Giving Centre Stage to Top-Down Inhibitory Mechanisms for Selective Attention

Sock Ching, Low

TESI DOCTORAL UPF / 2020

Thesis supervisor

Dr. Paul F.M.J. Verschure,

Departament de Tecnologies de la Informació i les Comunicacions

Universitat Pompeu Fabra Barcelona

In appreciation

As a pre-teen, I scoffed at my parents when they insisted that no man is an island; growing up meant realising that they were right. It is thus apt that they are the first people I would like to thank, for their unwavering emotional support and the security to follow my curiosity abroad. My extended family has also always made clear that I can depend on them, and for that I am grateful.

Another person that was pivotal to this thesis is my supervisor, Dr Paul Verschure. Through him, I got to know the SPECS lab, whose members, past and present, have given me a warm place to work (although this is, unfortunately, either too literal or too figurative for most of the year). They have been an undeniable pillar in my life.

Science may be 24/7, but my human body is not, at least not yet; many people have contributed to keeping my mind and soul healthy. It includes those whom I got to know through the CSIM master program. Thank you for putting in the effort to keep our ties alive, and for not trying very hard to do typically "cool" things when we do meet. Also, the Salty Frisbee tribe has consistently provided me with exercise, sand in my pockets, and good friends. Many of you have been a source of stability for me throughout the PhD journey, which is ironic considering how anarchic the group is. I also want to thank my friends who I left back in Singapore, for being there for me even if I can be very absent. Eating my way through Singapore whenever I visit is only as enjoyable as it is because you're there with me. There are also the people who were friends of friends, but who are now *my* friends. You have all kept me going with food, music, and conversation. Lastly, I need to mention Zoe, Luca, Neola, and Marcel, who have become pseudo-family. I'm sure we have all grown over these years together, and hopefully for the better.

Abstract

Selective attention determines the sensory signals that are processed at higher levels at the expense of others and is biased by higher-order brain regions which anticipate task-relevant stimuli and increase neural sensitivity to them in the sensory cortex. Often, this is thought to occur through excitation of selected neurons, but some studies have suggested that it is not the full description of the process. Increasingly, evidence has pointed to an alternative, top-down inhibitory biasing mechanism. Here, we investigated such an inhibitory model of attention. We first showed how sensitivity to stimulus features known to be task-irrelevant are reduced through top-down suppression. Secondly, we demonstrated a biologically grounded spiking model's ability to modulate information processing and benchmarked it to physiology. Lastly, we explored the interaction between the excitatory and inhibitory models of top-down attention in a foraging agent. Our results support the inhibitory model of top-down attention as a biological attentional mechanism and show how it fits into the current zeitgeist of top-down attentional mechanisms.

Abstract in Catalan

L’atenció selectiva determina els senyals sensorials que es processen a nivells superiors a costa dels altres. Està esbiaixada per regions cerebrals d’ordre superior que anticipen estímuls rellevants per a la tasca i augmenten la sensibilitat neuronal a l’escorça sensorial. Sovint, es creu que això es produeix mitjançant l'excitació de neurones seleccionades, però alguns estudis han suggerit que no és la descripció completa del procés. Cada vegada més, l’evidència apunta cap a un mecanisme alternatiu de polarització inhibitiva de dalt a baix. Aquí hem investigat, aleshores, un model d’atenció inhibitori. Primer, vam demostrar com es redueix la sensibilitat a les funcions d’estímul irrellevants per tasques mitjançant la supressió de dalt a baix. En segon lloc, vam demostrar la capacitat d’un model d’espiga basat en la biologia per modular el processament de la informació i l’hem comparat amb la fisiologia. Per últim, hem explorat la interacció entre els models excitadors i inhibidors d’atenció de dalt a baix en un agent de cerca d’aliments. Els nostres resultats donen suport al model inhibitori de l’atenció de dalt a baix com a mecanisme d’atenció biològica i mostren com s’adapta al ‘zeitgeist’ actual dels mecanismes d’atenció de dalt a baix.

Table of contents

Page

Abstract v

List of publications ix

List of figures xi

List of tables xii

1. INTRODUCTION 1

2. TOP-DOWN ATTENTION IS MODULATED BY DEMANDS ON COGNITIVE RESOURCES 11

2.1 Methods 13

a) Experimental protocol 16

b) Displacement detection task 18

c) Auditory span task 19

d) Outcome measures 20

e) Behavioural analysis 21

f) Visualisation of the anticipatory field 22

g) Generalised linear mixed-effects model 22

2.2 Results 24

a) Behavioural analysis 24

b) Clinical measures are correlated with explicit but not implicit responses 26

c) Generalised linear mixed-effects model 27

2.3 Conclusions 31

3. SACCADIC DYNAMICS ARE MODULATED BY WORKING MEMORY LOAD 35

3.1 Methods 37

a) Experimental protocol 38

b) Catch detection task 40

c) Sternberg auditory n-back task 41

d) Data analysis 42

3.2 Results 44

a) Catch detection is affected by size more than by hue 44

b) Participants remained engaged in both tasks 47

c) Performing both tasks concurrently led to task interference 48

d) Eye movements increase when working memory is loaded 50

e) Eye movements increase specifically during memory recall 51
3.3 Conclusions 52

4. THE NEURAL SUBSTRATE OF SELECTIVE ATTENTION 55
4.1 Methods 58
a) Network-level description 59
b) Neuron-level description 63
c) Mean-field model 65
d) Experimental configuration 66
4.2 Results 67
a) Performance in attentional tasks 67
b) Physiological benchmark 75
4.3 Conclusions 78

5. COMPLEMENTARY INTERACTIONS BETWEEN CLASSICAL AND TOP-DOWN DRIVEN INHIBITORY MECHANISMS OF ATTENTION 81
5.1 Methods 84
a) Control architecture 85
b) Task description 89
c) Model formalisation 92
d) Bottom-up sensory competition 93
e) Top-down biasing components 93
f) Data analysis 96
5.2 Results 97
a) Top-down attentional biasing increased probability of choosing a reward over a distractor 97
b) Attentional mechanisms led to quicker trials when there were more rewards than distractors 98
c) Combining excitatory and inhibitory attentional mechanisms improved performance regardless of reward proportion 99
5.3 Conclusions 101

6. CONCLUSIONS 105

Bibliography 109

List of publications

Contributions directly related to the thesis

Low, S.C., van Wijngaarden, J.B.G., Verschure, P.F.M.J., 2016. Modelling the Effect of Cognitive Load on Eye Saccades and Reportability: The Validation Gate. In: Conference on Biomimetic and Biohybrid Systems. Springer, pp. 459–466.

Low, S.C., Zucca, R., Verschure, P.F.M.J., 2017, September. The role of inhibition in selective attention. In: The International Conference on Artificial Neural Networks (ICANN). European Neural Network Society, Alghero/Sassari, Sardinia, Italy.

Low, S.C., Zucca, R., Verschure, P.F.M.J., 2018, February. Predictive coding gives us blinkers. In: Bridging Attention and Prediction. University of Barcelona, Barcelona, Spain.

Maier, M.*, Low, S.C.*, Ballester, B.R., Bañuelos, N.L., Oller, E.D. and Verschure, P.F., 2018, October. Depression Modulates Attentional Processing After Stroke. In: International Conference on NeuroRehabilitation . Springer, Cham. pp. 702-706.

Maier, M.*, Low, S.C.*, Ballester, B.R. and Verschure, P. F., 2020. Cognitive abilities modulate top-down function in an attentional task: The validation gate task as a novel diagnostic method. Manuscript in preparation.

Low, S.C. and Verschure, P.F. and Santos-Pata, D., 2020. Working Memory Load Increases Saccadic Eye Movements During Recall. Available at SSRN: https://ssrn.com/abstract=3671967 or http://dx.doi.org/10.2139/ssrn.3671967

Low, S.C., Verschure, P.F.M.J. and Santos-Pata, D., 2020. Working memory load increases saccadic eye movements during recall. Manuscript submitted for publication.

Low, S.C., Zucca, R. and Verschure, P.F.M.J., 2020. Putting the spotlight on inhibition: modelling selective attention through inhibitory and excitatory pathways. Manuscript in preparation.

