Optimal allocation of attention in a signal detection task

A Thesis

submitted to

Indian Institute of Science Education and Research Pune

in partial fulfillment of the requirements for the

BS-MS Dual Degree Programme

by

Sahiti Chebolu

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

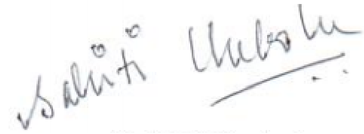
Pashan, Pune 411008, INDIA.

July, 2021

Supervisor: Dr. Kevin Lloyd

Certificate

This is to certify that this dissertation entitled " Optimal allocation of attention in a signal detection task " towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Sahiti Chebolu at Indian Institute of Science Education and Research under the supervision of Dr. Kevin Lloyd, Senior Research Scientist, MPI for Biological Cybernetics, during the academic year 2020-2021.



Sahiti Chebolu

Dr. Kevin Lloyd

Committee:

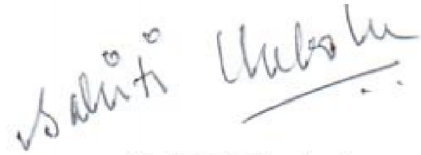
Dr. Kevin Lloyd

Prof. Collins Assisi

This thesis is dedicated to the silenced and the unheard

Declaration

I hereby declare that the matter embodied in the report entitled " Optimal allocation of attention in a signal detection task " are the results of the work carried out by me at the MPI for Biological Cyberneticsand Indian Institute of Science Education and Research, Pune, under the supervision of Dr. Kevin Lloyd and the same has not been submitted elsewhere for any other degree.



Sahiti Chebolu

Dr. Kevin Lloyd

Acknowledgments

First and foremost, I thank my supervisor, Dr. Kevin Lloyd for his invaluable scientific advice and unparalleled guidance. I thank Prof. Peter Dayan for hosting me in his lab and his insightful comments and suggestions throughout the project. I express my gratitude to Max Planck Institute for Biological Cybernetics for the infrastructure and financial support and also my home institution, IISER Pune for the fantastic training and career opportunities throughout the last five years. I also thank my expert advisor, Prof. Collins Assisi for his timely advice and guidance. My sincere thanks goes to the KVPY for the scholarships to aid my education. I am ever grateful to my parents, grandparents and brother for their constant encouragement, support and love. I thank my boyfriend for always being there for me and my friends and lab-mates from Pune and Tubingen for their company and camaraderie.

Abstract

Detection of temporally uncertain signals is a common motif in many real world and laboratory settings including tasks that require sustained attention. While the role of attention in these tasks has been studied extensively, there have not been many normative accounts of attentional allocation in such tasks. In this study, we investigated optimal behavior in a signal detection task with uncertain signal onset where we allowed attention to improve the quality of sensory information collected. However, paying attention came at a cost and so, attention had to be allocated wisely. Using dynamic programming, we estimated an optimal policy for allocating attention within each trial of the task. Interestingly, we found that a rational agent must pay attention only when there is enough (but not overwhelming) evidence in favor of a signal. Further, for the same amount of evidence, it was optimal to pay more attention later in the trial. Reward-cost trade-offs dictated that when a trial was too tough or when there was too much bias towards a certain hypothesis, there is no advantage in paying attention. The performance (as measured by the sensitivity index, d') was a result of complex interactions between factors like signal length, signal probability and attention costs. It increased with signal length, decreased with attention costs only at short signal lengths and remained unchanged with signal probability (despite a shift in the response criterion). Equivalent results have been shown experimentally in sustained attention tasks which are known to involve a similar temporal uncertainty.

Contents

List of Figures

Chapter 1

Introduction

1.0.1 Signal Detection

The real world is replete with instances where one has to make choices based on incomplete and imperfectly reliable information about the world. The signal detection paradigm provides a framework for modelling and analysing decision-making under uncertainty in a wide range of tasks and scenarios [Green and Swets, 1966]. A typical task may involve indicating whether a signal stimulus was present or absent on a given trial. Even in this simplest of settings, a variety factors will influence the decision-maker, such as the stimulus strength, the expected probability of a signal, and the relative payoffs for being correct or incorrect. Signal detection theory provides both a normative account of how decisions should be made in such settings, and a suite of tools to assess actual behaviour.

In this study, we are particularly interested in how attentional processes affect performance in signal detection tasks, particularly those involving temporal uncertainty about when a signal may arrive – if it arrives at all – within a trial. How to optimally integrate sensory evidence across a trial becomes surprisingly complex in the face of such temporal uncertainty. That human subjects find such uncertainty a challenge is suggested by experiments showing that even a seemingly modest amount of uncertainty about signal timing leads to a significant performance decrement [James P. Egan and Schulman, 1961, Green and Swets, 1966].

The challenge of temporal uncertainty to detection performance is central in *sustained attention* (or 'vigilance') tasks. These canonically involve prolonged periods of signal absence interleaved with short, weak signals occurring at unpredictable times; to perform well, a subject is taxed with having to continuously attend to the display in order to detect these unpredictable signals [Parasuraman, 1979]. A brief review of the empirical findings in these tasks therefore provides a perfect background for the current investigation into the role and effects of attention in detecting temporally unpredictable signals.

1.0.2 Behaviour in Sustained Attention tasks

Behavior in sustained attention tasks has been characterised in detail, both in humans [Warm et al., 2008] and rodents [McGaughy and Sarter, 1995]. A recurring finding is that there is a drop in detection performance over time on the task, a phenomenon referred to as 'vigilance decrement'. Early theories attributed this decrement to a lack of arousal due to low cognitive demands, due to both the scarcity of signals and repetitive nature of the task [Warm et al., 2008]. However, to the contrary, more recent studies suggest that sustained attention tasks are demanding on information-processing resources, and that increasing perceptual or cognitive demands in these tasks actually tends to worsen the decrement [Warm et al., 2008, Sarter et al., 2001].

Manipulation of several experimental variables have been shown to affect performance on sustained attention tasks. One such is *memory load*. In signal detection tasks, signals can be detected successively or simultaneously. Successive judgements involve distinguishing signal events from non-signal events using information stored in memory. For simultaneous judgements, all the information needed to detect a signal is found in the stimulus itself, leading to no memory requirement. It has been found that successive judgements lead to a greater vigilance decrement than simultaneous judgements [Warm et al., 2008]. Another way of increasing the memory load is to explicitly pair vigilance tasks with working memory tasks. [Caggiano and Parasuraman, 2004] showed that performance in a spatial vigilance task showed greater decrement with time when paired with a spatial working memory task.

Another important variable is the *event rate*, which refers to the rate of appearance of stimulus events among which the targets may occur. Larger event rates cause a reduction in hit rates (i.e., correct detections) and this has been argued to be due not to a decrease in sensitivity (i.e., ability to discriminate signal from non-signal), but due to a shift in response criterion (i.e., a increased response bias towards not reporting a signal) [Parasuraman et al., 1987].

Further variables of importance are signal salience/discriminability and event uncertainty. For example, if the contrast between a signal and it's background is decreased, so that it is harder to discriminate, then there is a greater vigilance decrement over time [Helton and Warm, 2008]. Similarly, making the time at which a signal appears more unpredictable ('event asynchrony') or increasing spatial uncertainty about where a signal will appear leads to greater vigilance decrement [Warm et al., 2008].

These results hint at which factors in vigilance tasks may be particularly important in engaging attention and information processing systems of the brain. They also constrain explanations for mechanisms that might be leading to the observed behavior. A leading theory is that sustaining attention for long periods leads to depletion of finite attentional/information-processing resources, and hence cause a decline in performance [Thomson et al., 2015]. This theory predicts that increasing work load through changing the different task factors mentioned above should lead to larger decrements due to a greater rate of resource depletion [Thomson et al., 2015].

However, it has also been shown that not all of these performance declines are due to a sensitivity decrement (that is, a decrement in signal detectability). Some of these effects are due to the shifts in the response bias (that is, the threshold for responding in favor of signal or non-signal hypotheses) [Parasuraman et al., 1987]. Further, not all manipulations of task difficulty lead to sensitivity decrements. For example, a reliable sensitivity decrement is observed only when a signal quality is low and the memory load is high (and not, for example, when memory load is increased at high signal quality) [Parasuraman et al., 1987]. Previous studies have not explained why resource depletion should lead to a diverse range of effects. Why do some manipulations of task difficulty result only in shifts in response criterion without an actual change in the sensitivity?

1.0.3 Attention

Sustained attention tasks, as their name suggests, involve the engagement of attentional processes for a long period of time. But how, particularly if such attention carries some cost, should attention be most effectively deployed in such tasks in order to successfully detect signals? Reviewing previous findings and theories about the effects of attention, especially in sustained attention tasks, may better equip us to locate the role of attention in our chosen signal detection task. We start by reviewing studies about the regulation and effects of attention in general, and then move to neuroscientific studies involving sustained attention tasks that particularly implicate the neuromodulator acetylcholine in the engagement of attention.

In simple terms, attention involves selecting particular information to concentrate on while ignoring the rest [James, 1890]. Such selection is thought to be necessary because there are limits on how much information brains can process in parallel at a time, commonly referred to as a 'bottleneck' [Anderson, 2004]. If information processing resources are so limited, then they must be 'allocated' strategically – the brain must select what to attend to and when.

The distribution of attention depends on 'top-down' and 'bottom-up' factors [Posner MI, 1990, Anderson, 2004]. Top-down regulation involves the use of prior knowledge of the task structure, parameters, prior expectations and future goals to direct attention to relevant stimuli at appropriate locations and times [Posner MI, 1990]. On the other hand, bottom-up processes are stimulus-driven shifts in attention to a salient target leading to detection [Posner MI, 1990, Sarter and Lustig, 2019].

What are the effects of concentrating attention on something particular like a stimulus or a location in space? There is a wealth of evidence that attention can improve behavioral performance by boosting perception. For example, in the well known Posner cueing paradigm, when the location of a target is cued in advance, this results in a faster response and also better detection of weak stimuli, indicating faster and enhanced processing [Posner et al., 1980, Prinzmetal et al., 2005]. It also has been shown that attention improves discriminability in simpler detection tasks [Moray et al., 1976]. In tasks requiring visual search of a target among distractors, enhancing attention by cueing or providing prior information improves performance and signal discriminability while impairing attention has an opposite effect [Verghese, 2001].

At a neural level, this might be an effect of improving neural activity to the target while inhibiting neurons responding to the distractors [Verghese, 2001]. [Cohen and Maunsell, 2010] suggest that attention improves reliability of sensory information encoded by neuronal populations by decreasing correlated variance. In addition, many other studies have demonstrated this improvement in signal-to-noise ratio that attention has at a neurophysiological level [Assad, 2003, Cohen and Maunsell, 2009, Reynolds and Chelazzi, 2004].

Similar effects of enhanced perception due to attention have also been shown specifically in tasks involving temporal uncertainty. For example, [Ghose and Maunsell, 2002] showed that sensory neuron responses in the visual cortex are modulated by attention depending on the expectancy of a visual stimulus – there is greater attentional modulation when a stimulus is expected with a high probability. [Jaramillo and Zador, 2011] showed that valid temporal expectations of an auditory signal improved speed and accuracy of responses and that this is mediated by enhanced representation of sounds in the auditory cortex when a signal is highly expected.

On the flip side, focusing attention on one particular source of information may come at the cost of attenuating, or ignoring entirely, other sources that may carry important information. For example, for an animal foraging in the wild, concentrating on searching for food on the ground may reduce its ability to monitor for aerial predators, thereby incurring greater predation risk. Indeed, behavioral ecologists have found evidence that this is the case [Clark and Dukas, 2003]. It has also been argued, particularly in economics, that allocating limited resources to a particular task or activity can lead to implicit 'opportunity costs', which are benefits lost from resources not dedicated to other tasks [Mackowiak and Wiederholt, 2009]. In addition, information processing during attention, including evidence accumulation and perception, detection and response selection come with physiological costs and are metabolically expensive [Laughlin et al., 1998, Gaulin and McBurney, 2003].

Neuromodulation and Sustained Attention

The neuromodulator acetylcholine (ACh) has been widely implicated in attentional control. Cholinergic signaling appears to operate at multiple timescales, with both a slowly varying 'tonic' mode of signalling, and a rapidly changing 'phasic' mode. It has been suggested that the former might support top-down stabilisation of attentional function, for example in response to expected changes in attentional challenges [Sarter and Lustig, 2019], while the latter appears to be more involved in detection of signals at faster, trial-level timescales [Sarter et al., 2016, Sarter and Lustig, 2019].

The earliest evidence for a role of ACh in sustained attention comes from lesion studies. For example, [McGaughy et al., 1996] in their Sustained Attention Task (SAT) specifically adapted for rodents, showed that animals with lesions of cholinergic neurons in the basal forebrain showed much worse detection performance (as measured by their sensitivity, d') than normal animals. While lesion studies demonstrated the necessity of ACh for attentional function, other evidence has come from studies that monitor extracellular ACh concentrations during attentional tasks. [Arnold et al., 2002] used microdialysis to compare ACh efflux in rats performing the SAT to two control tasks which had similar sensory and motor demands as the SAT but which did not specifically engage attention. The ACh efflux was significantly higher on the SAT compared to the other two tasks.

More recently, techniques have been developed which allow the measurement of ACh concentrations at faster timescales, and these have led to new findings that suggest an important role for fast ACh signalling in attention and signal detection. For example, using a newly developed amperometric technique, [Howe et al., 2013] observed transient ACh signals at a sub-second resolution that correlated with hit trials in the SAT. Intriguingly, this was not observed on all hit trials, but only on those that were preceded by misses or correct rejections. One speculation is that ACh transients might be mediating signal detection by shifting from a state of monitoring for signals to a cue-activated state that could help execute the appropriate, cue-associated response. Using optogenetics, [Gritton et al., 2016] showed that generating transients in the medial prefrontal cortex on signal trials improved detection of signals while generating transients on non signal trials increased false alarm rates. Furthermore, inhibiting the generation of transients on signal trials reduced the number of hits. This suggests a causal role of ACh transients in signal detection.

Since we are particularly interested in the dynamics of attention in sustained attention tasks, it is relevant to understand how ACh might be mediating signal detection at the neural level.

One idea is that ACh acts to boost thalamic input to the cortex and suppress input from other cortical areas, and hence facilitates processing of external stimuli [Hasselmo and Sarter, 2011]. ACh has also been shown to enhance the response of neurons to sensory input hence enhancing signal to noise ratios in sensory cortices [Newman et al., 2012]. ACh signals may contribute to active maintenance of stimuli and/or relevant task sets by enhancing persistent spiking activity (activity that persists for a period after input stimulation) in the cortex [Hasselmo and Sarter, 2011].

This rich set of data showing how ACh can mediate signal detection and attention shifts on a fast timescale motivates us to address the question of attentional regulation within a trial in our chosen signal detection tasks.

1.0.4 Aim of the study

In this study, our central goal is to investigate the computational problem of how attention should be “optimally allocated" in signal detection tasks with temporally uncertain signals.

Firstly, we aim to construct normative models of behavior in these tasks. Normative models aim to find the optimal solution to a given problem. For example, in a standard signal detection task, an optimal decision is one that maximises accuracy (or more generally, expected reward) in the task. Hence, these models can tell us how one ought to behave in a given case. We start by discussing the optimal inference and decisions in a signal detection task with unpredictable signal onset.

However, in reality, optimal decision making not only involves maximising final performance but also taking into account the various constraints there exist on resources and capabilities of a subject/ animal. In other words, optimal behavior must trade-off rewards and costs. This logic of ‘resource-rationality’ is growing in popularity in neuro-cognitive modelling due to its more realistic assumptions about animals’ abilities and its generation of possible explanations for apparently irrational behaviour [Lieder and Griffiths, 2020]. Taking the particular case of attention, it is often considered to be a limited faculty; some have viewed it as a finite resource that can be depleted [Parasuraman, 1979] while others have have associated some costs with paying attention as we have argued in the previous section.

