses in human frontoparietal cortex. *NeuroImage*, *56*(2), 744–52. doi:10.1016/j.neuroimage.2010.04.035

Wunderlich, K., Beierholm, U. R., Bossaerts, P., & O'Doherty, J. P. (2011). The human prefrontal cortex mediates integration of potential causes behind observed outcomes. *Journal of neurophysiology*, *106*(3), 1558–69. doi:10.1152/jn.01051.2010

Yang, T., & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, *447*(7148), 1075–80. doi:10.1038/nature05852

Yu, A. J., Dayan, P., & Cohen, J. D. (2009). Dynamics of attentional selection under conflict: toward a rational Bayesian account. *Journal of experimental psychology. Human perception and performance*, *35*(3), 700–17. doi:10.1037/a0013553

Collaboration in Senior Thesis Work
The Relationship between the Senior Thesis and Earlier Work

Please use this form to indicate the relationship between previous work and your senior thesis and to indicate whether your thesis involved collaboration with others.

Indicate below whether there is any overlap between your senior thesis and earlier work that you did for junior reports, junior papers, or papers for various courses.

Overlap __✔___

No overlap __ _____

If you checked the box indicating that there is overlap between your senior thesis and previous work, please describe the overlap on a separate page, and include it within the thesis after this form.

Readers of your thesis may, if they choose, ask to see earlier papers that you indicate have some overlap with your senior thesis.

Indicate below whether all or part of your thesis resulted from work done collaboratively with one or more other people.

Collaboration __✔_

No collaboration ______

If you checked the box indicating that your thesis work was done entirely, or in part, in collaboration with other people, describe the nature of the collaboration and what resulted from it on a separate page, and include it within the thesis after this form.

Overlap Between the Senior Thesis and Previous Work

Data from Experiment 1 were collected during my time as an undergraduate research assistant in the Niv Lab from Sophomore summer through Junior year. Some preliminary results of Experiment 1 were presented as a poster at the 2012 Society for Neuroscience Conference. This thesis, however, is the first time the results have been reported in writing.

Nature of Collaboration

This thesis is part of a larger ongoing project in the Niv Lab. As such, the specific experimental hypotheses and experimental design were formulated through discussion with different people in the lab, including Professor Yael Niv, Reka Daniel, Angela Radulescu and Andra Geana. Many of the analysis methods were also inspired by earlier work (e.g., Gershman, Cohen & Niv, 2010; Wilson & Niv, 2011; see reference page). The FHT task used in experiments 1 and 3 was first developed by Vivian DeWoskin for her senior thesis, though I have since made several modifications to the task.

I was involved in the design and execution of the experiments. I collected the behavioral data for Experiment 1. fMRI data for experiment 2 and 3 were collected with the help of Angela Radulescu. I analyzed the data using either scripts I wrote myself or scripts I obtained from other members in the lab.

Research Involving Human Subjects

Did your Senior Thesis involve research with human subjects? Yes

If your Senior Thesis DID involve research with human subjects, please indicate your IRB Case Number below.

5291 Approval Date: 3/25/2012