Low, S.C., Vouloutsi, V., Verschure, P.F.M.J., 2020, October. Complementary Interactions Between Classical and Top-Down Driven Inhibitory Mechanisms of Attention. In: Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence. BICA Society, Natal, Brazil.

Other contributions

Puigbò, J.-Y., van Wijngaarden, J., Low, S.C., Verschure, P.F.M.J., 2016. Synaptogenesis: Constraining synaptic plasticity based on a distance rule, in: International Conference on Artificial Neural Networks. Springer, pp. 28–35.

Santos-Pata, D., Zucca, R., Low, S.C. and Verschure, P.F., 2017. Size matters: How scaling affects the interaction between grid and border cells. Frontiers in computational neuroscience, 11, p. 65.

Moulin-Frier, C., Fischer, T., Petit, M., Pointeau, G., Puigbo, J.-Y., Pattacini, U., Low, S.C., Camilleri, D., Nguyen, P., Hoffmann, M., 2017. DAC-h3: a proactive robot cognitive architecture to acquire and express knowledge about the world and the self. IEEE Trans. Cogn. Dev. Syst.

List of figures

Page
Figure 1. Primate visual hierarchy. 2
Figure 2. Validation Gate hypothesis, visualised. 8
Figure 3. The set-up, the schema of task presentation and experimental protocol. 18
Figure 4. Schemata of a circle displacement and corresponding eye movement. 21
Figure 5. Proportions of displacements detected. 24
Figure 6. Correlations between explicit detection proportion, latency and neuropsychological test battery. 26
Figure 7. Partial correlation across cognitive loads for explicit detection proportion. 27
Figure 8. Probability of explicit detection in each cognitive load condition for the three groups of cognitive impairment. 28
Figure 9. Experimental equipment and protocol. 38
Figure 10. Catch detection performance is modulated by task variables and not Sternberg task stage. 45
Figure 11. Behavioural markers of task interference. 49
Figure 12. Eye movements increase in conjunction with demand for working memory. 51
Figure 13. Model architecture. 60
Figure 14. Performance and neuronal responses in a target detection task. 69
Figure 15. Performance in a 2D oddball task. 74
Figure 16. Thalamic response to increasing PFC weight. 77
Figure 17. Control architecture of the agent. 87
Figure 18. Experimental setup. 90
Figure 19. Selected concepts in the contextual layer. 95
Figure 20. Performance of the agent in the two experiments. 97

List of tables

Page
Table 1. Description of clinical scales and test batteries used to diagnose and measure cognitive impairment..... 14
Table 2. Patients' characteristics. 17
Table 3. Results of post hoc Monte Carlo permutation. 25
Table 4. Overview of generalised mixed-effects models (GLME) used in comparisons. 28
Table 5. Generalised mixed-effects model (GLME) comparisons executed. 30
Table 6. Final estimates of the model. 30
Table 7. Post-hoc Monte Carlo permutation statistics. 46
Table 8. Distribution characteristics for behavioural data. 47
Table 9. Distribution characteristics for eye-tracking data. 50
Table 10. Connection parameters for the target detection task..... 61
Table 11. Connection parameters for the oddball detection task... 62
Table 12. Thalamic neuron parameters. 65
Table 13. Cortical neuron parameters. 65
Table 14. Simulation parameters. 67
Table 15. Post-hoc Monte Carlo permutation testing statistics for mean differences in latencies of trials. 99
Table 16. Post-hoc Monte Carlo permutation testing statistics for mean differences in probability of success. 100

1. INTRODUCTION

Top-down and bottom-up interactions in selective attention

Sensory processing in humans, and many other non-human animals, is selective. This happens largely because we receive a wide array of sensory information about stimuli in our environment, but our capacity to process this information is limited (Broadbent, 1958; Duncan, 1984; Kahneman, 1973; Schneider and Shiffrin, 1977). The mechanism through which certain stimuli are selected for processing instead of others, allowing for differential processing of simultaneous sources of information (Johnston and Dark, 1986), is referred to as selective attention. While selective attention is not modality-specific, and many experiments do exploit cross-modality paradigms, most studies focus on the visual domain followed by the auditory (Hutmacher, 2019). For this reason, examples and descriptions in this thesis will heavily refer to visual attention but the concepts can be extrapolated to other sensory domains as well.

Selective attention is considered to be a result of the integration of two processing streams—top-down and bottom-up—with the winning stimulus having "further access to memory systems for mnemonic encoding and retrieval and to motor systems for guiding action and behaviour" (Kastner and Ungerleider, 2000). The concept of predictive models of the environment has a long history (Helmholtz, 1962) and has been developed into what is known as the Bayesian brain (Clark, 2013; Friston, 2012, 2005). In the context of selective attention, this implies that top-down processes bias bottom-up sensory signals towards those that are anticipated to be, information-wise, more valuable. Bottom-up information depends on the physical features of stimuli relative to the background, like their orientation and colour, while top-down information, stemming from higher-order brain regions, is based on prior knowledge and goals. In psychophysiology, these two processes have been reliably shown by tasks that dissociate them, like the oddball task, with one such study showing that evoked response potentials related to novel stimuli are subject to habituation within an experimental session while those related to typical stimuli are not (Debener et al., 2002). The processing of visual information is hierarchical; visual information

flows from the retina to the occipital lobe where it then moves across the neocortex to the frontal lobe (Figure 1-left). Each brain area represents stimuli and their features at its own level of abstraction, and top-down attentional influence on neural representation of stimuli can be observed at various stages of the visual hierarchy. For example, there is substantial evidence that attention modulates the rate of neuronal spiking, such as when firing rates of neurons encoding the attended stimulus are increased (Moran and Desimone, 1985), and neuronal oscillatory coherence, which has been shown to increase in the gamma frequency range in populations representing the attended stimulus at mid- to high-level stages of the visual hierarchy (Gregoriou et al., 2009; Saalmann et al., 2007; Taylor et al., 2005) while decreasing that in the V1 (Chalk et al., 2010).

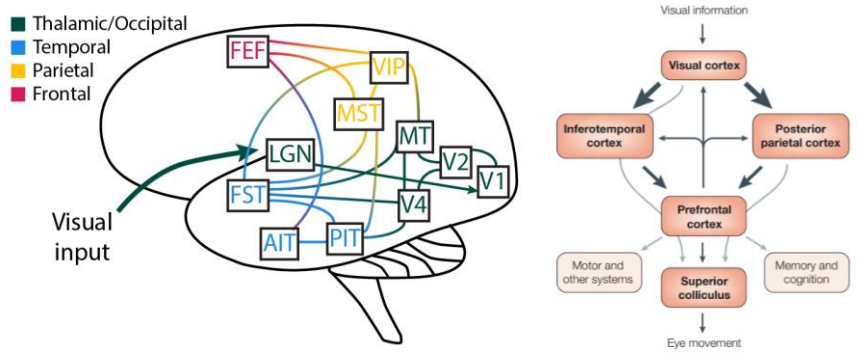


Figure 1. *Primate visual hierarchy.* (Left) Simplified anatomical hierarchy of the macaque visual sensory processing system, with lines between regions indicating reciprocal connections. The colours show the lobe that the area belongs to. Full matrix of connections and more can be found in Felleman and Van Essen (1991) as well as Hagmann et al. (2008). MT: medial temporal; PIT: posterior inferiotemporal; FST: floor of superior temporal; VIP: ventral intraparietal; AIT: anterior inferiotemporal; MST: medial superior temporal; FEF: frontal eye field (Right) Anatomical description from Itti and Koch (2001) of the proposed brain regions that participate in the deployment of visual attention, with the arrow from the visual cortex to the prefrontal cortex through the inferiotemporal cortex representing the ventral stream and that through the posterior parietal cortex representing the dorsal stream.