In this work, we conduct a resource-rational analysis of the sustained attention task: attention is conceptualized as a mechanism that improves sensory evidence but comes at a cost. We ask how attention is deployed optimally so as to maximise performance-related reward while also minimising the attention costs incurred.

Chapter 2

Methods

2.1 HMM Model

We formulate a generative model for a generic signal detection task that comprises signal and non-signal trials occurring with probabilities p_{signal} and $1 - p_{signal}$. Each trial has a fixed number of N time steps. Non-signal trials do not contain a signal while signal trials have a short signal embedded in them. The task is to report whether a signal was present (hypothesis H_1) or absent (hypothesis H_0).

We construct a Hidden Markov Model (HMM) to characterise the task. There are three underlying states (denoted as X_t at time step t): pre-signal state (0), signal state (1) and post-signal state (2). These are not directly accessible to the agent which instead, receives observations Y_t emitted by the latent states. Observations are discrete binary variables drawn from a Bernoulli distribution (dependent on the hidden states):

$$\begin{aligned} P(Y_t = 0|X_t = 0) &= \eta_0 \\ P(Y_t = 1|X_t = 1) &= \eta_1 \\ P(Y_t = 0|X_t = 2) &= \eta_2 \end{aligned} \tag{2.1}$$

For simplicity, $\eta_{0,1,2} = \eta$. In other words, and as will be demonstrated below, η controls the reliability of sensory evidence about whether there is currently a signal or not.

We assume there is a uniform probability of the signal occurring anywhere in the trial. This leads to the hazard function, $r(t)$ for the start time of the signal:

$$
\begin{aligned}
r(t) &= \frac{\text{probability density function}}{1 - \text{cumulative density function}} \\
&= \frac{\frac{p_{signal}}{N}}{1 - \frac{(t-1)\cdot p_{signal}}{N}} \\
&= \frac{1}{\frac{N}{p_{signal}} - t + 1}
\end{aligned}
\tag{2.2}
$$

where t is the time step within a trial. The signal is not assumed to have a fixed length. Instead, we introduce a constant probability, q per time step of turning off. Once the signal turns off, it cannot come on again. The reason we do this is because with a fixed signal length, the belief update is not Markovian. Non-Markovian systems were not desirable for us, especially when we do dynamic programming to optimise rewards and costs (this will be described in the next few sections). These rules specify the transition probabilities, $P(X_t|X_{t-1})$ between the hidden states as summarised in the table below:

		X_t		
	$P(X_t\|X_{t-1})$	0	1	2
	0	1-r(t)	r(t)	0
X_{t-1}	1	0	1-q	q
	2	0	0	1

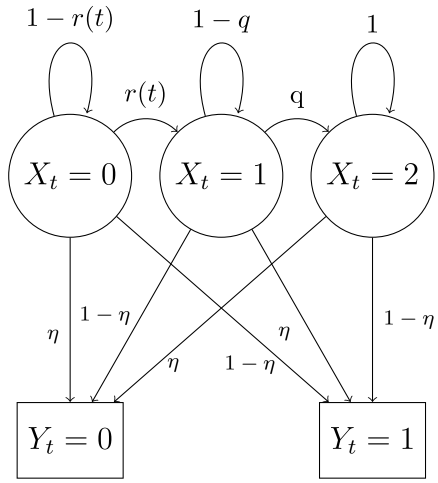


Figure 2.1: A schematic of the Hidden Markov Model

Assuming the agent knows all about the task structure and contingencies, it would then be able to estimate the posterior belief, $P(X_t|Y_{1:t})$ recursively in time through the trial (online inference) using Bayes' rule:

$$\begin{aligned} P(X_t \,|\, Y_{1:t}) &= \frac{P(Y_t \,|\, X_t, Y_{1:t-1}) P(X_t \,|\, Y_{1:t-1})}{P(Y_t \,|\, Y_{1:t-1})} \\ &= \frac{P(Y_t \,|\, X_t) \sum_{X_{t-1}} P(X_t \,|\, X_{t-1}, Y_{1:t-1}) P(X_{t-1} \,|\, Y_{1:t-1})}{P(Y_t \,|\, Y_{1:t-1})} \\ &= \frac{P(Y_t \,|\, X_t) \sum_{X_{t-1}} P(X_t \,|\, X_{t-1}) P(X_{t-1} \,|\, Y_{t-1})}{\sum_{X_t} P(Y_t \,|\, X_t) \sum_{X_{t-1}} P(X_t \,|\, X_{t-1}) P(X_{t-1} \,|\, Y_{1:t-1})} \end{aligned} \tag{2.3}$$

The posterior belief at the end of the trial, $P(X_N|Y_{1:N})$ will then give the most likely state at $t = N$ given all the observations. This should inform the decision of the agent as follows:

$$\text{If } P(X_N = 0 \,|\, Y_{1:N}) > [\, P(X_N = 1 \,|\, Y_{1:N}) + P(X_N = 2 \,|\, Y_{1:N}) \,], \text{ then choose } H_0$$
$$\text{else if } P(X_N = 0 \,|\, Y_{1:N}) < [\, P(X_N = 1 \,|\, Y_{1:N}) + P(X_N = 2 \,|\, Y_{1:N}) \,], \text{ choose } H_1$$

since $P(X_N = 1|Y_{1:N}) + P(X_N = 2|Y_{1:N})$ is the evidence that the signal is or was on, while $P(X_N = 0|Y_{1:N})$ is the evidence in favor of no signal in the trial.

Figure 2.2 shows a graphical representation of the HMM through the time steps.

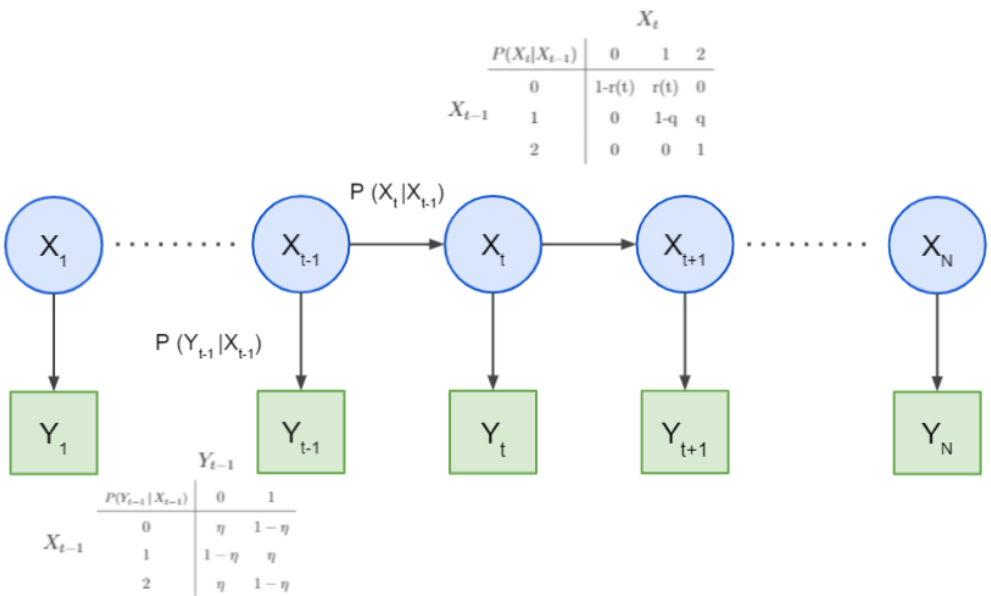


Figure 2.2: A graphical representation of the HMM through the time steps in a trial. The transition function ($P(X_t \mid X_{t-1})$) determines the transition between the latent states. The observations depend on the latent states through the emission probabilities ($P(Y_{t-1} \mid X_{t-1})$)

2.2 Costs of attention

The model above describes the Bayes optimal inference for an agent that can accumulate evidence at no costs – information comes for free. However, it is not unrealistic to consider that there may be costs associated with collecting evidence, and that these may scale with the quality of evidence collected. This is also intuitive – processes like attention presumably improve the quality of evidence but come at a cost as we have argued in Chapter 1. This raises an interesting question: how must an agent optimally modulate it's attention so as to maximise rewards from correct inference while minimising the attentional costs incurred? In order to answer this question, we formalise the notion of modulating attentional states in the following.

Let the belief state b_t be the posterior probability $P(X_t|Y_{1:t})$ at time step t in the trial.

This quantity sufficiently captures the evidence gathered until time step t. In addition to this, we introduce the internal state (IS_t) of the agent. This is defined as a binary variable which controls η_t and hence, the quality of evidence collected at t. $IS_t \in \{L, H\}$ where L is the 'low' state with $\eta = \eta_L$, while H is the 'high' state with $\eta = \eta_H$. H is the more informative state and so $\eta_H > \eta_L$. For simplicity, we let η_t be constant across the hidden states, X_t. Together, this defines the complete state of an agent at time t: (b_t, IS_t). The state space is therefore three dimensional: the belief space has two dimensions, say $b_t(1), b_t(2)$ (which fixes $b_t(0) = 1 - b_t(1) - b_t(2)$) and there is an additional dimension for the binary variable IS_t.

At each time step t, the agent can choose its internal state IS_t by taking an action a_t. Hence a_t are also binary such that $a_t \in \{\text{choose } IS_t = L, \text{ choose } IS_t = H\}$. The outcomes of the actions are assumed to be deterministic, so that taking action $a_t = L$ always leads to $IS_t = L$, and vice versa, irrespective of the current or past states of the agent or the environment. At the end of the final step N, the agent must make a decision in favor of hypothesis H_0 ('trial is a non-signal trial') or in favor of hypothesis H_1 ('trial is a signal trial).

Action a_t is assumed to incur a cost c_{ij} based on the internal state $IS_t = j$ chosen and the preceding internal state $IS_{t-1} = i$. This leads to the following cost table:

		IS_t	
	costs	L	H
	L	c_{LL}	c_{LH}
IS_{t-1}	H	c_{HL}	c_{HH}

On making a decision in favor of H_0 or H_1 at the end of the trial, the agent receives a reward (R_{ij}) based on the true identity of the trial. Correct responses are rewarded while incorrect responses have no consequence:

	R_{ij}	true trial H_0	H_1
	H_0	1	0
response	H_1	0	1

where i stands for the chosen hypothesis H_i and j stands for the true hypothesis H_j.

Given this formulation, our question about how an agent must modulate attention can be re-framed as follows: what is the optimal strategy for choosing actions at each time step t given the belief state and the internal state (b_{t-1}, IS_{t-1})? An optimal strategy is one that maximises the expected net reward. This can be found by maximising the value function, $V^{a^*}(b_t, IS_t)$ which is the expected reward from following optimal behavior in the state at t (b_t, IS_t) and thereafter. This quantity can be expressed recursively as the immediate reward obtained on optimal action (a^*) at t, plus the expected value over the possible next states at $t+1$ (assuming optimal behavior on future time steps):

$$\begin{aligned} V^{a^*}(b_t, IS_t) &= \max_a R(b_t, IS_t, a) + \sum_{b_{t+1}, IS_{t+1}} P(b_{t+1}, IS_{t+1} \mid b_t, IS_t, a) \cdot V^{a^*}(b_{t+1}, IS_{t+1}) \\ &= \max_a [\, Q(b_t, IS_t, a = L), Q(b_t, IS_t, a = H) \,] \end{aligned} \tag{2.4}$$

Here, $Q(b_t, IS_t, a = L)$ and $Q(b_t, IS_t, a = H)$ stand for the immediate plus expected future rewards from following the two possible actions at t, L and H, respectively. $V^{a^*}(b_t, IS_t)$ is the maximum of the two.

This is the Bellman optimality equation for the optimisation problem. Due to the existence of an optimal substructure in the problem, Dynamic Programming can be employed to solve the equation. Solving it gives the optimal value function for each state at each time point t and also the optimal set of actions (known as the optimal policy) that lead to the optimal values. The optimal policy is hence a complete specification of the actions an agent must take for each possible belief state at each point in time in order to maximise net expected reward.

2.2.1 Bellman equation for the task

Equation 2.4 lays out the general expression for the value function for the states of the type (b_t, IS_t) with fixed rewards. To adapt the equation to the task at hand, we will have to specify an expression for the transition function between the belief state-internal state pairs $P(b_{t+1}, IS_{t+1}|b_t, IS_t, a)$ and the reward function $R(b_t, IS_t, a)$. As described above, the reward function consists of costs of choosing an internal state (c_{ij}) as well as the payoffs at the end of the trial (R_{ij}).

The transition between internal states (IS_t to IS_{t+1}) is completely specified by the action (a^*) chosen. The transitions to the future belief state, b_{t+1} depends on the present belief state b_t and possible subsequent observations (Y_{t+1}). This is in turn dependent on the underlying task states (X_{t+1}) and consequently, the structure of the task as specified by the emission and transition probabilities. This sequence of simplifications on the expression is shown below:

$$V^{a^*}(b_t, IS_t) = R(b_t, IS_t, a^*) + \sum_{b_{t+1}, IS_{t+1}} P(b_{t+1}, IS_{t+1}|b_t, IS_t, a^*) \cdot V^{a^*}(b_{t+1}, IS_{t+1}) \tag{2.5}$$

$P(b_{t+1}, IS_{t+1}|b_t, IS_t, a^*))$ can be written as follows:

$$P(b_{t+1}, IS_{t+1}|b_t, IS_t, a^*)) = P(b_{t+1}|IS_{t+1}, b_t, IS_t, a^*) \cdot P(IS_{t+1}|b_t, IS_t, a^*) \tag{2.6}$$

Since IS_{t+1} is fixed given IS_t and $a(= a^*)$, $P(IS_{t+1}|b_t, IS_t, a^*) = 1$ for a single IS_{t+1} and is 0 for the rest. equation 2.4 thus reduces to:

$$V^{a^*}(b_t, IS_t) = R(b_t, IS_t, a^*) + \sum_{b_{t+1}} P(b_{t+1}|b_t, IS_t, a^*) \cdot V^{a^*}(b_{t+1}, IS_{t+1}) \tag{2.7}$$

The quantity $P(b_{t+1}|b_t, IS_t, a^*)$ can be conditioned on the possible future observations (Y_{t+1}) hence giving:

$$P(b_{t+1}|b_t, IS_t, a^*) = \sum_{Y_{t+1}} P(b_{t+1}|b_t, IS_t, Y_{t+1}, a^*) \cdot P(Y_{t+1}|b_t, IS_t, a^*) \tag{2.8}$$

By the Bayesian update equation 2.3, b_{t+1} is completely determined by the current belief, b_t, future internal state I_{t+1} (also given by I_t, a^*) and the future observation, Y_{t+1}. Hence,

$$P(b_{t+1}|b_t, IS_t, Y_{t+1}, a^*) = \begin{cases} 1, & given \quad b_t, Y_{t+1}, I_t, a^* \\ 0, & \text{otherwise} \end{cases} \tag{2.9}$$

Hence, the Bellman equation reduces to:

$$V(b_t, IS_t) = R(b_t, IS_t, a^*) + \sum_{Y_{t+1}} P(Y_{t+1}|b_t, IS_t, a^*) \cdot V(b_{t+1}, IS_{t+1}) \tag{2.10}$$

where both b_{t+1} and IS_{t+1} are fixed. Now, all that remains to be specified is $P(Y_{t+1}|b_t, IS_t, a^*)$ which can be simplified by invoking the task structure (given by the emission and transition probabilities):

$$P(Y_{t+1}|b_t, IS_t, a^*) = \sum_{X_{t+1}} P(Y_{t+1}|X_{t+1}, IS_t, a^*) \cdot \sum_{X_t} P(X_{t+1}|X_t) b_t \tag{2.11}$$

2.2.2 Solving for optimal policy

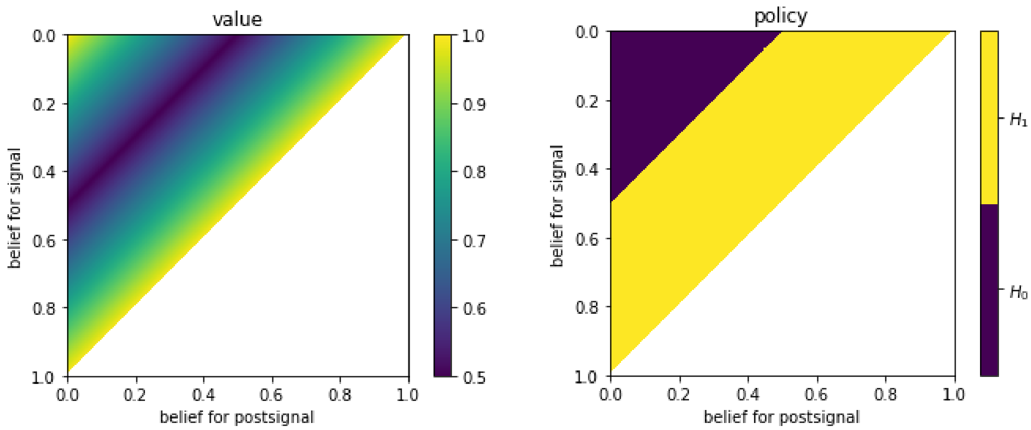
The optimal policy (which is a complete description of the optimal behavior for the agent for each possible state at each time point) can be found by solving the Bellman equation. We assume that the structure of the task, including emission and transition probabilities for the different latent states and the reward and cost functions are known. Hence, equation 2.10 can then be solved for the optimal value and actions at each b_t, IS_t.

We solve the equation numerically by discretising the two-dimensional belief space. The belief states, running from $b = 0$ to $b = 1$ along each axis ($b_t(1)$ and $b_t(2)$ axes), are discretised in steps of $db = 0.01$. This results in a two dimensional grid for the belief states. The other variables (IS, X, Y) are already discrete. Given this completely discrete space, we apply backward induction (dynamic programming) to solve for $V^{a^*}(b_t, IS_t)$ backward in time from $t = N$ to $t = 1$.

For this, $V^{a^*}(b_N, IS_N)$ must be given as an initial condition in order to work backwards. After the final time point, the agent must make a decision in favor of either hypothesis- H_0 : (non-signal trial) or H_1 : (signal trial). This choice is made based on the final belief, b_N and the payoffs, R_{ij} (reward from choosing H_i when H_j is true) according to the following decision rule: choose H_1 if $[b_N(1)+b_N(2)]\cdot R_{11}+b_N(0)\cdot R_{10} > b_N(0)\cdot R_{00}+[b_N(1)+b_N(2)]\cdot R_{01}$ and H_0 otherwise. These quantities are the immediate expected payoffs from choosing H_0 or H_1 (this simply means that the reward from a decision in favor of a hypothesis is weighted by the posterior belief for the hypothesis). Hence, the value at N is the maximum of the expected payoffs for each hypothesis:

$$\begin{aligned} V^{a^*}(b_N, IS_N) = \max_{H_0, H_1} [\, & (b_N(1) + b_N(2)) \cdot R_{11} + b_N(0) \cdot R_{10}, \\ & b_N(0) \cdot R_{00} + (b_N(1) + b_N(2)) \cdot R_{01} \,] \end{aligned} \tag{2.12}$$

Since this doesn't depend on IS_N, the value function at $t = N$ is the same for both internal states. This results in a (2D) V-shaped value function with highest value for extreme belief states $b(X_N = 0) = 1$ or $b(X_N = 1) + b(X_N = 2) = 1$ and the lowest for $b(X_N = 0) = b(X_N = 1) + b(X_N = 2) = 0.5$. This function is shown in a 2D plot in Figure 2.3a. Figure 2.3b shows the corresponding regions where H_0 and H_1 have the highest expected payoff as per equation 2.12. The function is not defined for the lower right triangle of belief states where $b(X = 1) + b(X = 2) > 1$.



(a) Value function after $t = N$ for choosing between H_0 and H_1

(b) Policy function after $t = N$ for choosing between H_0 and H_1

Figure 2.3: Initial condition for the value function (and the corresponding policy) to be used for backward induction. This is obtained by finding the immediate expected payoff from deciding in favor of H_0 or H_1 after the final time step, N

Given this initial $V^{a^*}(b_N, IS_N)$, $V^{a^*}(b_{N-1}, IS_{N-1})$ can be calculated using the Bellman equation 2.10 for each b_{N-1} and IS_{N-1}. $V^{a^*}(b_{N-1}, IS_{N-1})$ can then be used to calculate $V^{a^*}(b_{N-2}, IS_{N-2})$ and so on, all the way back to $t = 1$. The actions that lead to the maximum value at each state form the policy function.

Numerical issues While calculating $V^{a^*}(b_t, IS_t)$ by backward induction, the Bayesian belief update equation 2.3 gives the possible future beliefs b_{t+1}. These future beliefs are then used to get the future values, $V^{a^*}(b_{t+1}, IS_{t+1})$ to be used in the Bellman equation. As the belief space has been discretised, $V^{a^*}(b_{t+1}, IS_{t+1})$ is defined only at the belief values in the 2D grid at steps of $db = 0.01$. Since b_{t+1} obtained from the belief update may not fall on the belief grid, interpolation of the value function (at $t+1$) from the neighboring corners of the grid is used to find the value at b_{t+1}.

While finding the maximum between $Q(b_t, I_t, a = L)$ and $Q(b_t, I_t, a = H)$ in order to get the value, $V^{a^*}(b_t, I_t)$, we use a softmax rule (with $\beta = 50$) instead of the max function. The softmax function is given by:

$$softmax(\, x_i\,) = \frac{e^{\beta x_i}}{\sum_j e^{\beta x_j}} \tag{2.13}$$

This finds a smooth maximum between the input numbers, x_i. When $\beta = \infty$, the softmax is equivalent to the max (technically, it is equivalent to the $\arg max$). Doing this prevents small numerical errors from being picked up as actual differences between the values.

2.3 Forward runs

Given the value/policy functions for the entire belief space at all internal states at each time point, the agent can choose internal states (L or H) using the optimal policy as evidence is collected (hence, belief is updated) through the trial. Starting at a prior belief, $b_{t=0}$, and initial internal state, IS_0, an action a_1 is chosen based on the value function. This sets the η_1 parameter for the emission functions, and we can accordingly sample an observation y_1 from state x_1. x_1 is reached from x_0 according to the transition probabilities. The sampled observation leads to a belief update and the cycle continues, leading to a sampled 'forward run' of the policy.

Algorithm 1: Forward run

Result: $b_{1:N}$, $IS_{1:N}$
Require transition functions, emission probabilities;
set initial b_0, IS_0;
for $t \leq N$ **do**
    choose optimal action a_t using $V^{a^*}(b_{t-1}, IS_{t-1})$;
    update internal state IS_t based on a_t;
    observe y_t based on IS_t and x_t;
    update b_t based on y_t;
end

2.4 Measures of performance

In the results, we report average performance of the model using the signal detection measures of sensitivity index (d') and response criterion (c). d' is a measure of how detectable a signal is, that is, how separated the underlying signal distribution is from the non-signal distribution. This can be calculated as: $z(\text{ Hit Rate }) - z(\text{ False alarm Rate })$. z is the inverse of the standard cumulative normal distribution.

c is a measure of how likely it is to respond in favor of the signal. It is calculated as: $-0.5 \cdot [z(\text{ Hit Rate }) + z(\text{ False alarm Rate })]$. Lower the c, greater is the tendency to respond in favor of the signal.

These two quantities are independent measures of performance in the task.

Chapter 3

Results

3.1 Task description

We begin by providing a brief summary of the task we modelled. Details of the task structure and the Hidden Markov Model (HMM) of the task may be found in Chapter 2.

The signal detection tasks we consider consist of trials of fixed length in which a signal may be present with probability (p_{signal}). Signal trials have a signal embedded in them while non-signal trials have no underlying signal throughout the trial. The task for an agent is to decide in favor of either of the hypotheses:

$$\begin{aligned} &H_0 : \text{This is a non-signal trial} \\ &H_1 : \text{This is a signal trial} \end{aligned} \tag{3.1}$$

We don't fix the signal length for reasons described in Chapter 2; instead, we allow a signal to end with constant probability q once it comes on. The signal has a uniform probability of coming on anywhere in the trial (which leads to an increasing hazard function). Figure 3.1 shows a schematic of a trial in the task.

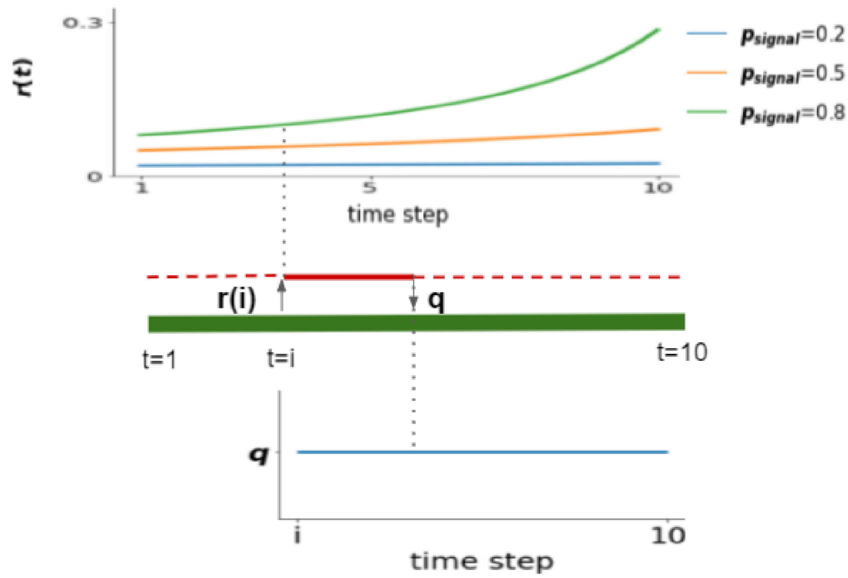


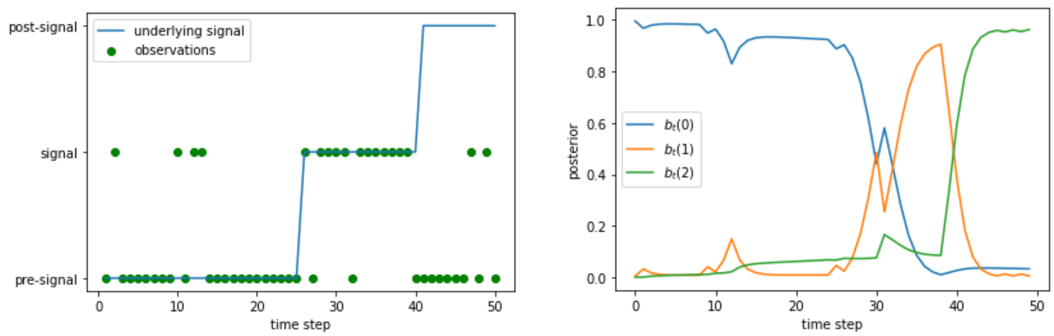
Figure 3.1: A schematic of a trial in the task. The trial is of fixed length $N = 10$. The probability per time step of the signal coming on is given by the hazard rate, $r(t)$ which is shaped differently for different p_{signal} as shown in the figure. The probability per time step of the signal turning off is given by the constant q. Once the signal turns off, it cannot come on again

3.2 Bayes optimal agent

Figure 3.2 shows the plots for the underlying signal, observations and inferred posterior probabilities within a single example trial. In this particular example, the signal comes on before $t = 25$ and turns off around $t = 40$. This is accompanied by a rise in $b_t(1)$ which drops after $t = 40$ after which $b_t(2)$ rises.

To further understand the behavior of the model, consider the average d' as a function of different parameters, namely trial length (N) and q.

Figure 3.3 shows that as q decreases, d' decreases. This is because smaller q leads to signals staying on for longer, allowing better detection on signal trials. On non-signal trials, when small signal lengths are expected at high q, a few noisy observations can easily be confused with a short signal – leading to more false alarms.



(a) Hidden states and observations

(b) Posterior probabilities

Figure 3.2: An example trial of the HMM with no attention costs. Parameters: $p_{signal} = 0.5$, $q = 0.1$, $\eta = 0.7$, $N = 50$

Trial length interacts with q. For smaller q, d' increase as trial length rises – a larger trial provides more chances for longer signals allowing better detection of the signal. However, for large values of q, d' decrease as trial length increases. This might be because signals tend to be short even as the trial is longer – where it is easier to miss a signal and confuse noise with signals.

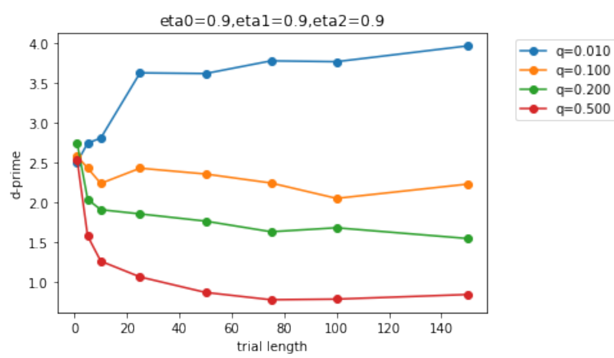


Figure 3.3: Performance (d') of the HMM with no attention costs as a function of trial length (N) and q. Parameters: $p_{signal} = 0.5, \eta = 0.9$

3.3 Resource limited agent

Given that the task structure is known, equation **??** gives the optimal inference strategy in the task and hence guarantees the best performance possible. However, this optimality is unbounded: it doesn't take into account realistic limits on information processing resources of an agent.

Here, we consider a resource-limited agent that experiences costs when accumulating evidence. These costs depend on the quality of evidence an agent is collecting. In our task, the greater the value of η, the greater is the quality of evidence – and hence the greater the cost. One can imagine that attention improves the quality of the evidence, but at a cost.

As described in Chapter 2, for simplicity, we allow our agent to collect evidence at two levels, η_H and η_L in two internal (or attentional) states (IS_t): high ($IS_t = H$) and low ($IS_t = L$). We assume the agent has control on the internal state at every time point. $\eta_H > \eta_L$ and so, the cost of collecting evidence at $IS_t = H$ is higher than the cost of choosing $IS_t = L$. Further, choosing an IS_t also depends on IS_{t-1}: the cost of switching between different internal states might be higher than maintaining a particular state, for example.

Given these assumptions, we can find the optimal policy via the (Bellman) equation 2.10.

3.3.1 Signal stays on: A first model

The HMM consists of three states: the pre-signal(0), signal(1) and post-signal(2) states. Since $b_t(1) + b_t(2) = 1 - b_t(0)$, the corresponding belief space is two dimensional.

As a first case, we set $q = 0$. This means that there is zero probability that state 1 will go to state 2 (i.e., there is no post-signal state – if a signal comes on, it stays on until the end of the trial). Effectively, this reduces to an HMM with 2 states with the following transition matrix:

$P(X_t \mid X_{t-1})$		X_t 0	1
	0	1-r(t)	r(t)
X_{t-1}	1	0	1

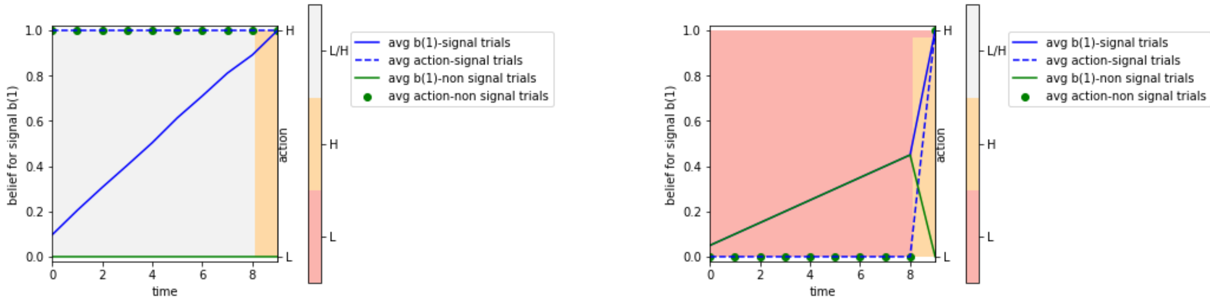
where $r(t) = \frac{1}{\frac{N}{p_{signal}} - t}$. This leads to a one-dimensional belief space where $b_t(1) = 1 - b_t(0)$. We start by characterising this model before moving on to the more complex case where the belief space is two dimensional.