The visual cortex's (V4) ability to segregate stimuli, which is indicative of top-down attention (McAdams and Maunsell, 1999), is also improved by stimulation of higher-order regions that have been implicated with top-down attention, such as the frontal eye field (Armstrong and Moore, 2007). Furthermore, priming attention by providing evidence for a target's upcoming presentation leads to increased pre-stimulus activity in the visual cortex when measured

with magnetoencephalography or functional magnetic resonance imaging. For example, this has been found for representations of spatial locations (Simpson et al., 2011), stimulus colour (Giesbrecht et al., 2006), and object category such as when differentiating between faces and houses (Esterman and Yantis, 2010). These findings were operationalised in the biased-competition model of selective attention (Desimone and Duncan, 1995), which describes how multiple, concurrently presented stimuli activate sensory neuronal populations that compete through inhibition in the sensory cortex (Reynolds and Heeger, 2009). Top-down processes from higher-order brain regions can bias this competition to favour neurons responding to the relevant stimuli (Kastner and Ungerleider, 2001), often selected through its task-relevance (Hopfinger et al., 2000). Hence, neural signals are strongly dependent on predictions from previous trials than only from the current stimulus itself (Clark, 2013; Marcos et al., 2013) and modality-specific predictions are integrated into selective attention according to their reliability (Mengotti et al., 2018).

As a result of such top-down biasing, selected stimuli have the most active neuronal populations which dominate the biased competition and thus are attended to. Indeed, it has been shown that attending to a stimulus leads to a higher quality of its representations in the inferior temporal lobe (Zhang et al., 2011). Brain regions in both the dorsal and ventral streams have shown attentional feedback modulation, and while the feedforward flow of visual information is relatively well-established, it is less certain whether the feedback pathways for attentional modulation is corticocortical (Saalmann et al., 2007) or corticothalamocortical (Saalmann et al., 2012; Van Essen, 2005).

Framed by such a winner-takes-all approach to the selection of stimuli for processing (Koch and Ullman, 1985), it is often thought that top-down signals are excitatory, boosting the activity of selected neurons to ensure that their activity surpasses that of the others and effectively increasing sensitivity to the preferred stimuli (Itti and Koch, 2001; Martinez-Trujillo and Treue, 2004; Moore and Zirnsak, 2017). In such models, inhibition acts mostly as a tool to sharpen the

contrast between selected stimuli and their close competitors (e.g. Kidd et al., 2005; Reynolds and Desimone, 2003; Schwartz and David, 2018). In this sense, even when through inhibitory means, the attentional system biases sensory processing based on what is task-relevant (e.g. Otazu et al., 2009; Serences et al., 2004).

The study of selective attention in humans is driven in large part by a desire to understand it so as to better treat attentional disorders or ameliorate their negative impact on people's quality of life. Developing a comprehensive model of selective attention would also offer insight into many aspects of daily life with far-reaching consequences, such as lapses in professional judgement, for example when radiologists miss superimposed gorillas in computed tomography scans (Drew et al., 2013), and the problem of habitual multi-tasking, which has been shown to have grave consequences such as when a driver is distracted by a mobile phone (Sanbonmatsu and Strayer, 2013) and is less aware of how safely they drive (Sanbonmatsu et al., 2016). It is therefore important to note that the standard and well accepted excitatory models of attention face several challenges, which undermines those objectives.

For example, at the psychophysical level, perceptual neglect in the form of inattentional and change blindness has been observed despite strong bottom-up saliency of the anomalous stimuli (e.g. Simons and Levin, 1997). In addition, the error rate of the detection of appearance (Chabris et al., 2011) and displacement (Mathews et al., 2015) of salient stimuli has been shown to be modulated by working memory load, implicating the top-down attentional system in these lapses. Similarly, increased working memory load led to reduced signatures of distraction in event-related brain potentials in a multi-modal experiment (SanMiguel et al., 2008). These phenomena suggest that top-down processes can also bias attention through an inhibitory-focused mechanism; it selectively decreases, instead of increases, sensitivity to regions, objects or features. As a mechanism of signal processing, inhibition has the functional advantage of increasing signal-to-noise ratio and reducing the volume of information to an amount that processing capacity can handle. In addition, inhibition necessarily has a lower limit, or a maximum effect, which is simply

the silencing of neural representations, whereas excitation has no theoretical upper limit which could lead to unsustainable levels of activity.

The excitatory model of attention also faces a neurophysiological challenge as it predicts that attentional signals should propagate through the hierarchy of cortical structures serving perception. This would imply that top-down signals should appear first in frontal cortical areas and from there proceed towards the primary sensory cortices. However, there is evidence that the attentional hierarchy is not that clear-cut. For instance, lesions in the frontal cortex affect attentional processing but do not abolish the effect of top-down signals in earlier visual areas (Paneri and Gregoriou, 2017): for instance, healthy monkeys had representations of salient stimuli in the frontal cortex and the parietal cortex with no temporal latency even though the frontal cortex is considered downstream from the parietal cortex (Katsuki and Constantinidis, 2012). These results suggest that attention might rely on a variety of mechanisms beyond the specific enhancement of task-relevant features.

At their core, these findings address the fact that a stimulus' bottom-up saliency and its features' predictability is independent of its task-relevance. In many visual search tasks, the number of task-relevant stimulus features is limited. For example, if the task is to find red objects out of various coloured objects, the task-relevant feature is limited to a single colour, which is red. However, this is not always the case in natural environments. Sometimes, the task-relevant stimulus feature is unspecific and therefore less predictable than task-irrelevant features. In this case, if the task is to find non-red objects, the task-relevant feature has now expanded to all colours other than red, while the task-irrelevant feature is what is limited to a single colour. As a result, the task-relevant features are now more unpredictable than the task-irrelevant feature. It would be more efficient for attentional processes to target smaller feature-spaces. Thus, in traditional search tasks (i.e. find the red object) where the task-relevant feature is in the smaller feature-space, top-down attention can target task-relevant features for enhancing. In contrast, in the inverse scenario (i.e. find the non-red object) where the task-

irrelevant features are now in the smaller feature-space, top-down attention can instead target task-irrelevant features for inhibition.

The concept of selectively inhibiting potential distractors is not new, with an early proposal describing how analysis of distractors can lead to their inhibition as a supplement to the more popular mechanism of amplifying representation of target stimuli (Houghton and Tipper, 1994). Subsequent neurophysiological findings on the connections between the frontal cortex, the thalamus, and the thalamic reticular nucleus have led to the proposal that such an inhibitory mechanism could be grounded in thalamocortical substrates (e.g. Saalmann et al., 2012; Sherman, 2016; Zikopoulos and Barbas, 2006). In support of this, it has been shown that mental fatigue decreases the ability to suppress irrelevant information (Faber et al., 2012), that monkeys' inhibitory thalamocortical pathways are activated for sensory modalities that they know to be task-irrelevant in a visual-auditory task (McAlonan et al., 2006), and that genetically modified mice expressing reduced activity in the proposed inhibitory pathways exhibited symptoms similar to attention-deficit/hyperactivity disorder (Wells et al., 2016).

In this inhibitory-focused approach, top-down attention functions more as a gate on feedforward processing rather than as a specific enhancer. This thesis addresses the functional and biological plausibility of such a mechanism for top-down attention, specifically building on the Validation Gate hypothesis that proposes that there is a threshold of acceptable errors in top-down prediction of feature spaces for inhibition, called anticipatory fields (Mathews et al., 2015). Revisiting the bottleneck of sensory processing, it is known that loading working memory leads to a reduction in sensory processing capacity (e.g. Mathews et al., 2015). In such situations, the Validation Gate hypothesis describes how top-down processes dynamically increase the inhibitory threshold of the anticipatory fields to filter out larger feature spaces. This reduces sensory perception to a level that is manageable for the available processing resources. In this thesis, we take a closer look at the described inhibitory mechanism of attention and how it interacts with working memory through the lens of the Validation Gate hypothesis.

The Validation Gate hypothesis

A similar processing and integration bottleneck can be found in information theory, which formulates the problem as finding the optimal trade-off between accuracy and complexity when describing a variable given its joint probability distribution with another, observed variable (Tishby et al., 2000). This was addressed in an autonomous and interactive system tracking multiple users in real time (Mathews et al., 2012), where multi-sensory input is associated to stimuli, or 'concepts', and subsequent input due to these stimuli is anticipated spatio-temporally. The anticipatory fields were projected as regions within the dimensions that describe the stimuli, thus providing a threshold for prediction errors that are within a range determined to be acceptable by the system. Each stimulus, in this case each user, had their own anticipatory field, also called a validation gate. The saliency of stimuli within the validation gates are inhibited, creating a selection bias towards unexpected input and away from noise. In this way, the validation gates defined regions in the input space that had lower potential for information gain, thus directing available processing resources to the spaces outside of them (Figure 2). When the available processing resources were further reduced, for example when different levels of the cognitive hierarchy, mainly between perception and task sets, compete for them, the threshold of the validation gates were dynamically increased in response.