Behavior of the model

Consider the edge case when $\eta_H = 1.0$ and $\eta_L = 0.5$ – the H state provides completely reliable information while the low state gives completely unreliable information. In addition, all the costs are set to 0. In this case, $Q(b_t, I_t, a = H) > Q(b_t, I_t, a = L)$ for all belief values at $t = N$ and $Q(b_t, I_t, a = H) = Q(b_t, I_t, a = L)$ for all time points before as seen from the policy function in Figure 3.4a. Increasing the costs of choosing H even slightly will cause $Q(b_t, I_t, a = H)$ to dip below $Q(b_t, I_t, a = L)$ for $t < N$ as seen in the policy function from Figure 3.4b. In other words, getting completely certain information at the final time point is enough to detect a signal and the belief on the previous steps do not matter. This is because $q = 0$ and so, a signal is guaranteed to stay on once it turns on.

More realistically, what if the H state doesn't provide exact information (i.e. $\eta_H < 1$)? In this case, it is not necessarily sufficient to choose H only on the final time step since there are belief states earlier in time where choosing H has higher value as seen in Figure 3.5a. Therefore, as the quality of information reduces, it is optimal to pay for attention earlier in time. In addition, if choosing H is made costlier or if the L state is more informative, then the decision to choose H can be pushed to a later time point. Figure 3.5b and c shows these results.

In summary, for the case when $q = 0$, it is optimal to choose H as late in the trial as possible, simply because the signal doesn't turn off once it comes on.



(a) $costs = 0$ - L and H have the same value before $N = 10$

(b) $costs_{iH} = 0.03$ - Optimal to choose H only on final time point

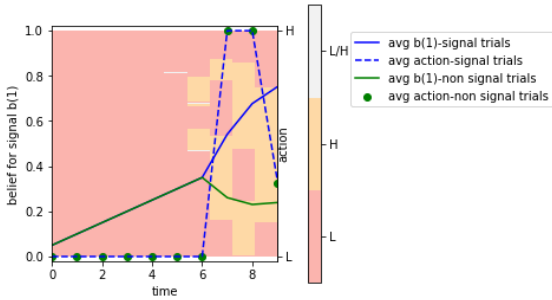
Figure 3.4: Optimal policy overlaid by average forward runs (average posterior and average action) at each time step when $\eta_H = 1.0, \eta_L = 0.5, N = 10$. The optimal policy is shown for all belief states at each time step. The label 'L' refers to the region where it is optimal to choose the low state, 'H' refers to where it is optimal to choose the high state and 'L/H' refers to where both actions have equal value

3.3.2 The three state model

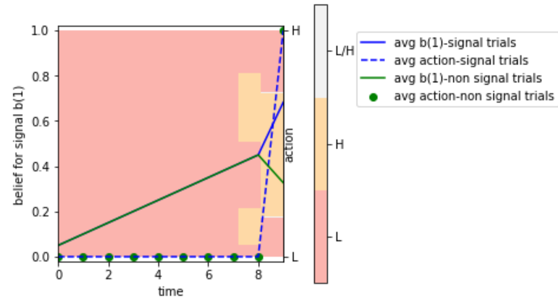
With intuition from examining inference from the no-cost case and also the optimal policy and inference for the simple case with $q = 0$, we now turn to investigating the behavior of the full model.

We begin by fixing the task parameters to the following: $N = 10$, $q = 0.2$, $p_{signal} = 0.5$, $\eta_H = 0.9$, $\eta_L = 0.6$ and costs $\begin{pmatrix} c_{00} = 0.0 & c_{01} = 0.04 \\ c_{10} = 0.0 & c_{11} = 0.02 \end{pmatrix}$. It was natural to allow costs of switching to the H state (c_{01}) to be greater than the costs of maintaining the H state (c_{11}). We allowed there to be equal probability for a signal and non-signal trial ($p_{signal} = 0.5$). Finally we set $q = 0.2$ because the expected signal length ($\frac{1}{q} = 5$) is not too high or low compared to our trial length $N = 10$. These form our default set of parameters for the rest of the report. When the value of any of the parameters is not explicitly mentioned, they can be taken to be the above.

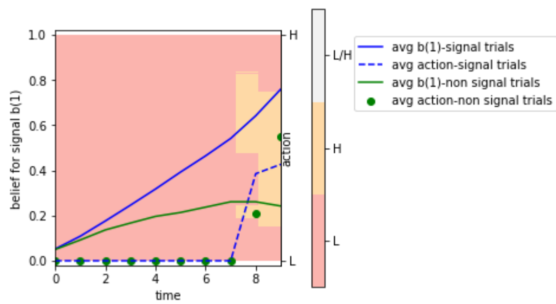
We use the Bellman equation 2.10 to solve for the time-dependent optimal value and policy functions for each belief and internal state.



(a) $costs_{iH} = 0.03, \eta_H = 0.8, \eta_L = 0.5$ - decreasing η_H makes it optimal to choose H earlier in time



(b) $costs_{iH} = 0.05$ - increasing costs of choosing H makes it optimal to choose H later in time



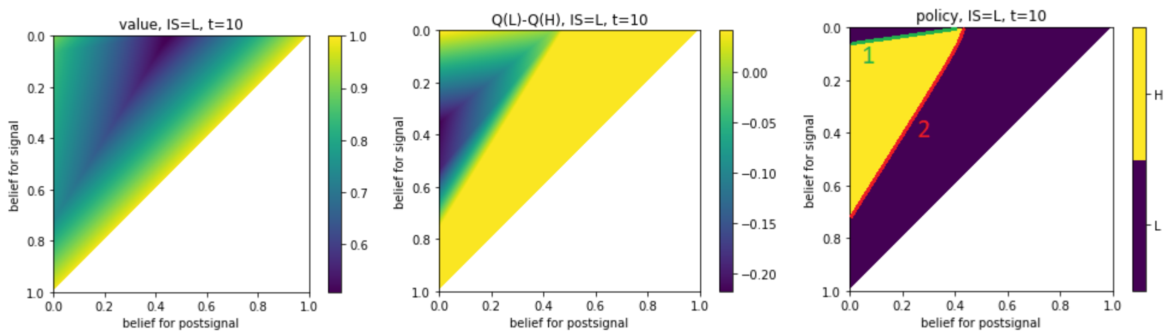
(c) $costs_{iH} = 0.05, \eta_L = 0.7$ - increasing η_L makes it optimal to choose H later in time

Figure 3.5: Optimal policy overlaid by average forward runs (average posterior and average action) at each time step when $N = 10$ at different costs and $\eta_{H,L}$. The optimal policy is shown for all belief states at each time step. The label 'L' refers to the region where it is optimal to choose the low state, 'H' refers to where it is optimal to choose the high state and 'L/H' refers to where both actions have equal value

To understand the policy, we begin by examining the value and policy functions for the 3D state space at the final time step. Figures 3.6, 3.7 show these functions for all belief states and for both, $IS = L$ and $IS = H$ as shown in Figures 3.6, 3.7. To recall,

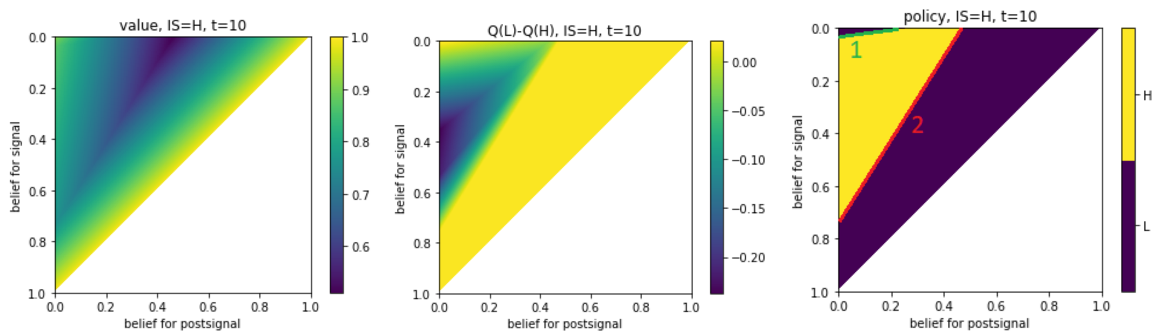
$$\begin{aligned} V^{a^*}(b_t, IS_t) &= \max_a \{Q(b_t, I_t, a = L), Q(b_t, I_t, a = H)\}, \\ policy(b, IS) &= argmax_a \{Q(b_t, I_t, a = L), Q(b_t, I_t, a = H)\} = a^*, \text{ and,} \\ \Delta Q &= Q(b_t, I_t, a = L) - Q(b_t, I_t, a = H) \end{aligned} \tag{3.2}$$

Based on the sign of ΔQ (that is, if $Q(L) > Q(H)$ or the opposite), the belief space (at each internal state) is partitioned into different regions: regions where it is optimal to



(a) Value Function (b) ΔQ Function (c) Policy Function

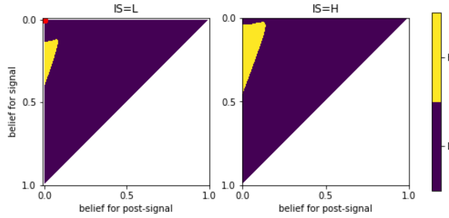
Figure 3.6: Value, ΔQ and policy functions (with $IS = L$) at the final time step, $t = 10$ using the standard parameter range



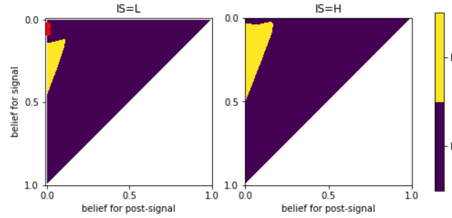
(a) Value Function (b) ΔQ Function (c) Policy Function

Figure 3.7: Value, ΔQ and policy functions (with $IS = H$) at the final time step, $t = 10$ using the standard parameter range

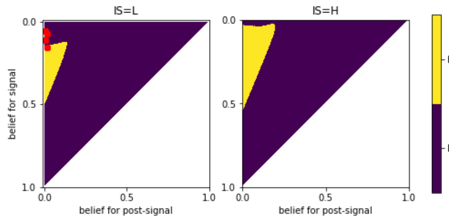
choose H, and regions where it is optimal to choose L. This is shown by the policy function as shown in Figures 3.6c, 3.7c. There exist two boundaries that separate these regions. The first boundary on the top-left corner (shown in green) represents the belief states above which $b(X_t = 0)$ is large enough (and where $b(X_t = 1), b(X_t = 2)$ are low enough), that it is not worth paying the costs for better information (in H state) on the next time step. The second boundary (shown in red) are the belief states beyond which $b(X_t = 1), b(X_t = 2)$ are high enough such that a costlier state which provides better information is not worth choosing (in other words, there is enough evidence in favor of the signal trial). It is only in an intermediate range of the belief space where it is advantageous to choose the H state.



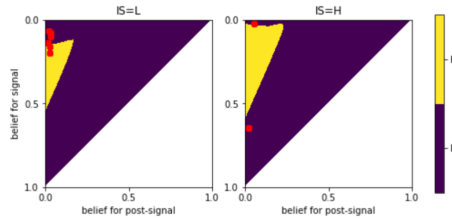
(a) $t = 1$



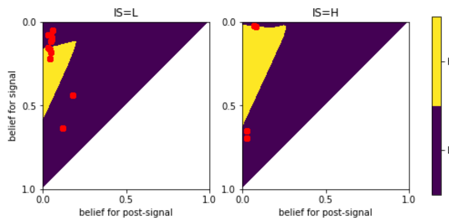
(b) $t = 2$



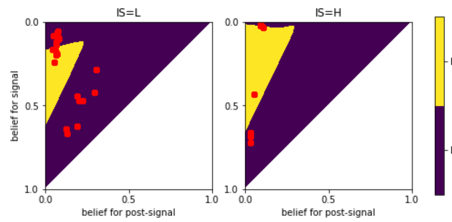
(c) $t = 3$



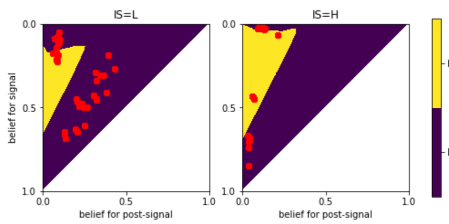
(d) $t = 4$



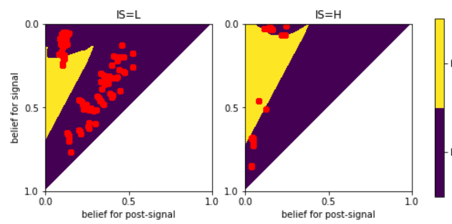
(e) $t = 5$



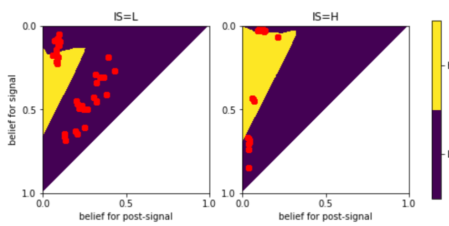
(f) $t = 6$



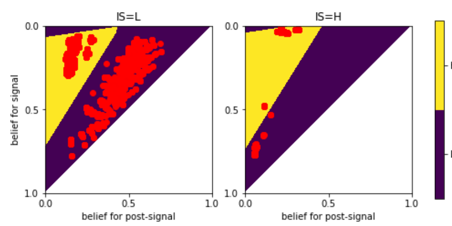
(g) $t = 7$



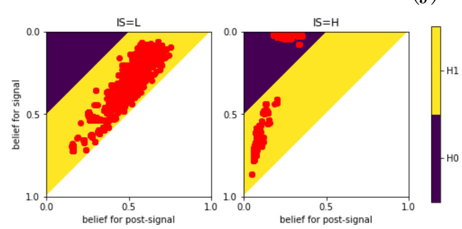
(h) $t = 8$



(i) $t = 9$



(j) $t = 10$



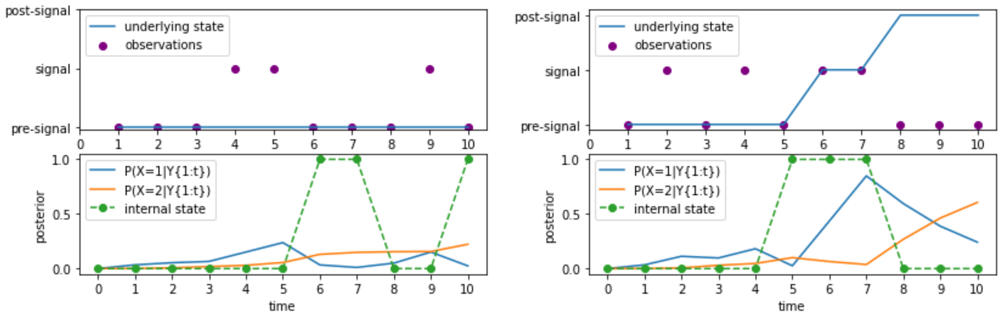
(k) Final decision (between H_0 and H_1)

Figure 3.8: Policy functions overlaid by scatter plots showing the states occupied by actual forward runs of the model. This is shown for all time steps for the entire state space (the 2D belief space and the two internal states, $IS = L$ and $IS = H$). The forward runs are obtained by staring at the initial belief: $b(0) = 1$,internal state: $IS = L$ and using the standard parameters

Figure 3.8 shows the policy plots for all $N = 10$ time steps, overlaid by scatter plots showing the states occupied by actual forward runs of the model obtained using this optimal policy (procedure for this is discussed in Chapter 2). They are obtained by staring at the initial belief: $b(0) = 1$,internal state: $IS = L$ and using the standard parameters. Each scatter point exists in a 3D space (2D belief state and a variable for the internal state). The region where it falls in the space determines what action is chosen (according to the policy function) and hence the internal state for the next time point. The next belief state is found by a Bayesian belief update. Notably, the policy is time-dependent: the boundaries move further apart in time leading to the 'choose H' region expanding – in general, it is more useful to choose $IS_t = H$ later in the trial because q is still low and so there is a decent probability for signals to stay on till the end of the trial.

To get a better feel for the trends in our trajectories, we first consider single example trials as shown in Figure 3.9. Figure 3.9a shows an example non-signal trial. The trial begins at an initial belief of $b_0(0) = 1$ and internal state, $IS_0 = L$. For the first five time points, the belief states lie below the first boundary (i.e., there is not enough evidence in favor of the signal/ post-signal states to trigger shifting to the H state). Beyond this, the belief states fall in the regions where it is optimal to choose the H and hence obtain better evidence about the underlying signal.

Figure 3.9b shows an example of a signal trial. Similar to the previous case, $IS = H$ is chosen only after the fourth time step. However, in this case, a signal comes on at $t = 6$. This leads to increased evidence in favour of a signal and so $b(X_t = 1) + b(X_t = 2)$ rise leading the belief states to cross the second boundary at $t = 8$. Due to the nature of the transition function between the 3 underlying signal, states, there is always a transfer of probability density in favor of $X_t = 1, 2$ through time – this means that $(b(X_t = 1) + b(X_t = 2))$ increases with time irrespective of the observations. Hence, crossing the second boundary means that the 'choose H' region will not be reached again. In other words, once the belief in favor of the signal trial crosses a boundary, there is no need to pay for better information since the agent can be sure that a signal is/was present – a ‘decision’ is already reached before the end of the trial!



(a) Example non-signal trial (b) Example signal trial

Figure 3.9: Example trials generated using the standard parameters. Each plot shows the following at each time step through the trial: underlying signal state (in blue) and observations (as purple scatter points) in the top graph; posteriors (for $X_t = 1$ in blue and $X_t = 2$ in orange) and the internal state (in green) in the bottom graph

Now that we have looked at single example trials, we now examine average performance. In a fixed duration signal detection task, the main measures of behavior are the hit (correct signal detection) and false alarm (false signal detection) rates. Alternatively, the correct rejection (correct response to non-signal trials) and miss (failure to detect signal) rates may also be reported. The hit and false alarm rates can be used to find the sensitivity index (d') which measures how well an agent can distinguish a signal from noise. In addition, they can be used to find the response bias (c) which represents the criterion for responding in favor of the signal. As seen in table 3.1, the d' is well above 0 meaning that the signal is really being discriminated from the noise. c is slightly over 0 implying there is a bias for reporting in favor of the non-signal (irrespective of the underlying signal state).

hit= 0.725	FA= 0.201	$d' = 1.435$
miss= 0.275	CR= 0.799	$c = 0.119$

Table 3.1: Average performance obtained using the standard set of parameters

We also examine the ‘attentional dynamics’, or how likely it is to choose the ‘H’ internal state through time within a trial. Some natural questions include the following:

1. What fraction of a trial is spent in $IS = H$? For an experimenter, this is akin to asking how attentive a subject is, on an average.

2. Where in the trial is $IS = H$ chosen the most? In an experimental terms, does the subject pay attention uniformly through the trial?

3. How do these measures differ between signal and non-signal trials?

4. What is the relationship between accuracy and the occupancy of $IS = H$? That is, do subjects perform better when they pay more attention?

We find the frequency of choosing $IS = H$ at each time step averaged across multiple trials. Figure 3.10 shows this expected occupancy of the H state at each time step. As

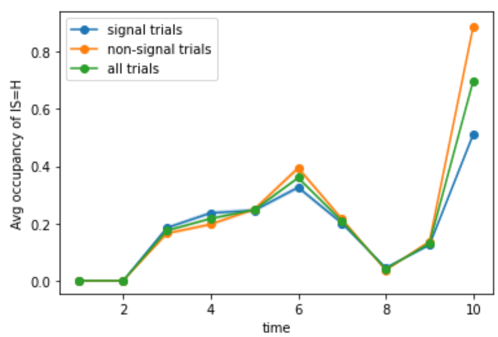


Figure 3.10: Average occupancy of $IS = H$ at each time step in the trial, found using the standard parameters. This is shown for signal, non-signal and all trials

expected from the policy plots in Figure 3.8, occupancy in the earlier time steps is low and rises in time. This is followed by a drop and quick rise at the final time step. This means that it is optimal to pay the most for better quality information at the final time point. At $q = 0.2$, the expected signal length is $\frac{1}{q} = 5$ and so there is a good chance of detecting a signal at the final time point. However, as there is also a significant probability of a signal switching off before the end, it is advantageous to pay attention at intermediate time points as well. A possible reason for a dip at $t = 7$ to 9 might be as follows: if a signal has not been detected when $IS = H$ was chosen before, then it is a non-signal trial or the signal has already ended or is yet to come on. For the first and last cases, increasing attention in the final time step should be enough to infer correctly in most of the trials. If the signal has already ended, then there is no utility in attending to the signal. If a signal has already been detected, there is again not much of a gain in paying more attention.

Figure 3.10 compares the occupancy through the trial for signal and non-signal trials. They are mostly similar except for some small differences. The occupancy is slightly higher at earlier time points and lower at later ones (especially the end point) for the signal trials: this may again be explained by the observation that if a signal is detected earlier (which can happen for early signals), then there is no need to pay for more information later.

To calculate the average occupancy of the H state, we find the percentage of a trial that is spent in $IS = H$, averaged over multiple (5000) trials. The average rate of occupancy for the entire trial is: 0.206 ± 0.001. This means that, on an average, 0.206 fraction of a trial is spent in the high attentional state. For signal trials, the proportion is 0.186 ± 0.002, and for non-signal trials it is 0.226 ± 0.002. There is also no real difference in occupancy of $IS = H$ between accurate trials (where the average occupancy is 0.215 ± 0.095) and inaccurate trials (with average 0.205 ± 0.092). Intuitively, one would expect that spending more time points in the higher attention state would lead to higher certainty about the underlying signal. However, there are many confounds to this: for example, once an early signal is detected there is not much added utility of going to $IS = H$. Hence if a more certain belief state is reached quickly, attention can be turned off without much effect on the accuracy.

Having looked at the performance and 'attention' trends under the default parameters, the next natural question is how the behavior is affected by varying these parameters.

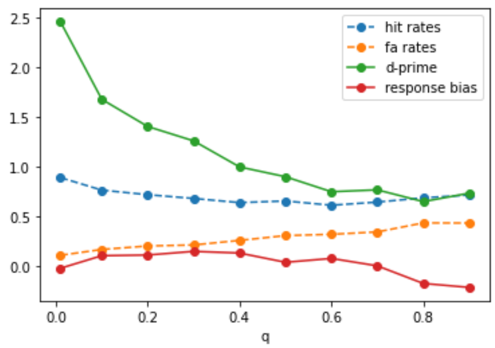
Effects of Signal Length

The first parameter we look at is q – the probability per time step that a signal turns off once on. If q is higher, then signals will tend to be shorter. Since q is a constant number, the resultant signal lengths follow a geometric distribution with expected value $1/q$ – this therefore gives the expected signal length. In effect, the tail of the geometric distribution is always truncated because the trial length is finite. It is hence good to keep in mind that the expected signal length of $1/q$ is only an approximate estimate.

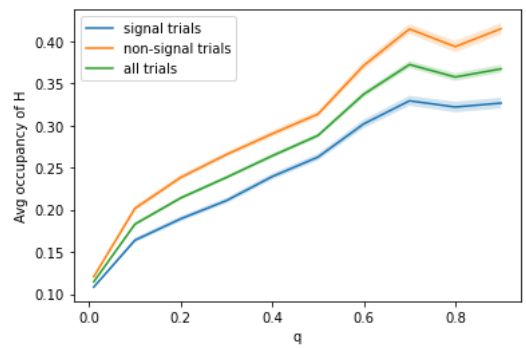
What should be the effect of increasing q (that is, reducing the average signal length) on behavior? Firstly, we would expect that shorter signals are harder to detect – because a shorter signal implies a shorter interval where there are more signal than non-signal observations and so it is harder to detect the signal from the background noise. Intuitively, as a signal gets harder to detect, we predict that attention should be engaged more to improve chances of detection.

In order to test our hypotheses, we simulate our model at different values of q, keeping all other parameters at their default values. Figure 3.11 shows the overall performance measures including d', c, hit and false alarm rates as q increases. The sensitivity, d' falls with increase in q. This is due to a decrease in hit rates and also a rise in false alarms. This trend is just as how we predicted – the shorter the signal, harder it will be to detect. The response bias is constant and above 0 for low q and then it decreases (to below 0) for large q. This implies that the propensity to respond in favor of the signal rises with q.

What about the occupancy of $IS = H$? As predicted, as q rises, the average occupancy of $IS = H$ also increases. This is because, when a signal is harder to detect, $IS = H$ should be chosen more often to improve detection. However, for very large values of q, this effect saturates and there is no rise in occupancy beyond this. This might be because when signals get too short, choosing $IS = H$ more will not significantly improve performance.



(a) Performance curves for different q



(b) Occupancy of $IS = H$ for different q

Figure 3.11: Average behavior for different values of q with all other parameters fixed at default values. The plots include hit, false alarm rates, d' and c curves, and average occupancy rates for signal, non signal and all trials

The average trends are a good window into the effects of changing signal length on the model dynamics. To dig a bit deeper, we analyse the temporal pattern of occupancy of $IS = H$ within a trial and also the optimal policy functions.

Figure 3.12 shows the average occupancy at every time point within a trial at $q = 0.6$. Quite visibly, the occupancy of $IS = H$ is high even at earlier time points in contrast to $q = 0.2$ where it is the highest at the final time point by a huge margin (shown in Figure 3.10). This is because signals tend to be very short at $q = 0.6$ and so it is required to pay attention almost throughout the trial to effectively detect signals.

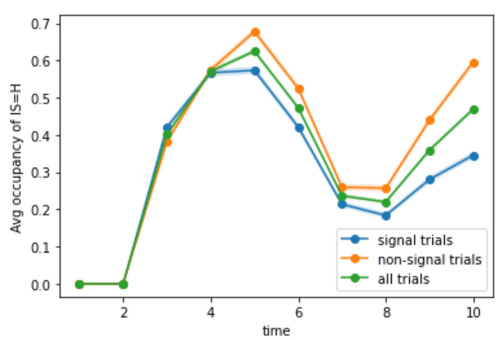
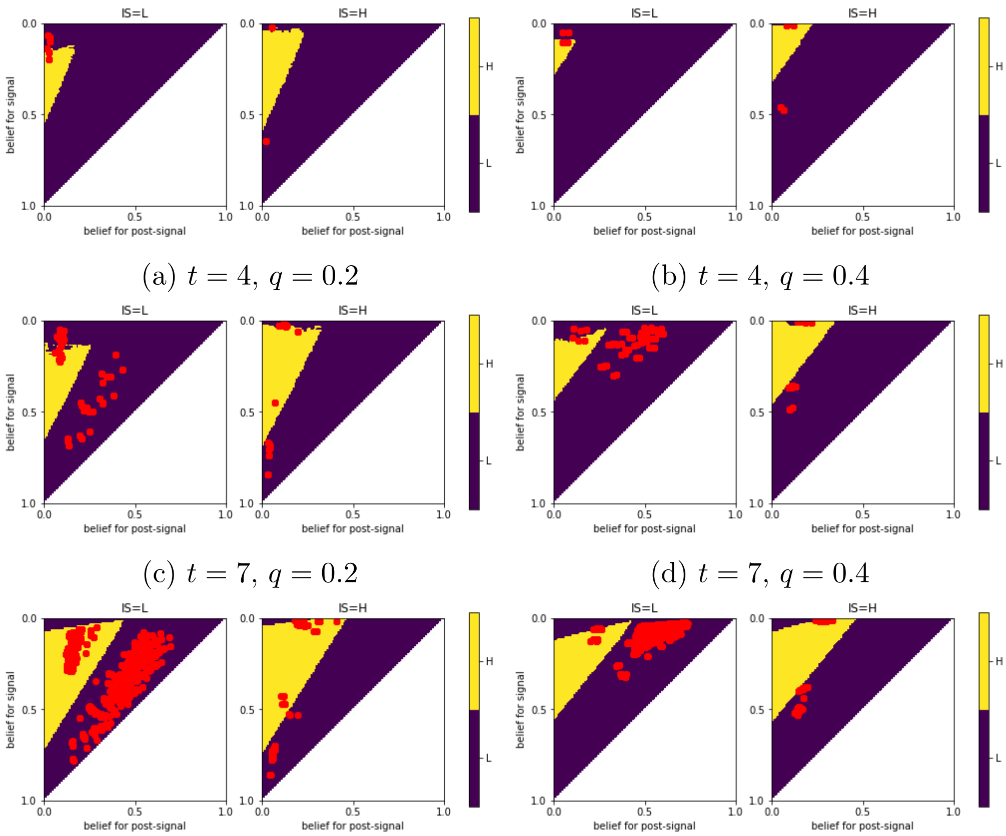


Figure 3.12: Average occupancy of $IS = H$ at each time step in the trial, found using $q = 0.6$. This is shown for signal, non-signal and all trials combined

It is interesting to interrogate the differences between the optimal policy at $q = 0.6$ and

$q = 0.2$. Figure 3.13 shows the policy plots for the entire state space at three different time steps in the trial. Evidently, the position of the two critical boundaries differ between the two cases (as compared in Figure 3.13). When $q = 0.6$, the first boundary is moved up towards smaller values of $b(1), b(2)$ – this means that attention is engaged even with low evidence in favor of the signal. The second boundary is also moved up towards smaller $b(1), b(2)$ – this means that the beliefs in favor of the signal need not be very high to decide in favor of it.



(a) $t = 4$, $q = 0.2$ (b) $t = 4$, $q = 0.4$

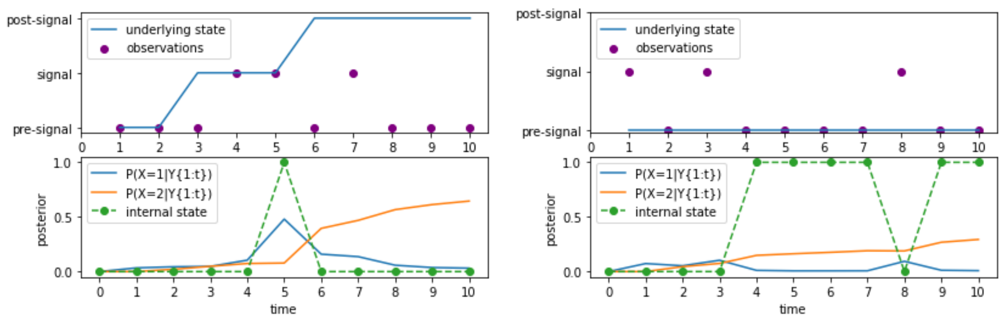
(c) $t = 7$, $q = 0.2$ (d) $t = 7$, $q = 0.4$

(e) $t = 10$, $q = 0.2$ (f) $t = 10$, $q = 0.4$

Figure 3.13: A comparison of policy function plots for selected time points for the whole state space at $q = 0.2$ and $q = 0.6$. Note the differences in positions of the two boundaries (marked in green and red) between the two cases. These are overlaid by scatter plots obtained from forward runs

To illustrate these effects better, we include single example signal and non-signal trials as shown in Figure 3.14. Figure 3.14a shows an example signal trial. The $IS = H$ is chosen even with a small rise in $b(1), b(2)$ at $t = 4$. Another signal observation is enough for the belief states to cross the second boundary – there is sufficient evidence in favor of the signal trial. This strategy makes sense because when q is larger, signals tend to be very short so a short sequence of signal observations should be enough evidence for the signal.