The functional success of this probabilistic model led to the hypothesis that it could be a component of an inhibitory mechanism of selective attention, which was then directly tested, again in the spatial domain, in a psychophysiological study on humans using a displacement detection task combined with a secondary task to load participants' working memory (Mathews et al., 2015). This task ensures that task-relevant features are unpredictable while task-irrelevant ones are predictable. The authors demonstrated that low-latency saccadic responses and key presses could be used as proxies for bottom-up attention and volitional attention respectively, as increasing working memory load, also referred to as cognitive load, led to a decrease in performance of key presses but not low-latency saccades. Critically, it was found that the decrease in performance of key presses was due to a decrease in sensitivity of spatial

displacements, which is the task-relevant feature in this task, supporting the prediction made by the Validation Gate hypothesis.

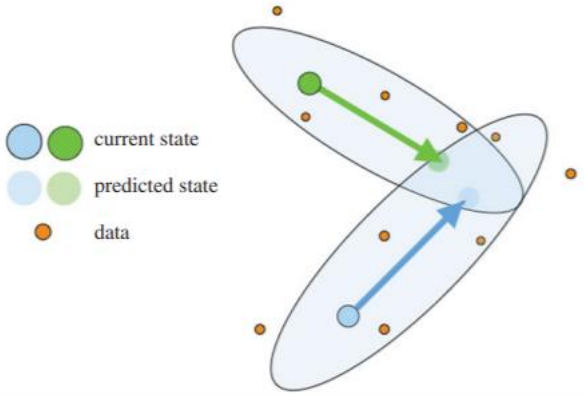


Figure 2. *Validation Gate hypothesis, visualised.* Reproduced with permission from (Verschure, 2016). If we follow dots moving along linear trajectories, data (orange dots) is classified relative to areas in input space where it is expected to occur given the properties of stimuli, or their validation gate (light blue area). Resources are only allocated to data which falls outside of the validation gates or when validation gates overlap, i.e. resolving novelty and ambiguity respectively.

Using the same task, another study replicated this effect of working memory load and additionally isolated brain regions that contribute to the process with functional magnetic resonance imaging (Malekshahi et al., 2016). These were the parietal and prefrontal regions, along with the precuneus, the caudate nucleus and the thalamus. Based on activity within these regions, the authors showed that prediction errors leading to key presses were processed mainly in late processing stages and that the activity of higher-order regions in the brain modulates the detection of these errors. The findings bolster the earlier proposal of a thalamocortical substrate of top-down driven inhibition of predictable distractors.

Considering the discussed findings, this thesis, therefore, investigates the hypothesis that there is a top-down driven inhibitory mechanism of selective attention. In **chapters 2 and 3**, we test the generalisability of the Validation Gate hypothesis to people with cognitive impairment and to a stimulus feature other than spatial location respectively. Using psychophysical tasks, we show that working memory load modulates task-specific error sensitivity in

both healthy participants as well as participants with cognitive impairment, confirming the proposal that it is an outcome of the integration of top-down and bottom-up attentional processes. Using the same task, we also demonstrate the functional importance of saccadic eye movements in the retrieval of items in working memory distinct from their more intuitive role in active sampling.

Against the backdrop of these results, in **chapter 4** we analyse a computational model of top-down inhibition for selective attention. The model is grounded in biophysically plausible interactions between specific brain regions that would allow for an inhibitory modulation of stimuli representation in the sensory cortex, as suggested by the Validation Gate hypothesis. It consists of populations of spiking neurons, allowing us to compare changes in spiking behaviour with what was found in mouse models of the brain, in addition to demonstrating the model's functional capabilities through the execution of two typical attentional tasks.

Attention models should lead to advantageous behaviours for the agent hosting it, implying that the proposed top-down inhibitory mechanism of selective attention should improve an agent's ability to behaviourally select target objects over distractors. Therefore, **chapter 5** describes how both top-down inhibitory and excitatory attentional biasing can be implemented in an artificial agent through the operationalisation of cognitive functions in the Distributed Adaptive Control cognitive architecture. The agent's behaviour selection is dependent on sensory processing that can be biased by top-down processes, and its performance in a simple foraging task when using different top-down attentional mechanisms is measured. We found that any mechanism of top-down attention improves behaviour selection compared to when behaviour is driven by purely bottom-up, competitive interactions. The described excitatory and inhibitory mechanisms of top-down attention also indicate a possible trade-off in terms of reaction time and accuracy, and the most robust form of top-down attention is the combination of both of excitatory and inhibitory biasing.

Together, this thesis demonstrates how, in certain situations, top-down attentional processes in humans select stimuli for inhibition to

bias the competition between stimuli in the sensory cortices. This selection is made using predictable features that are task-specific, and the threshold of inhibited features is modulated by sensory processing capacity. The neural substrate of this mechanism is shown to be possibly thalamocortical, and evidence is provided for its behavioural relevance and complementary role with the excitatory model of top-down attention.

The concluding chapter places the results in the broader context of our understanding of attention. It describes how a lot of research centre on search paradigms, even though, in most animals, that accounts for only a portion of the functional usefulness of selective attention. For example, task-relevant stimuli are not always more predictable than task-irrelevant ones, and attention is also closely coupled with other cognitive processes that have long-term implications, such as learning and memory. A more comprehensive model of attention would also lead to more effective therapies and interventions for people who suffer from attentional deficits.

2. TOP-DOWN ATTENTION IS MODULATED BY DEMANDS ON COGNITIVE RESOURCES

Selective attention plays such a pivotal role in our lives that dysfunctions in some part of its network of brain regions is associated with a multitude of disorders—for example, autism (Belmonte and Yurgelun-Todd, 2003), attention-deficit/hyperactivity, and depression (Kertzman et al., 2010)—as well as syndromes like neglect (Corbetta and Shulman, 2011). It closely interacts with working memory, and an individual's working memory capacity is predictive of higher-order cognition (e.g. Daneman and Carpenter, 1980) that closely relates to real-world activities (Engle, 2002). Selective attention enlists multiple brain regions, summarised as a) the dorsal frontal-parietal network that controls attention and eye movements and encodes stimulus saliency, and b) the ventral frontal-parietal network that underlies reorienting and detection of novel behaviourally relevant events (Corbetta and Shulman, 2011). As a result, stroke-induced lesions often disrupt the attentional network and post-stroke cognitive impairments, which are frequent (Leśniak et al., 2008) and can severely impact survivors' ability to live independently (Paolucci et al., 1996), include attentional deficits.

As predicted by the Validation Gate hypothesis, top-down attentional processes and working memory draw from overlapping pools of cognitive resources, if not the same pool. This leads to a need for top-down attention to inhibit larger feature spaces in sensory processing when working memory takes up more of these resources. As these inhibited regions indicate states of features predicted to be task-irrelevant, they are called anticipatory fields. While there is ample evidence that top-down attention is modulated by working memory load (Gazzaley and Nobre, 2012), how cognitive impairment, which implies a restriction on cognitive resources, affects top-down attention is less clear. However, clinical scales are notoriously poor at giving insight to the mechanistic deficiencies causing the impairments that they measure. Indeed, there are many clinical pen-and-paper scales used to diagnose cognitive impairment, and their diagnoses do not always converge. Correlating the clinical measures to performance in a task known to dissociate top-down and bottom-

up attentional processes would clarify the specificity of these clinical measures.