Figure 3.14b shows an example non-signal trial. As before, $IS = H$ is chosen even with a small rise in $b(1), b(2)$. In absence of signal observations, $b(1), b(2)$ stay low and so the agent stays in the H state almost through out the entire trial. This means that the agent must stay in a monitoring state throughout in order to not miss a signal.



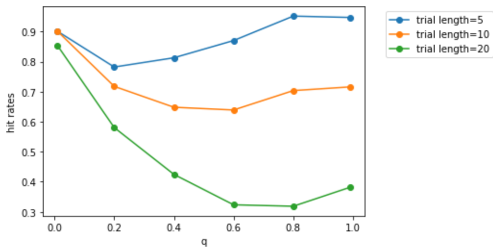
(a) Example signal trial (b) Example non-signal trial

Figure 3.14: Example signal and non-signal trials using $q = 0.6$. The plots show the underlying signal, observations, posteriors and internal states at every time step through the trial

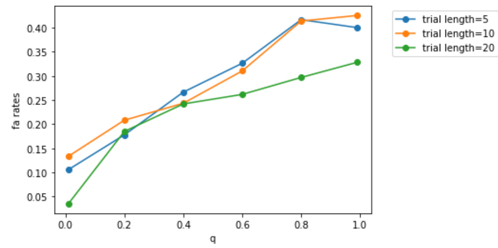
Does only the absolute value of q (which determines the average signal length) effect the detectability of a signal? One could imagine that the signal length *relative* to the trial length determines how detectable a signal is. For example, at the same expected signal length, the longer the trial, the more opportunity to confuse signal with noise, making it harder to detect. Hence, for the same q, we expect performance to be lower for longer trial length. Figure 3.15 shows this is indeed the case.

What is the effect of trial length on occupancy of the H state? We predict that for a q, larger the trial length, the more the H state needs to be engaged in order to detect the signal.

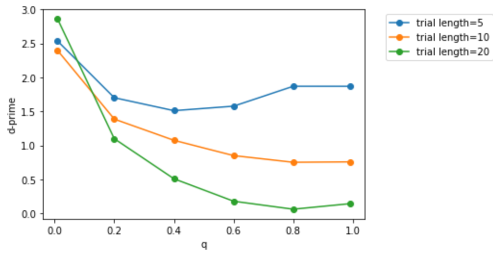
We observe the opposite effect with H being chosen the most at the lowest trial lengths.



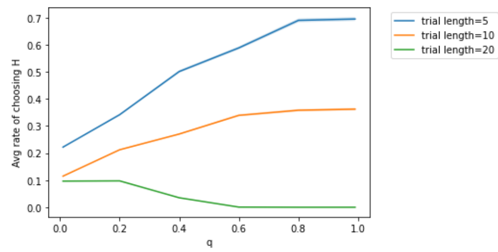
(a) Hit rates for different values of q at different trial lengths



(b) False alarms for different values of q at different trial lengths



(c) d' for different values of q at different trial lengths



(d) Average occupancy for different values of q at different trial lengths

Figure 3.15: Average behavior for different values of q and trial length with all other parameters fixed at default values. The plots include hit, false alarm rates, d' curves, and average occupancy rates for signal, non signal and all trials

This is due to a decrease in occupancy on an average for each point in time, as N increases. This happens because as N becomes larger, there is a smaller probability of a signal coming on at each time step.

Effects of signal probability

Another crucial parameter in the model is p_{signal}, the probability of a signal trial. There are multiple ways which this could be manipulated in general vigilance tasks – increasing the (non-target) event rate could reduce the perceived signal probability; the target rate itself can be directly reduced.

If the signal rate is low (< 0.5), then we would expect the response criterion to be biased in favor of non-signal trials and the opposite if signal rate is high. Using signal detection theory, we predict the following trend in the criterion with p_{signal}: Say the signal and non-signal trials are drawn from normal distributions with $\mu_{NS} = 0, \mu_S = 1$ and $\sigma = 1$ for both. The Bayes criterion on observation x is given by:

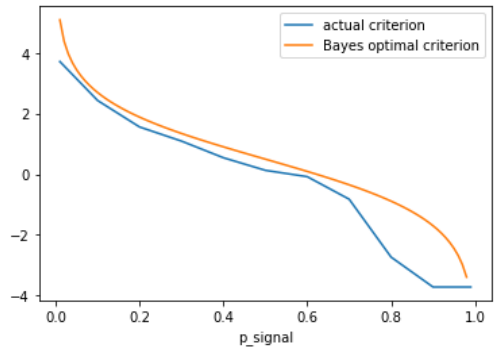
$$\begin{aligned} \frac{p(x|\mu_S, \sigma)}{p(x|\mu_{NS}, \sigma)} &= \frac{(1 - p_{signal}) * (R_{00} - R_{01})}{p_{signal} * (R_{10} - R_{11})} \\ e^{\frac{-(x-1)^2 + x^2}{2}} &= \frac{1 - p_{signal}}{p_{signal}} \\ x &= ln(\frac{1 - p_{signal}}{p_{signal}}) + 0.5 \end{aligned} \tag{3.3}$$

This Bayes criterion depends on p_{signal} by a function as shown in Figure 3.16a.

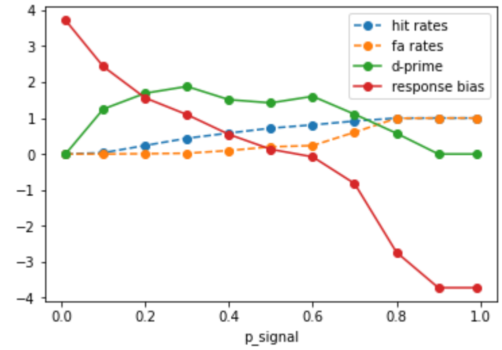
How should the signal rate effect performance? Since sensitivity is independent of the response bias, we predict that there should be no effect on overall performance (d'). We predict that the occupancy of the H state is the highest when there is equal probability for a signal and non-signal trial to occur. If the prior is biased in favor of one trial type, then there is less utility in improving evidence at a cost.

We ran simulations to test our predictions. For the performance curves, as p_{signal} increases, the response bias decreases, thereby increasingly favouring the reporting of a signal. Therefore, at the extreme probabilities, the hit and false alarm rates either fall towards 0 or 1. Fig 3.16a compares the optimal Bayes criterion (for different signal priors) and the actual criterion obtained from the model. The model criterion has a similar shape to the optimal criterion, in line with our predictions.

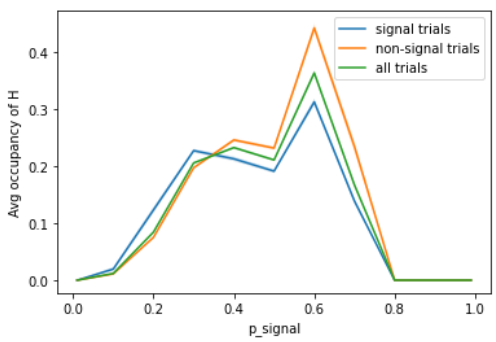
Surprisingly, d' doesn't stay constant throughout but instead drops at the extreme p_{signal} – this means that when a signal trial is very likely or very rare, the detectability of the signal is lower. The reason for this lies in the occupancy (of $IS = H$) curve for the different signal probabilities, shown in Fig 3.16c. At the extreme signal probabilities, the occupancy falls to 0. This implies that all of the inference at these p_{signal} is done at $\eta_L = 0.6$ and so the detectability of signal is low at extreme p_{signal}. The occupancy is an inverted-U as we predicted. However, the peak of the curve doesn't lie at $p_{signal} = 0.5$ but at $p_{signal} = 0.6$.



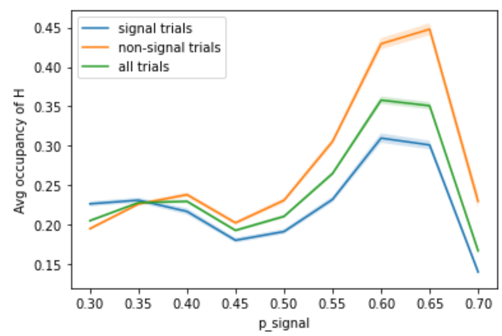
(a) Comparing criterion form simulations to the Bayes optimum criterion



(b) Performance curves for different p_{signal}



(c) Occupancy curves for different p_{signal}

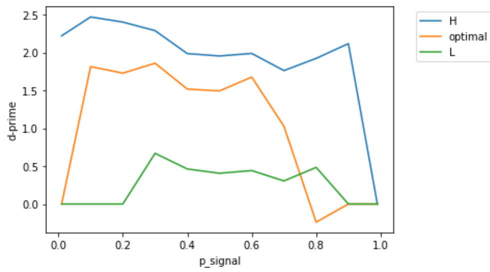


(d) Occupancy curve zoomed in at $p_{signal} = 0.3$ to $p_{signal} = 0.7$. This demonstrates that the double peak in the occupancy is a real effect.

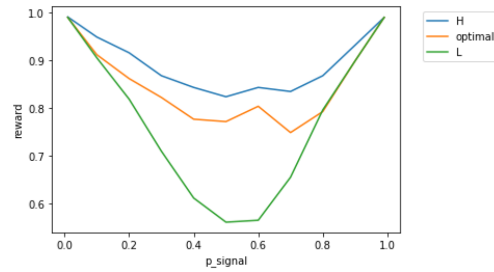
Figure 3.16: Average behavior for different values of p_{signal} with all other parameters fixed at default values. The plots include hit, false alarm rates, d' and c curves, and average occupancy rates for signal, non signal and all trials

Why is attention engaged much less at extreme probabilities? This is because the costs of choosing $IS = H$ enough number of times to improve performance over-weigh the rewards obtained from a marginal rise in performance. This is well demonstrated in Figure 3.17. At extreme signal probabilities, choosing H doesn't confer much advantage over L and consequently, it is optimal to simply not choose H and pay the costs. This is akin to saying that when the prior is already very high in favor of signal or non-signal trials, then collecting better quality evidence doesn't improve performance much.

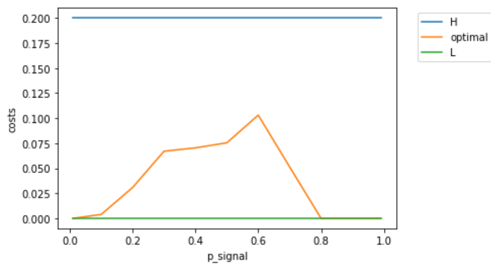
How do the temporal dynamics of occupancy look at different signal probabilities? At



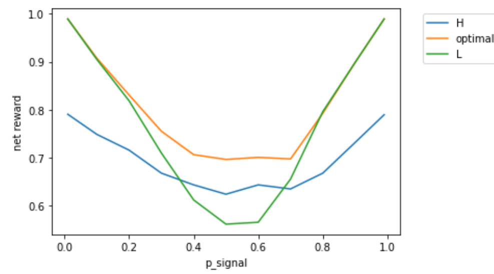
(a) d' for the three conditions



(b) Average total reward for the three conditions



(c) Average costs per trial for the three conditions



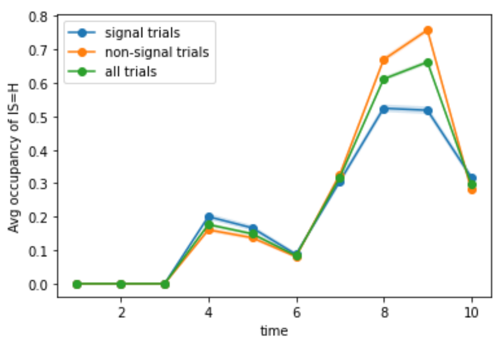
(d) Net Reward for the three conditions

Figure 3.17: Behavior of the model with different p_{signal} under the three conditions: choosing $IS = H$ throughout, choosing IS according to the optimum policy (optimal condition), choosing $IS = L$ throughout. The plots show d', average total reward, average costs and net reward for the three conditions. While d' and so total reward is the highest when $IS = H$ is always chosen, the corresponding costs are also high. Hence, the net reward is the highest for the optimal case

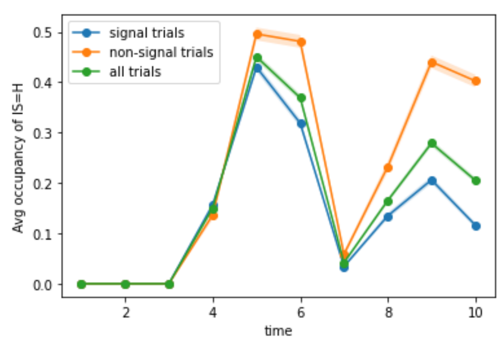
higher p_{signal}, $IS = H$ is chosen quite often even earlier in the trial instead of being concentrated at the final time point. At lower p_{signal}, the H state is occupied mostly towards the end of the trial. These effects are shown in Figure **??** for two example signal probabilities.

This can be explained by how the boundaries which determine the policy, shift with change in p_{signal}. As this probability increases, the first boundary is shifted to lower $b(1), b(2)$ as shown in Figure 3.23 hence activating the H state for smaller rises in belief in favor of the signal (this makes sense – if signals are more probable, there is a greater probability for $X = 0$ to switch to $X = 1$ at every time point). Hence H tends to be activated at early time points as much as the later ones. An opposite effect occurs at the lower signal probabilities. If this is the case, why does the occupancy decrease for higher probabilities? This might be because with larger signal probabilities, the trajectories quickly cross the second boundary, hence not keeping the H state active for long.

This effect is well demonstrated by an example run as shown in Figure 3.20. The underlying signal is the same for both the runs. For $p_{signal} = 0.4$, the H state is activated only later in time and doesn't turn off despite rise in belief in favor of $X = 1, 2$. On the other hand, for $p_{signal} = 0.7$, the H state is quickly activated and is also turned off by the evidence gathered from only a single signal observation in the high attention state.



(a) Average occupancy of $IS = H$ at $p_{signal} = 0.4$



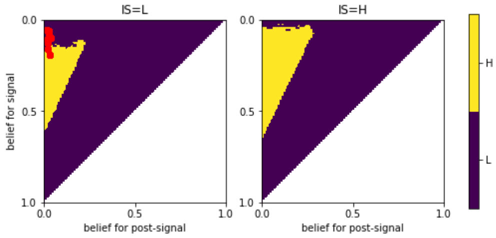
(b) Average occupancy of $IS = H$ at $p_{signal} = 0.7$

Figure 3.18: Average occupancy of $IS = H$ at each time step in the trial, found using $p_signal = 0.4, 0.7$. This is shown for signal, non-signal and all trials combined

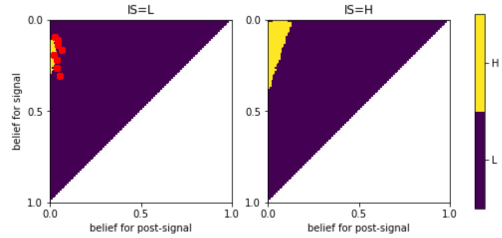
Interaction between q and signal probability

The previous sections showed the effect of varying q and p_{signal} separately. Here we look at how they interact by varying both. We predict that d' will decrease with q at all values of p_{signal} and will follow an inverted-U shape at each p_{signal}. We predict that the occupancy rate will increase with q at every signal probability while following an inverted-U shape at each p_{signal}.

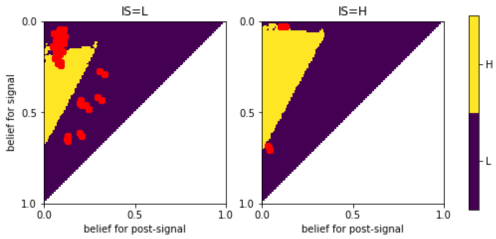
Figure 3.21 shows the trends observed from the simulations. d' shows the patterns that we predicted, except that the extent of decrease in d' with q also depends on p_{signal} – more extreme p_{signal} values cause a greater decline in d'.



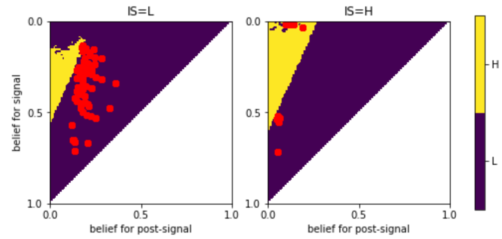
(a) $t = 4$, $p_{signal} = 0.4$



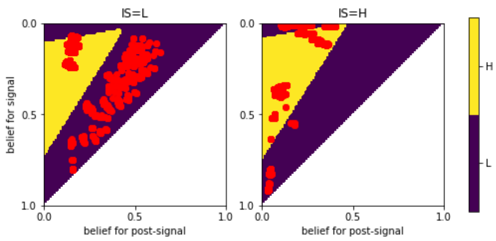
(b) $t = 4$, $p_{signal} = 0.7$



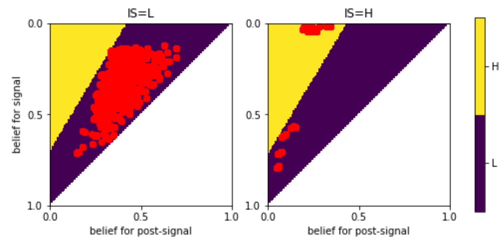
(c) $t = 7$, $p_{signal} = 0.4$



(d) $t = 7$, $p_{signal} = 0.7$



(e) $t = 10$, $p_{signal} = 0.4$

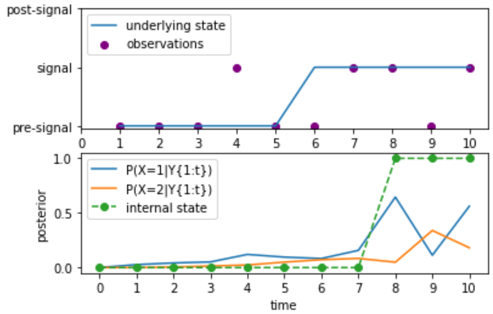


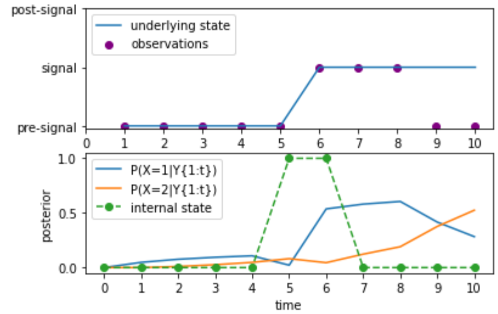
(f) $t = 10$, $p_{signal} = 0.7$

Figure 3.19: A comparison of policy function plots for selected time points for the whole state space at $p_signal = 0.4, 0.7$. Note the differences in positions of the two boundaries (marked in green and red) between the two cases. These are overlaid by scatter plots obtained from forward runs

Surprisingly, the effect on occupancy rate is opposite to what we predicted: occupancy shows a decreasing trend for all p_{signal} values except $p_{signal} = 0.5$! Why is this the case? As discussed previously, for extreme signal probabilities, it is not worth paying the costs of the high attentional state for a marginal improvement in performance (since with a biased prior, one can already be sure about the trial). As q increases, and the signal gets harder, this effect dominates, hence causing the occupancy to reduce.

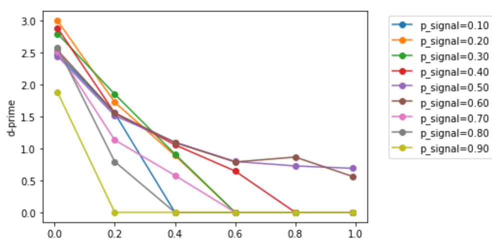
This means that short (therefore, hard-to-detect) signals that are very rare or very frequent are not detected well. This is perhaps because they engage the least amount of

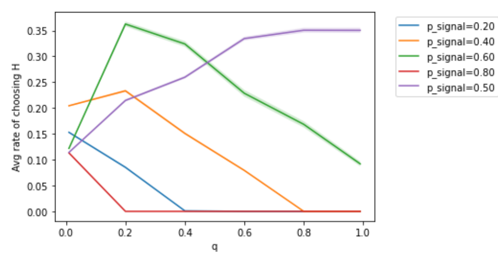




(a) Example signal trial, $p_{signal} = 0.4$ (b) Example signal trial, $p_{signal} = 0.7$

Figure 3.20: Example signal and non-signal trials using $p_{signal} = 0.4, 0.7$. The plots show the underlying signal, observations, posteriors and internal states at every time step through the trial





(a) d' curves for different q and p_{signal} (b) Occupancy curves for different q and p_{signal}

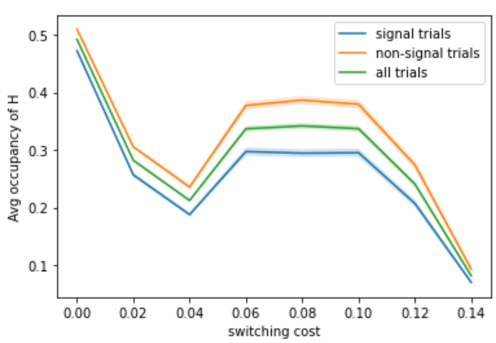
Figure 3.21: Average behavioral trends with varying q and p_signal. The plots of d' and occupancy curves show an interaction between q and p_{signal}

attention.

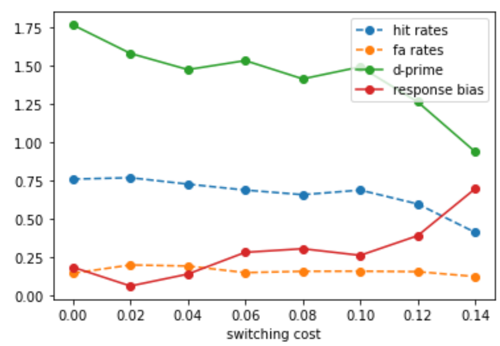
Effects of costs

Another important parameter is the cost of choosing $IS = H$. In our default parameter settings, this cost is asymmetric, such that choosing H from the L state has higher cost than choosing it from the H state. In other words, the switching cost to $IS = H$ is higher than the maintenance cost. We can change these costs in multiple combinations in order to investigate their effects. We choose to change the switching costs while keeping the maintenance costs fixed. What might be the expected effects? Firstly, the occupancy of $IS = H$ should decrease if the costs of choosing it are higher. Secondly, if H is chosen a fewer number of times due to increasing costs, then the detectability of the signal would drop leading to worse performance.

We simulate the model at different switching costs and obtain the average occupancy and performance, in order to test our hypotheses. Surprisingly, the occupancy of $IS = H$ doesn't decrease monotonically with increasing costs. In fact, there is a range where it increases with rising costs, as shown in Figure 3.22.



(a) Average occupancy of $IS = H$ for different switching costs

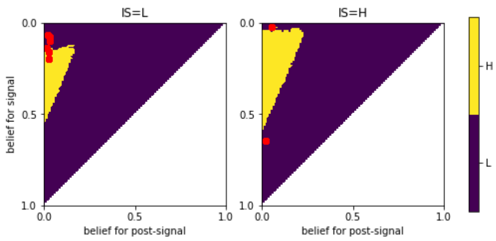


(b) Performance curves for different switching costs

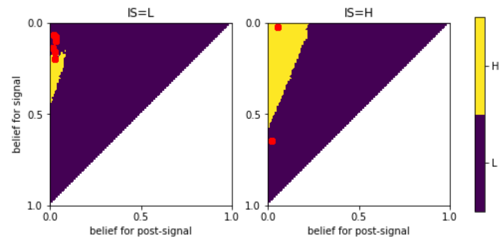
Figure 3.22: Average behavior for different switching costs with all other parameters fixed at default values. The plots include hit, false alarm rates, d' curves, and average occupancy rates for signal, non signal and all trials

To understand why this happens, we compare the optimal policy functions for two example switching costs $= 0.04, 0.08$. As visible from the plots in Figure 3.23, for the higher cost, the first boundary at $IS = H$ is shifted to very low $b(1), b(2)$ – it becomes optimal to maintain the $IS = H$ state if belief in favor of states 1 and 2 are low. Hence, the overall occupancy increases in this case. If switching costs are increased further, the frequency of choosing $IS = H$ is low enough to outweigh this lowered threshold for maintenance, hence ultimately causing the occupancy to reduce.

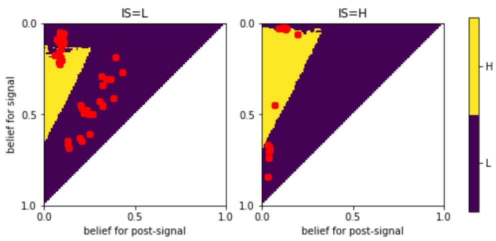
How do the performance curves look for increasing switching costs? d' shows a decreasing trend with with increasing costs, as we had predicted. This drop is mainly due to reduced



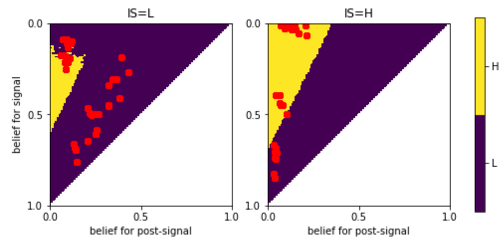
(a) $t = 4$, switching costs $= 0.04$



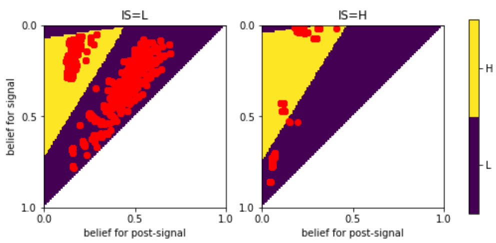
(b) $t = 4$, switching costs $= 0.08$



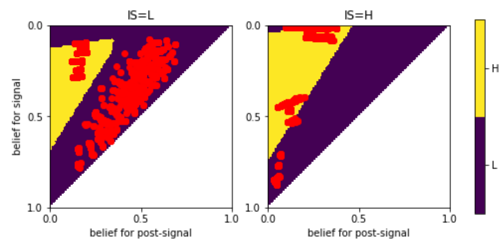
(c) $t = 7$, switching costs $= 0.04$



(d) $t = 7$, switching costs $= 0.08$



(e) $t = 10$, switching costs $= 0.04$



(f) $t = 10$, switching costs $= 0.08$

Figure 3.23: A comparison of policy function plots for selected time points for the whole state space at switching costs $= 0.04, 0.08$. Note the differences in positions of the two boundaries between the two cases. These are overlaid by scatter plots obtained from forward runs

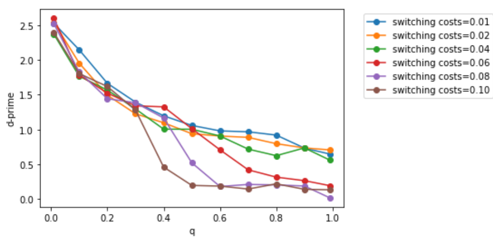
hit rates as seen in Figure 3.22.

Interactions with switching costs

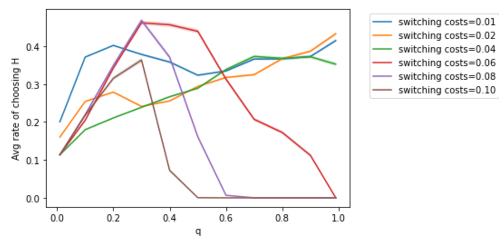
How does changing switching costs interact with changes in q and also in p_{signal}? We predict that in addition to the effects we expect with each of the parameters, there wouldn't be any additional interactions.

Figure 3.24 shows the effects of changing switching costs and p_{signal} on d' and the occupancy rates. The performance curves clearly show a decreasing trend with switching costs.

However, there is no apparent trend with p_{signal} except for the decline at very high signal probabilities. On the other hand, there exists a decreasing trend in the occupancy with increasing switching costs at all signal probabilities except at $p_{signal} = 0.5$ where the trend was is not monotonic. For a specific switching cost, the occupancy has an inverted-U trend with p_{signal}. At high costs, the decay in occupancy at extreme signal probabilities is higher than at lower costs.



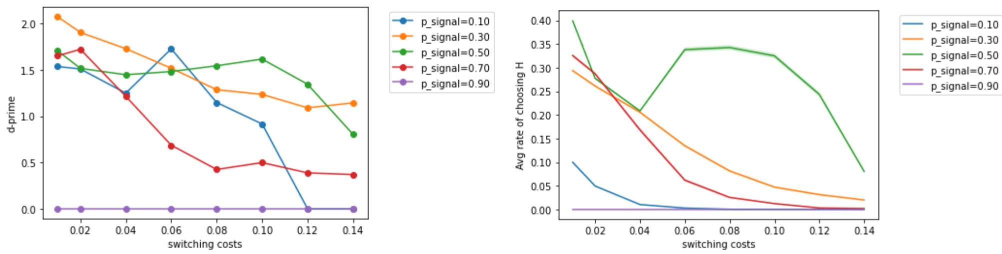
(a) d' curves for different switching costs and q



(b) Occupancy curves for different switching costs and q

Figure 3.24: Interaction between switching costs and q

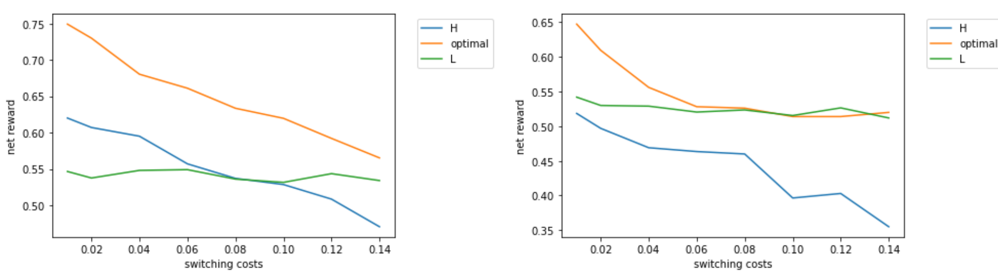
Figure 3.25 shows the effects of changing switching costs and q on d' and the occupancy rates. The d' curves show a decreasing trend with q. d' decreases with switching costs only high q. This is because when costs are high and signals are short, the costs required to improve performance outweigh the rewards from the slight improvement in performance. This is shown in Fig 3.26 where at higher q, the net reward from the optimal policy is no better than choosing the lower attention state. The greater the switching costs, greater is the decline in performance with q. The occupancy curves show different trend with different costs. For lower costs, the occupancy increases with q since it is still advantageous to choose H to detect harder signals. However, if the costs are prohibitively high, they lead to a decline in occupancy with q.



(a) d' curves for different switching costs and p_{signal}

(b) Occupancy curves for different switching costs and p_{signal}

Figure 3.25: Average behavioral trends with varying switching costs and $p_s ignal$. The plots of d' and occupancy curves show an interaction between switching costs and p_{signal}



(a) $q = 0.2$

(b) $q = 0.6$

Figure 3.26: Behavior of the model with different switching costs under the three conditions: choosing $IS = H$ throughout, choosing IS according to the optimum policy (optimal condition), choosing $IS = L$ throughout. The plots show d', average total reward, average costs and net reward for the three conditions

Chapter 4

Discussion and Conclusion

In this work, we considered optimal behavior in a signal detection task with temporal uncertainty, assuming that the task structure (i.e., signal probability, reward contingencies, observation reliabilities, and trial structure) is known. We conducted a resource-rational analysis, taking into account putative attentional costs associated with improving the quality of evidence. Optimal behavior is the one which maximises overall reward while also minimising attentional costs. We conducted this optimisation at a within-trial timescale using dynamic programming.