In this work, we use the attentional, displacement detection task to test the effect of cognitive impairment on both explicit and implicit responses. We hypothesise that cognitive impairment should affect the same top-down processes that are hindered by working memory load while leaving the bottom-up processes intact. This is measured by changes in proportion of explicit displacements detected and the latency of detections, which has been shown to depend on cognitive load, while both proportion and latency of implicit detections are unaffected. In addition, we expect that explicit detections would be modulated by cognitive impairment through changes in sensitivity in task-specific domains due to the top-down processes' anticipatory contribution, which, in this case, is spatial. Hence, cognitive impairment, possibly induced, for example, by a stroke, can be expected to have a similar effect on explicit responses as working memory load does in healthy participants. It is particularly pertinent to understand cognitive impairments caused by a stroke as not only does it happen frequently (Leśniak et al., 2008), but it is also known to lower quality of life (Park et al., 2013).

We fit the data to a generalised linear mixed-effects model to ascertain the contribution of stimuli attributes, working memory loads and cognitive deficits on explicit responses as well as their interaction effects. Also, we expect to find that working memory load modulates explicit responses in stroke patients similar to healthy participants. At the same time, their implicit saccadic responses should not change between the working memory load conditions, indicating the successful dissociation between top-down and bottom-up processes.

2.1 Methods

The data was obtained from patients participating in a previous randomised controlled pilot trial that investigated the use of a VR-based stroke rehabilitation system (Rehabilitation Gaming System, Eodyne Systems S.L., Barcelona) to improve cognitive abilities (Maier et al., 2020). Twenty-nine patients (14 female, age=64.69±6.78 years) were recruited from Hospital de l'Esperança, Barcelona, and the inclusion criteria were: a) cognitive impairment (Montreal Cognitive Assessment (MoCA) < 26) due to a first-ever stroke over six months ago, and b) aged between 45 and 75 years old. Patients with hemianopia, spasticity, severe cognitive, physical, or perceptual impairments that could interfere with the execution of the experiment were excluded. The study was approved by the local Ethical Committee and registered at ClinicalTrials.gov (NCT02816008). The scales and tests carried out that are related to cognitive impairment are described in greater detail in Participants were seated in a chair at approximately 75 cm in front of a 24" touchscreen All-in-One PC (Sony Vaio) with a refresh rate of 60 Hz (Figure 3-A). The experiment was implemented using Unity software (Unity Technologies, U.S.) and the in-game resolution was 150 × 266 pixels. A keyboard was placed in front of the participant to log their button presses and a Tobii T120 eye tracker (Tobii Technology AB, Sweden) was placed directly below the screen and recorded the participants' eye movements at 120 Hz. The participants listened to auditory cues through the PC's built-in speakers. Before the experiment started, participants were able to freely adjust the volume of the speakers.

Participants could also adjust the chair's height and back-tilting configuration for comfort and they were instructed to remain as still as possible throughout the experiment. They were oriented to face blank walls behind the screen, and the experiment was conducted in an empty room at the hospital. The experimenter remained in the room throughout the session.
Table 1 and the patients' characteristics and the clinical scales collected can be found in Table 2. The Hamilton Depression Rating Scale was only available for a subgroup of patients (HAM-D, n=19).

Participants were seated in a chair at approximately 75 cm in front of a 24" touchscreen All-in-One PC (Sony Vaio) with a refresh rate of 60 Hz (Figure 3-A). The experiment was implemented using Unity software (Unity Technologies, U.S.) and the in-game resolution was 150 × 266 pixels. A keyboard was placed in front of the participant to log their button presses and a Tobii T120 eye tracker (Tobii Technology AB, Sweden) was placed directly below the screen and recorded the participants' eye movements at 120 Hz. The participants listened to auditory cues through the PC's built-in speakers. Before the experiment started, participants were able to freely adjust the volume of the speakers.

Participants could also adjust the chair's height and back-tilting configuration for comfort and they were instructed to remain as still as possible throughout the experiment. They were oriented to face blank walls behind the screen, and the experiment was conducted in an empty room at the hospital. The experimenter remained in the room throughout the session.

Table 1. *Description of clinical scales and test batteries used to diagnose and measure cognitive impairment.*

Name	Abbreviation	Protocol summary
Montreal Cognitive Assessment (Nasreddine et al., 2005)	MoCA	Test battery assessing short-term/working memory, visuospatial abilities, sustained attention, language, abstract reasoning, and orientation to time and place.
Mini-Mental State Examination (Folstein et al., 1975)	MMSE	Test battery assessing short-term/working memory, attention, language, and orientation to time and place.
Corsi block-tapping test (Corsi, 1973)	Corsi F	Repeat an observed sequence of blocks "tapped" in the original order.
Backward Corsi (Isaacs and Vargha-Khadem, 1989)	Corsi B	Like the Corsi F, but the order of the observed sequence must be reversed.
Frontal Assessment Battery (Dubois et al., 2000)	FAB	Bedside test battery assessing conceptualization, mental flexibility, motor programming, sensitivity to interference, and inhibitory control.
Rey Auditory Verbal Learning Test (Hall et al., 2010)	RAVLT	Five presentations of a 15-word list are given, each followed by attempted recall. This is followed by a second 15-word interference list (list B), followed by recall of list A.
RAVLT, immediate	RAVLT I	The score of the RAVLT for only immediate recall.
RAVLT, delayed recall	RAVLT D	The score of the RAVLT for only delayed recall.
Star cancellation test (Wilson et al., 1987)	Star	The subject is shown a sheet of paper covered with two sizes of stars, alphabets and words and asked to mark out only one of those four objects e.g. only small stars.
Trail Making test (Tombaugh, 2004)	TMT A	The subject has to connect 25 numbers distributed on a sheet of paper in ascending order and as quickly as possible.
	TMT B	The subject has to connect numbers and letters in ascending order, alternating between the numbers and letters (i.e. 1-A-2-B-3-C…) as quickly as possible.
Wechsler Adult Intelligence Scale IV (Kaufman and Lichtenberger, 2002)	WAIS	An IQ test assessing verbal comprehension, working memory, perceptual organisation, and processing speed through 13 separate tests.
	WAIS F	The score from the WAIS forward digit span test.
	WAIS B	The score from the WAIS backward digit span test.
	WAIS C	The score from the WAIS digit symbol coding test.

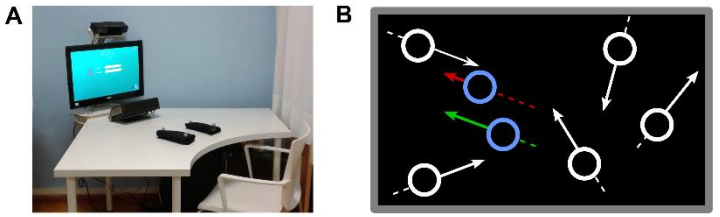
a) Experimental protocol

Before the experiment began, participants were briefed on the two tasks they were required to perform concurrently—a displacement detection task (Figure 3-B) and an auditory span task—and gave their written informed consent. They were familiarised with the hardware and the screen's brightness was adjusted to each participant's preference. They could adjust the keyboard and chair to a comfortable position before the eye tracker was calibrated using both the Tobii Pro Eye Tracker Manager (Tobii Technology AB, Sweden) as well as a 5-point routine built into the experiment. Previous work studying eye movements has used similar screen-based eye trackers with comparable accuracy and precision (e.g. Faber et al., 2018; Lloyd et al., 2017; Manning et al., 2014). The experiment was carried out in Spanish, as that was the preferred language of all the participants.

The participants carried out the displacement detection task in all blocks. Three levels of working memory load, or cognitive load (CL), were implemented using the auditory span task—without the auditory span task (low CL), with the auditory span task using short sentences (medium CL), and with the auditory span task using long sentences (high CL). In addition, there was an experimental block with no auditory span task and only one circle in the displacement detection task (no CL) which acted as control for subsequent analyses. The experiment consisted of 10 blocks, the first two of which were practice blocks. The order of the remaining eight blocks, the experimental blocks, was randomised in groups of four, each group consisting of one block per load condition (Figure 3-C). The duration of the blocks was dependent on the number of trials in the displacement detection task. There were five displacements in the practice blocks and each experimental block summed 30 displacements. The experiment lasted approximately 30 minutes and subjects received neither financial nor material remuneration.

Table 2. *Patients' characteristics.* BI, Barthel Index; Corsi B, Corsi Block Tapping Test Backward; Corsi F, Corsi Block Tapping Test Forward; FAB, Frontal Assessment Battery; FM-UE, Fugl-Meyer Upper Extremity Assessment; HAM-D, Hamilton Depression Rating Scale; MMSE, Mini-Mental State Examination; MoCA, Montreal Cognitive Assessment; MRC, Medical Research Council Scale for stroke assessment; RAVLT, Rey Auditory Verbal Learning Test; RAVLT I, RAVLT Immediate; RAVLT D, RAVLT Delayed Recall; Star, Star Cancellation Test; TMT A, Trail Making Test A; TMT B, Trail Making Test B; WAIS, Wechsler Adult Intelligence Scale IV; WAIS F, WAIS Digit Span Forward; WAIS B, WAIS Backward; WAIS C, WAIS Digit Symbol Coding, SD, standard deviation.

Characteristics	**(n = 29)**
	n (%)
Gender, female	14 (48.28%)
Impaired limb, right	13 (44.83%)
Aetiology	
Ischemic	17 (58.62%)
Hemorrhagic	11 (37.93%)
Capsulo lenticular	1 (3.45%)
	Mean (SD) – Median [2.5th and 97.5th percentile]
Age, years	64.69 (6.78) – 63.00 [53.68 – 76.00]
Days after stroke	1200.31 (1249.34) – 911.00 [190.45 – 5287.05]
Clinical Scales	
MoCA	21.00 (3.41) – 21.00 [12.45 – 25.00]
MMSE	27.24 (1.94) – 28.00 [23.23 – 30.00]
MRC	3.69 (0.66) – 4.00 [2.23 – 5.00]
FM-UE	52.69 (16.67) – 60.00 [7.25 – 66.00]
BI	94.31 (8.32) – 100.00 [76.13 – 100.00]
HAM-D (n = 19)	5.58 (5.14) – 4.00 [0.00 – 15.00]
Neuropsychological test battery	
Corsi F	5.59 (1.59) – 6.00 [3.00 – 9.00]
Corsi B	4.52 (1.88) – 5.00 [1.23 – 8.55]
FAB	16.31 (1.87) – 17.00 [11.23 – 18.00]
RAVLT I	33.10 (9.60) – 34 [11.13 – 46.55]
RAVLT D	5.07 (2.75) – 5.00 [0.00 – 10.00]
Star	51.38 (5.70) – 53.00 [29.05 – 54.00]
TMT A	71.79 (40.80) – 66.00 [29.00 – 171.05]
TMT B	214.86 (130.82) – 173.00 [48.50 – 402.00]
WAIS F	5.14 (1.19) – 5.00 [3.00 – 7.00]
WAIS B	3.45 (0.95) – 3.00 [2.00 – 5.00]
WAIS C	28.66 (13.58) – 27.00 [10.23 – 58.55]



C

Training		Evaluation (2x)			
		No CL	Low CL	Med CL	High CL
Fixed		Fixed	Randomized		
5 trials	5 trials	30 trials	30 trials	30 trials	30 trials
4 circles	4 circles LS	1 circle	6 circles	6 circles SS	6 circles LS

Figure 3. *The set-up, the schema of task presentation and experimental protocol.* A) The clinical set-up consisted of a table, an All-in-One PC and an eye tracker. B) The displacement detection task consisted of white circles that move in a linear trajectory over a black background (RGB: 43, 46, 49). At the edge of the screen, the circles 'bounced' off and continued moving in the direction of the original vector reflected about the axis parallel to the screen's edge. Occasionally one of the circles (indicated here in blue) displaced to a randomly sampled location within a radius of 3–40 pixels and continued moving (red arrow) in its original velocity (green arrow). C) The experimental protocol consisted of a training phase and an evaluation phase. The evaluation phase consisted of four different levels of cognitive load which were repeated twice, leading to eight experimental blocks of 30 trials each. CL, cognitive load; LS, long AST sentences; SS, short AST sentences; Med, medium.

Participants read the on-screen instructions for the experiment at the start, reiterating the verbal instructions, and further written reminders and block-specific instructions were shown on-screen between blocks. Before each block, interaction with the experiment was suspended for 1.5 s after which a button press would start the block.

b) Displacement detection task

The patients observed off-white circles (external diameter 3.2 cm (2.4 °), white border thickness 0.8 mm (0.61 °)) moving linearly on a dark background (Figure 3-B). Each circle had its own trajectory, randomly initialised at the start of every block, and when it reached the side of the screen its trajectory was reflected. Every 2.5–4 s, one of the circles displaced to a random position within a radius of 3–40 pixels from the previous location. After the displacement, it resumed its original trajectory. The patients were instructed to report every

displacement that they detected by pressing the spacebar on the keyboard.

Displacements did not occur if the participant's gaze could not be tracked at that instant, if it was too close to the screen's sides such that the displacement would move it out of the screen's dimensions, or if the circle was located too far from the gaze's position (>22 °). The circles moved at a speed of 0.19 °/s.

c) Auditory span task

The Auditory Span Task (AST) (Conway et al., 2005) was used to induce load on participants' working memory with two levels of difficulty based on the length of the sentence, long or short. The task required participants to listen to a sentence and respond if it was semantically correct immediately after. In the sentence database (Loboda, 2012), all the sentences consisted of an actor performing an action in a context. For semantically incorrect sentences, one of those three parts leads to a violation of expectations based on the other two parts (e.g. Mr Jones asked his son to water the cats and mow the lawn. Actor: Mr Jones' son; action: to water; context: the cats). However, the sentences often had additional information that did not interfere with this violation (mowing the lawn, in the previous example). The long sentences were thus the full ones, while the short sentences stayed the same but with the additional information removed.

Including the variables from the displacement detection task, there were the following experimental conditions: 1) presented one circle only (no CL), 2) presented six circles (low CL), 3) presented six circles while listening to short sentences (medium CL), and 4) presented six circles while listening to long sentences (high CL). Each condition was repeated twice during the experiment pseudo-randomly, totalling eight experimental blocks (Figure 3-C). The two practice blocks presented participants with four circles in the displacement detection task at the start; one included the auditory span task (short sentences, medium CL) while the other did not.

d) Outcome measures

Before the experiment, all patients included in this study were evaluated by a neuropsychologist with a neuropsychological test battery as well as secondary clinical scales that cover various cognitive abilities, motor functioning, and depression (Table 2).

For each cognitive load condition, we extracted displacements that are reported by either key presses (i.e. explicit detections) or saccadic responses (i.e. implicit detections), as well as the latencies of these reporting methods. An explicit detection was a key press that occurred 200 ms—3 s after a displacement. To identify implicit, saccadic detection of displacements, we followed the procedures and applied the scripts proposed by Wass et al. (2014, 2013) as they have been demonstrated to be robust to noisy eye tracking data. The raw eye-tracking data was first cleaned and missing values for both eyes were interpolated where possible. Specifically, we identified missing values at every timestep, marked outliers using the inbuilt MATLAB-function 'isoutlier' and merged the valid left and right eye data by taking their average position. The resulting data was smoothed using a simple down-sampling procedure: the data points were chunked into consecutive window-sized segments using a 100 ms window size, and a single median average was calculated per window. For windows in which fewer than 50% of samples were available, the entire window was returned as empty and excluded from further analysis (Wass et al., 2014). In addition, when data outside of the valid screen range was detected, we corrected for the shift by subtracting the maximum from all data points, before converting the dataset into pixels.

Next, a preliminary classification of eye movement into saccades and fixations was performed based on angular velocity: valid eye data was flagged as a saccade when their angular velocity exceeded 30 °/s at each time step. Then, small gaps of missing positional data (up to 75 ms or 9 samples long) were linearly interpolated. These small gaps were also labelled as saccades if their angular velocity exceeded the threshold. Lastly, we excluded saccades that were flagged right before or right after the remaining gaps of data, as they were most likely data fragments (interrupted saccades). We also cleaned out saccades that were either shorter than 75 ms in total duration or

covered less than 0.5 ° in angular distance. All thresholds used in the pre-processing and labelling pipeline were based on reported values from the original equipment manufacturer (Olsen, 2012).