4.1 Optimal policy

Doing this, we find an optimal policy – the optimal set of actions to take at every possible state of the world. In our case, this included how attentive to be at each particular time step – hence controlling the quality of information (observations) received. We find that the policy is time-varying. That is, for the 'same' (time-independent) state of the world, the optimal attention level indeed depends on the time of occupying that state within a trial. In general, we found that the region in the state space where it is optimal to pay attention grows with time. This means that as the time of the final decision gets nearer, more attention should be paid. Furthermore, we found that there existed two critical boundaries in the belief space (for a particular attentional state) which determined the choice of attention at each time step: a 'lower' boundary that sets a minimum amount of evidence needed in favor of the signal being on (or that it has passed) for it to be worth paying attention; and a second, 'higher' boundary that sets the maximum amount of evidence that is enough to stop paying attention – the latter sets the threshold beyond which one can be sure that the signal is or was on.

In practice, given the initial condition that every trial begins with no signal, combined with a particular parametrization of the model, not all regions of the belief/state space are reachable. Furthermore, some regions are occupied more often than others. To understand how attention is allocated in an actual trial, we averaged measures of attention across multiple trials. In addition, to gain insights into model behavior, we found averages of various performance measures, costs and rewards across multiple trials. Chapter 3 contains a detailed exposition of the effects of various parameters on these measures.

We next summarise the most important results and also discuss the insights they provide, their implications and significance.

4.2 Trends in attention

We looked at the average occupancy of the 'high' attention state in a trial as a measure of the amount of attention that needs to be paid. This was calculated by averaging (across multiple trials) the percentage of a trial that is spent in the H state. The trends that we found were complex with many non-monotonic and nonlinear patterns with different factors like signal length, attention costs and signal probability.

A naive prediction was that more attention needs to be paid to detect short signals. However, we found that the amount of attention that is paid as signal length reduces depended on the overall ‘toughness’ of the trial. For a relatively easy trial, it made sense to pay the costs of attention to improve performance (and hence final task rewards). However, for a trial that was tougher (due to being longer or having noisier observations), the costs required to improve performance at shorter signals far outweighed the resultant rewards hence making it not worth to pay attention.

We also found a similar effect when the prior signal probability was too high or too low. If the certainty about the trial identity was already very high, paying the costs of attention for a marginal improvement in detection was not ‘worth it’. The slight decline in performance at the these extreme probabilities (described in the previous section) can be explained by this rational inattention.

4.3 Attention allocation within a trial

We were also able to show how attention must be distributed within a trial across various task factors. We reported the following results from averaging the occupancy of the high attentional state (at each time step within a trial) across multiple trials.

Importantly, the optimal policy involved a non-uniform distribution of attention through the trial. The tendency was for attention to be concentrated as late in the trial as possible. If it was possible to get to a future state where more information could be obtained by paying the costs, then there was no advantage in paying for attention earlier. The effects of signal length on attention allocation is a perfect example of this – when signals were known to be long, more attention was paid later in the trial simply because signals were expected to stay on for longer. Even when attention had to be paid earlier in the trial (for example, to not miss a short signal), the distribution was not uniform across a trial. Usually, there is an early and a late peak in attention so that signals that might not be caught earlier can simply be caught later in the trial without having to pay much attention in the middle.

4.4 Trends in performance

We described optimal inference (not yet including resource constraints) in our chosen decision problem. We replicated the results of [James P. Egan and Schulman, 1961]: as the trial length increases, the discriminability, d' declines. Furthermore, we found that longer signal lengths lead to better performance as measured by d', although the decrement with trial length still occurs. This drop in performance is predicted even when no costs of attention are included. This points to a fundamental limit on which signals can be detected. As [James P. Egan and Schulman, 1961] speculate, the longer the trial, the more noise can be confused with the actual signal, hence leading to performance decrement – the signal is harder to detect.

We then considered the case including attentional costs. While similar patterns as above were seen with changing signal and trial length, the absolute performance itself did not reach the same levels. This is simply because the agent must trade-off rewards from correct detections with costs from collecting better quality evidence.

We found interesting interactions between increasing costs and changing signal lengths. There was no performance decline with increasing costs at large signal lengths but a decline with costs occurred only when signals are short. We demonstrated that this effect occurred because of how trade-offs between rewards and costs play out at different parameter ranges: when costs are high and signals are short, the costs required to improve performance outweigh the rewards from the slight improvement in performance. However, increasing costs at large signal lengths caused a shift in hit and false alarm rates and therefore, in the response criterion.

Our result is similar to the following effects which have been shown in previous studies: changing memory load and background noise (which changes signal discriminability) causes the overall performance (d') to decline significantly only when both of these are high. A temporal decline in detection performance ('vigilance decrement') also occurs only in this parameter range [Parasuraman, 1979]. Increasing memory load when signal quality is already high only causes a temporal shift in response bias with no vigilance decrement [Parasuraman, 1979]. Changing the quantities of signal quality and memory load in a variety of other ways also yield similar results [Parasuraman et al., 1987]. This comparison is justified because in our model, signal length determines it's strength (and so, it's discriminability) while attention costs can be likened to increased resource demands from manipulations like increasing memory load.

We showed that increasing signal probability shifts the response criterion to lower values (more liberal criterion). d' stays largely constant except at very high or low signal probabilities. Previous studies have also shown that increasing signal probability shifts response criterion to lower values while the performance (d') remains unaffected [Davies and Parasuraman, 1977].

Finally, we showed an interaction effect between signal length and signal probability. We found that decreasing both of them caused the response criterion to be shifted to higher values while also decreasing d'. We believe this change is similar to the manipulation of event rate in previous studies. Event rate is a complex parameter that presumably causes an increase in background noise (and so decreases signal quality) and a decrease in perceived signal (target) probability. It has been observed that increasing event rate shifts criterion to higher values while also decreasing the d' [Davies and Parasuraman, 1977].

4.5 Limitations and future work

While we talked about optimal behavior within a trial, we did not touch upon optimality across trials. We would like to extend our analysis beyond a trial because there have been interesting across-trial effects shown in these tasks that we did not provide a rational account of. A notable example is the famous vigilance decrement. However, we can provide descriptive explanations for these. For example, if we allow a parameter like attentional costs or perceived signal probability to change across trials, this could explain vigilance decrement. An important future step would be to construct an across-trial normative model to find rational behavior across trials.

While we described interesting attention patterns in our model, we did not provide a qualitative or quantitative comparison with real attention data (either behavioral observations or neural data). However, we believe that investigating optimal attention allocation at the within-trial level is an important first step to make progress in this area. A natural next step would be to turn to the effects of phasic Acetylcholine which has been shown to mediate signal detection in sustained attention tasks [Howe et al., 2013, Gritton et al., 2016]. These studies have shown in rats performing a signal detection task that phasic ACh signals cause signal detection on signal trials. They are necessarily preceded by thalamic-glutamatergic signals which are evoked by signals [Sarter et al., 2016]. This is similar to an effect in our model: switching to the higher attentional state happens only when there is a minimum amount of evidence in favor of the signal. In the future, we would like to investigate further the possible role of phasic ACh in shifting attentional states at fast timescales.

4.6 Significance

Leading accounts of vigilance behavior attribute the observed performance decline trends to limits on information processing abilities of animals/ subjects [Thomson et al., 2015]. However, they qualitatively and intuitively explain the effects of resource demands and so do not explain why these performance declines occur only for some manipulations of task difficulty and are in some cases due to response criterion shifts (and not changes in sensitivity) [Parasuraman et al., 1987].

In this work, we are able to provide a rational account of some of these effects in terms of a cost-benefit analysis. We also show response or sensitivity decrements (or both) occur for different manipulations of task parameters without having to invoke fundamentally different mechanisms for the two. In our model, all effects fall out of assuming attention is costly. We not only demonstrate that these complex trends and interactions can occur using a relatively simple model, but also show that this behavior is resource-optimal.

We are also able to theoretically show how optimal allocation of attentional resources should look like within a trial and how it must vary across different task parameters. We believe this is an exciting first step in trying to understand attention allocation in the brain and particularly, the role of neuromodulators, especially ACh in shifting attentional states in the brain.

4.7 Conclusion

In summary, we investigated optimal behavior in signal detection tasks with temporal uncertainty, assuming resource limitations on attention. We estimated an optimal policy for allocating attention within each trial of the task. Interestingly, we found that a rational agent must pay attention only when there is enough (but not overwhelming) evidence in favor of a signal. Additionally, attention allocation was not uniform within a trial: for the same amount of evidence, it was optimal to pay more attention later in the trial. The distribution of attention varied with different factors: when a signal was longer or less probable, it was optimal to pay most attention very late in the trial. The overall level of attention per trial was also a result of complex interactions between factors like signal length and detectability, signal probability and attention costs. When a trial was too tough or when there was too much bias towards a certain hypothesis, reward-cost trade-offs dictated that there was no advantage in paying for more information through attention.

We also found interesting performance patterns with our model. Performance (as measured by the sensitivity index, d') declined with decreasing signal length while it decreased with attention costs only at short signal lengths. The sensitivity remained almost unchanged with signal probability while the response criterion was shifted. Decreasing signal length, however, led to sensitivity decline at extreme signal probabilities. Equivalent results have been shown experimentally in vigilance tasks which are known to involve a similar temporal uncertainty.

References

[Anderson, 2004] Anderson, J. (2004). Cognitive Psychology and Its Implications, vol. 6th ed,. Worth Publishers.

[Arnold et al., 2002] Arnold, H., Burk, J., Hodgson, E., Sarter, M. and Bruno, J. (2002). Differential cortical acetylcholine release in rats performing a sustained attention task versus behavioral control tasks that do not explicitly tax attention. Neuroscience *114*, 451–460.

[Assad, 2003] Assad, J. A. (2003). Neural coding of behavioral relevance in parietal cortex. Current Opinion in Neurobiology *13*, 194–197.

[Caggiano and Parasuraman, 2004] Caggiano, D. and Parasuraman, R. (2004). The role of memory representation in the vigilance decrement. Psychonomic Bulletin and Review *11*, 932–937.

[Clark and Dukas, 2003] Clark, C. W. and Dukas, R. (2003). The behavioral ecology of a cognitive constraint: limited attention. Behavioral Ecology *14*, 151–156.

[Cohen and Maunsell, 2009] Cohen, M. and Maunsell, J. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nature neuroscience *12*, 1594–600.

[Cohen and Maunsell, 2010] Cohen, M. and Maunsell, J. (2010). A Neuronal Population Measure of Attention Predicts Behavioral Performance on Individual Trials. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 15241–53.

[Davies and Parasuraman, 1977] Davies, D. R. and Parasuraman, R. (1977). Cortical Evoked Potentials and Vigilance: A Decision Theory Analysis pp. 285–306. Boston, MA: Springer US.

[Gaulin and McBurney, 2003] Gaulin, S. J. C. and McBurney, D. H. (2003). Evolutionary Psychology, pp. 81–101. Prentice Hall.

[Ghose and Maunsell, 2002] Ghose, G. and Maunsell, J. (2002). Attentional modulation in visual cortex depends on task timing. Nature , *419*, 616–620.

[Green and Swets, 1966] Green, D. M. and Swets, J. A. (1966). Signal detection theory and psychophysics. John Wiley.

[Gritton et al., 2016] Gritton, H. J., Howe, W. M., Mallory, C. S., Hetrick, V. L., Berke, J. D. and Sarter, M. (2016). Cortical cholinergic signaling controls the detection of cues. PNAS *113*, 1089–1097.

[Hasselmo and Sarter, 2011] Hasselmo, M. E. and Sarter, M. (2011). Modes and Models of Forebrain Cholinergic Neuromodulation of Cognition. Neuropsychopharmacology Reviews *36*, 52–73.

[Helton and Warm, 2008] Helton, W. S. and Warm, J. S. (2008). Signal salience and the mindlessness theory of vigilance. Acta Psychologica *129*, 18–25.

[Howe et al., 2013] Howe, W. M., Berry, A. S., Francois, J., Gilmour, G., Carp, J. M., Tricklebank, M., Lustig, C. and Sarter, M. (2013). Prefrontal Cholinergic Mechanisms Instigating Shifts from Monitoring for Cues to Cue-Guided Performance: Converging Electrochemical and fMRI Evidence from Rats and Humans. Journal of Neuroscience *33*, 8742–8752.

[James, 1890] James, W. (1890). The Principles of Psychology, vol. 1,. New York: Henry Holt.

[James P. Egan and Schulman, 1961] James P. Egan, G. Z. G. and Schulman, A. I. (1961). Interval of Time Uncertainty in Auditory Detection. The Journal of the Acoustical Society of America *33*, 771.

[Jaramillo and Zador, 2011] Jaramillo, S. and Zador, A. M. (2011). The auditory cortex mediates the perceptual effects of acoustic temporal expectation. Nature neuroscience *14*, 246—251.

[Laughlin et al., 1998] Laughlin, S., Steveninck, R. and Anderson, J. (1998). The metabolic cost of neural information. Nature neuroscience *1*, 36–41.

[Lieder and Griffiths, 2020] Lieder, F. and Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. Behavioral and Brain Sciences *43*, e1.

[Mackowiak and Wiederholt, 2009] Mackowiak, B. and Wiederholt, M. (2009). Optimal Sticky Prices under Rational Inattention. American Economic Review *99*, 769–803.

[McGaughy et al., 1996] McGaughy, J., Kaiser, T. and Sarter, M. (1996). Behavioral vigilance following infusions of 192 IgG-saporin into the basal forebrain: selectivity of the behavioral impairment and relation to cortical AChE-positive fiber density. Behavioral neuroscience *110*, 247—265.

[McGaughy and Sarter, 1995] McGaughy, J. and Sarter, M. (1995). Behavioral vigilance in rats: task validation and effects of age, amphetamine, and benzodiazepine receptor ligands. Psychopharmacology *117*, 340–357.

[Moray et al., 1976] Moray, N., Fitter, M., Ostry, D., Favreau, D. and Nagy, V. (1976). Attention to Pure Tones. Quarterly Journal of Experimental Psychology *28*, 271–283.

[Newman et al., 2012] Newman, E., Gupta, K., Climer, J., Monaghan, C. and Hasselmo, M. (2012). Cholinergic modulation of cognitive processing: insights drawn from computational models. Frontiers in Behavioral Neuroscience *6*, 24.

[Parasuraman, 1979] Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. Science *205*, 924–927.

[Parasuraman et al., 1987] Parasuraman, R., Warm, J. S. and Dember, W. N. (1987). Vigilance: Taxonomy And Utility pp. 11–32. New York, NY: Springer New York.

[Posner et al., 1980] Posner, M., Snyder, C. and Davidson, B. (1980). Attention and the detection of signals. Journal of Experimental Psychology: General , *109*, 160–174.

[Posner MI, 1990] Posner MI, P. S. (1990). The attention system of the human brain. Annual Reviews Neuroscience , *13*, 25–42.

[Prinzmetal et al., 2005] Prinzmetal, W., McCool, C. and Park, S. (2005). Attention: Reaction Time and Accuracy Reveal Different Mechanisms. Journal of experimental psychology. General , *134*, 73–92.

[Reynolds and Chelazzi, 2004] Reynolds, J. and Chelazzi, L. (2004). Attentional Modulation of Visual Processing. Annual review of neuroscience , *27*, 611–47.

[Sarter et al., 2001] Sarter, M., Givens, B. and Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. Brain Research Reviews , *35*, 146–160.

[Sarter and Lustig, 2019] Sarter, M. and Lustig, C. (2019). Cholinergic double duty: cue detection and attentional control. Current Opinion in Psychology , *29*, 102–107.

[Sarter et al., 2016] Sarter, M., Lustig, C., Berry, A. S., Gritton, H., Howe, W. M. and Parikh, V. (2016). What do phasic cholinergic signals do? Neurobiology of Learning and Memory , *130*, 135–141.

[Thomson et al., 2015] Thomson, D. R., Smilek, D. and Besner, D. (2015). Reducing the vigilance decrement: The effects of perceptual variability. Consciousness and Cognition , *33*, 386–397.

[Verghese, 2001] Verghese, P. (2001). Visual Search and Attention: A Signal Detection Theory Approach. Neuron , *31*, 523–535.

[Warm et al., 2008] Warm, J. S., Parasuraman, R. and Matthews, G. (2008). Vigilance Requires Hard Mental Work and Is Stressful. Human Factors , *50*, 433–441.