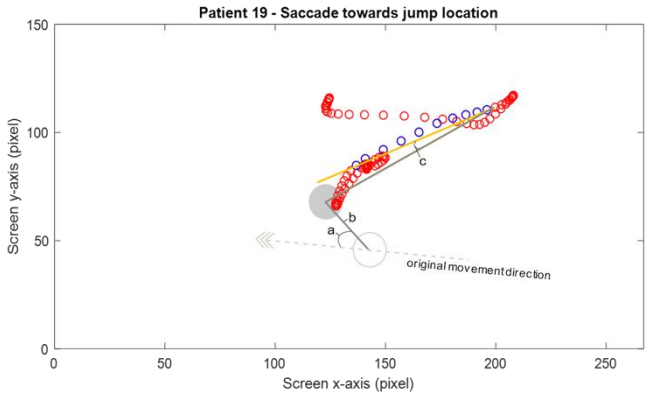


Figure 4. *Schemata of a circle displacement and corresponding eye movement.* The dotted line represents the eye movement before and right after a displacement. Red indicates positions labelled as fixations, and blue dots were labelled as saccades. To define whether a saccadic movement was towards a displacement position (filled grey circle), the angle between the ideal vector and the saccade vector, c, was calculated and needed to be less than 25 degrees. For each displacement, we also extracted the distance, b, and the angle, a, relative to the circle's movement direction before the displacement took place.

After cleaning and labelling the data, we proceeded to identify saccades that occurred 100–800 ms after a displacement. We extracted its magnitude (rho) and the angle it made relative to the ideal vector from the last fixation point to the displaced circle's current position (theta). The final criteria for an implicit detection was a theta within ±25 °. An example of the described analysis can be seen in Figure 4, which illustrates the eye movements of a random participant during a randomly selected displacement.

e) Behavioural analysis

The proportion of displacements that were detected by explicit and implicit responses were averaged across all participants per load condition. The resulting mean performance for explicit and implicit detections were first compared across loads using the non-parametric Friedman's ANOVA test (statistic χ^2_F). In case of significance, a post hoc analysis was performed using estimation statistics based on Monte Carlo permutation (Ho et al., 2019), with the paired mean

difference figures including a bootstrap 95% confidence interval to illustrate the effect size. The confidence interval accounts for skewness in the distributions with a bias-corrected and accelerated bootstrap (Efron, 1987). Each test was based on 5000 resamples; we report the results in terms of confidence intervals, mean difference values and the p-value of the permutation test. The explicit and implicit responses were correlated with the neuropsychological test battery and the secondary clinical scales using the Spearman's correlation (coefficient r_s). In the case of significant zero-order correlations, a partial-correlation was performed. Also, we relied on non-parametric tests because most data were non-parametric. All data processing and analysis were carried out using MATLAB 2017b (The MathWorks Inc, Natick, Massachusetts).

f) Visualisation of the anticipatory field

To determine how cognitive load affects sensitivity to displacement magnitudes, we divided the radial location of all displacements into a 12 by 5 grid and binned the displacement data of all participants in each dimension (angle and distance) per load condition. We did so using percentiles to account for a higher concentration of smaller displacements, which was an artefact of experimental limitations. We then interpolated each bin's detection rate to obtain a radial distribution of probabilities per load condition.

g) Generalised linear mixed-effects model

We estimate the contribution of the displacement dimensions on detection rate using a regression model, following recommendations proposed for psychophysical tasks (Moscatelli et al., 2012). Based on previous work, we expect to find that cognitive load has a greater modulatory effect on smaller displacements (rho) than larger ones, giving rise to the previously described anticipatory fields. A regression model also allows us to explore whether the clinical outcomes can serve as additional predictors of detection rate across the different cognitive load conditions. To be considered as predictors in the regressor, the clinical scale had to have shown a strong correlation with the detection ability beforehand. As the data

was repeatedly sampled across subjects, the observations of the response probability were dependent and may vary consistently between patients. To deal with this pseudoreplication and to fit the response outcome (i.e., number of detections) per displacement magnitude bin (rho) adequately, we applied a generalised linear mixed-effects model for binomial distributions, using the 'probit link' function, Φ, the Laplace approximation and the subjects as random effects on intercept and the displacement magnitudes. The model is defined by Eq. 1 as follows:

$$
\begin{aligned}
&DDetected_{ij} \sim Binomial(\prod_{ij}, n_{ij}) \\
&\Phi^{-1}\left[P\left(DDetected_{ij} = 1\right)\right] = \eta_{ij} \\
&\eta_{ij} = \beta_0 + Participant_i^0 + Rho_{ij}(\beta_1 + Participant_i^1) \\
&\qquad\qquad + Scale_{ij}\beta_2 + LowCL_{ij}\beta_3 + HighCL_{ij}\beta_4 \\
&Participant_i \sim N(0, \sigma^2)
\end{aligned}
\qquad \textit{Eq. 1}
$$

Where $DDetected_{ij}$ is the response variable, described by its mean Π and size n, and includes the number of detected displacements for subject i (where i = 1, 2, …, 29), and displacement magnitude index j (where j = 1, 2, …, 10). As the response type is binary, the probability of it being 1 is equivalent to the response variable that is to be modelled. The probit link function therefore links individual participants, η_{ij}, to the probability of a detection. Rho_{ij} contains the displacement magnitude values (binned in percentiles and centred, with upper limits = -0.35, -0.31, …, 0.65), $Scale_{ij}$ represents the subject's score in the respective clinical scale included, $LowCL_{ij}$ and $HighCL_{ij}$ are the dummy variables for the cognitive load condition (using the no cognitive load condition as the reference group), $\beta_{0, \ldots,} \beta_4$ are the fixed-effects coefficients, and $Participant_i^0$, $Participant_i^1$ are the random-effect coefficients accounting for participant-specific variation in ability to respond.

We used the Akaike Information Criterion (AIC) (Hastie et al., 2009) to determine whether the inclusion of a random-effect parameter is justified and compared the different models using the Likelihood-Ratio test to determine if the addition of a fixed effect is statistically relevant (nested models).

2.2 Results

a) Behavioural analysis

First, we analysed the explicit (key presses) and implicit (saccades) detections across all patients. Although there was a significant main effect across all load conditions, we found no difference between medium CL and high CL. We therefore combined the data of these two load conditions (henceforth referred to as high CL), since both required the patient to listen to the AST as a secondary task.

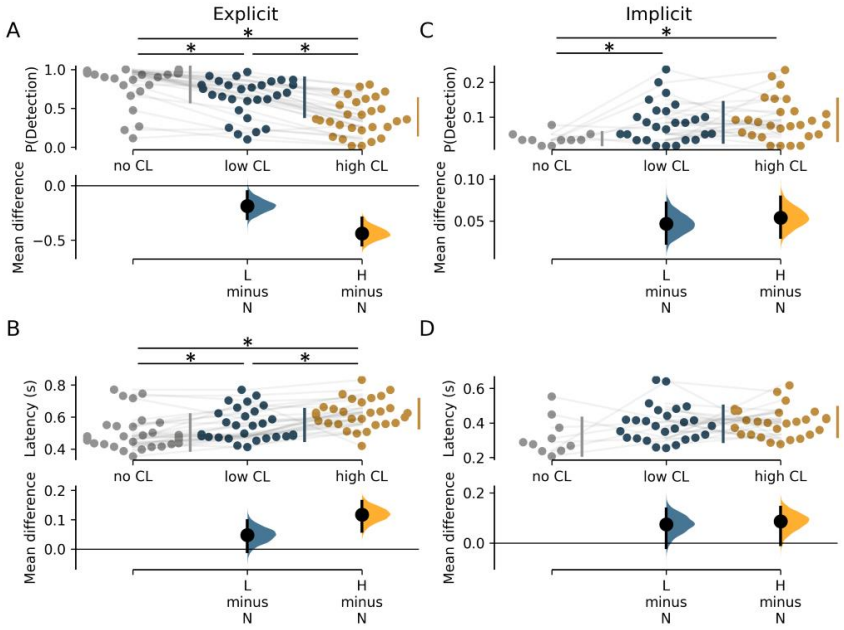


Figure 5. *Proportions of displacements detected.* A) The proportion of all displacements that were reported by key presses and thus are considered explicit detections. B) The latency of explicit detections. C) The proportion of all displacements that were detected by saccades towards the displacement position. D) The latency of the implicit detections. Mean differences with respect to the no load condition are shown in the contrast plots below the swarm plots, with the confidence intervals indicating the distributions of the means of the resamples. Individual dots within the swarms indicate individual participants. CL: cognitive load.

Consequently, cognitive load significantly modulated the explicit detection probability (Friedman; $\chi^2_F(2)=23.82$, $p<0.001$, Figure 5-A). Their reaction time also increased with cognitive load (Friedman; $\chi^2_F(2)=35.38$, $p<0.001$, Figure 5-B). The proportion of implicit

detections was, in general, higher when there was a cognitive load (Friedman; $\chi^2_F(2)=24.73$, $p<0.001$, Figure 5-C), but there was no significant difference between low and high load conditions ($p=0.37$). There were also no significant differences in implicit detection latencies across all load conditions (Friedman; $\chi^2_F(2)=5.25$, $p = 0.07$, Figure 5-D). More details on the statistical test results are in Table 3. The no cognitive load condition had one circle moving on screen, which led to mostly smooth pursuit eye movements instead of saccades. Also, there were low proportions of implicit detections in the low and high cognitive load conditions as there were not many saccades made in general, which is a finding consistent with literature. This meant that explicit and implicit measures-were-not suitable-for direct comparison (e.g. explicit detection in no CL vs implicit detection in no CL).

Table 3. *Results of post hoc Monte Carlo permutation.* N: sample size; CI (95%): 95% confidence interval; CL: cognitive load.

Comparison	N	Mean difference	CI (95%)	p-value
Proportion of explicit responses				
no CL vs low CL	29	-0.19	-0.23 and -0.14	<0.001
no CL vs high CL	29	-0.44	-0.51 and -0.35	<0.001
low CL vs high CL	29	-0.25	-0.30 and -0.20	<0.001
Latency of explicit responses				
no CL vs low CL	29	0.047	0.022 and 0.073	0.0012
no CL vs high CL	29	0.12	0.061 and 0.15	<0.001
low CL vs high CL	29	0.070	0.030 and 0.096	<0.001
Proportion of implicit responses				
no CL vs low CL	29	0.063	0.042 and 0.085	<0.001
no CL vs high CL	29	0.072	0.051 and 0.095	<0.001
low CL vs high CL	29	0.0099	-0.010 and 0.033	0.37
Latency of implicit responses				
no CL vs low CL	8	0.065	-0.030 and 0.19	0.33
no CL vs high CL	10	0.081	-0.034 and 0.16	0.15
low CL vs high CL	25	0.010	-0.046 and 0.062	0.71

b) Clinical measures are correlated with explicit but not implicit responses

There were significant correlations between explicit detections and neuropsychological test battery outcomes, as well as the HAM-D. However, no such correlations were found with implicit detections and no other secondary outcomes were consistently correlated with explicit detections. The tests within the neuropsychological test battery that correlated significantly with explicit detection proportion across all load conditions were Corsi F, Corsi B, TMT A, TMT B and WAIS C (Figure 6-A). Corsi F and WAIS C also showed a consistent positive correlation with the latency in detecting displacements (Figure 6-B). Besides Corsi F (Spearman's; r_s=-0.61, p<0.001), only TMT A and TMT B correlated positively with HAM-D (Spearman's; TMT A: r_s=0.46, p<0.05, TMT B: r_s=0.47, p<0.05). HAM-D was inversely correlated with the ability to report stimuli for low and high CL (Spearman's; low CL: r_s=-0.47, p<0.05, high CL: r_s=-0.47, p<0.05).

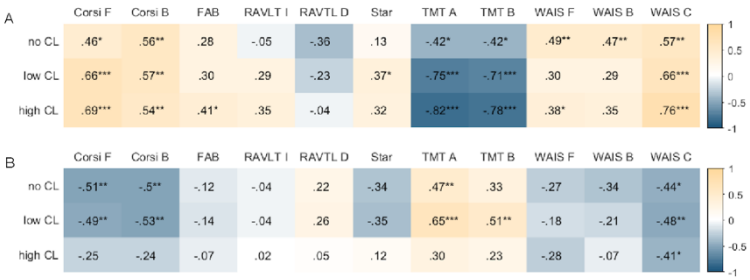


Figure 6. *Correlations between explicit detection proportion, latency and neuropsychological test battery.* A) Correlations with proportion of explicit detections (key presses). B) Correlations with the latency of explicit detections. Significant p-values are indicated * p<0.05, ** p<0.01, *** p<0.001 and the number and colour represent the correlation coefficient (Spearman's *r*). Corsi B: Corsi Block Tapping Test Backward; Corsi F: Corsi Block Tapping Test Forward; FAB: Frontal Assessment Battery; RAVLT: Rey Auditory Verbal Learning Test; RAVLT I: RAVLT Immediate; RAVLT D: RAVLT Delayed Recall; Star: Star Cancellation Test; TMT A: Trail Making Test A; TMT B: Trail Making Test B; WAIS: Wechsler Adult Intelligence Scale IV; WAIS F: WAIS Digit Span Forward; WAIS B: WAIS Backward; WAIS C: WAIS Digit Symbol Coding.

To further evaluate the relationships between the TMT tests, HAM-D, and explicit detection proportion in low and high load conditions, we performed a partial correlation while controlling for the effect of either TMT A, TMT B or HAM-D as independent variables. For this

analysis, we used the subgroup for which the HAM-D is available (n=19). However, as this subgroup did not exhibit a significant correlation between detection rate and Corsi-F, the partial correlation was carried out only for TMT A and TMT B but not for Corsi-F. Also, we only considered low and high load detection rate as HAM-D did not correlate with the no CL detection rate; we found HAM-D was no longer significantly correlated with explicit detection proportion after controlling for the effect of TMT A and TMT B (Figure 7). On the other hand, explicit detection proportions remained significantly correlated to TMT A and TMT B even after controlling for the effect of HAM-D for both low and high cognitive load conditions.

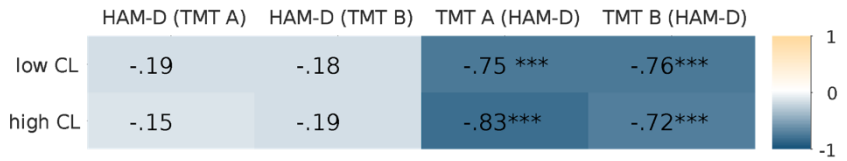


Figure 7. *Partial correlation across cognitive loads for explicit detection proportion.* The controlled variable is given in brackets. Significant p-values are indicated as * p<0.05, ** p<0.01, *** p<0.001 and the number and colour represent the coefficient (Spearman's r). CL: cognitive load; HAM-D: Hamilton Depression Rating Scale; TMT A: Trail Making Test A; TMT B: Trail Making Test B.

c) Generalised linear mixed-effects model

To investigate if the patients had a spatial bias when responding to displacement events, we visualised their explicit response probability in the Cartesian space of the displacement events across the different cognitive load conditions (Figure 8, more details available in section 2.1-f), with white representing jumps that are always explicitly reported and black those that are never explicitly reported. We split the participants according to whether they demonstrated cognitive impairment as measured by TMT A and/or TMT B. The impairment cut-off was based on normative data obtained from a healthy group, which corresponds to a response time less than 64 seconds in TMT A and less than 157 seconds in TMT B (Peña-Casanova et al., 2009). Participants labelled as 'no impairments' had response times below the normative cut-offs for both TMT A and TMT B.

Preliminary visual inspection of the anticipatory fields warranted closer, quantitative inspection of the data. We thus modelled the data using a generalised linear mixed-effects model to quantify the effects of each factor. The models tested can be found in Table 4 and the comparisons performed can be found in Table 5. The most basic model, GLME 0, described the relationships previously shown to be significant (Mathews et al., 2015), with cognitive load and displacement magnitude as the main effects (represented by the intercept and slope of Eq. 1 respectively).

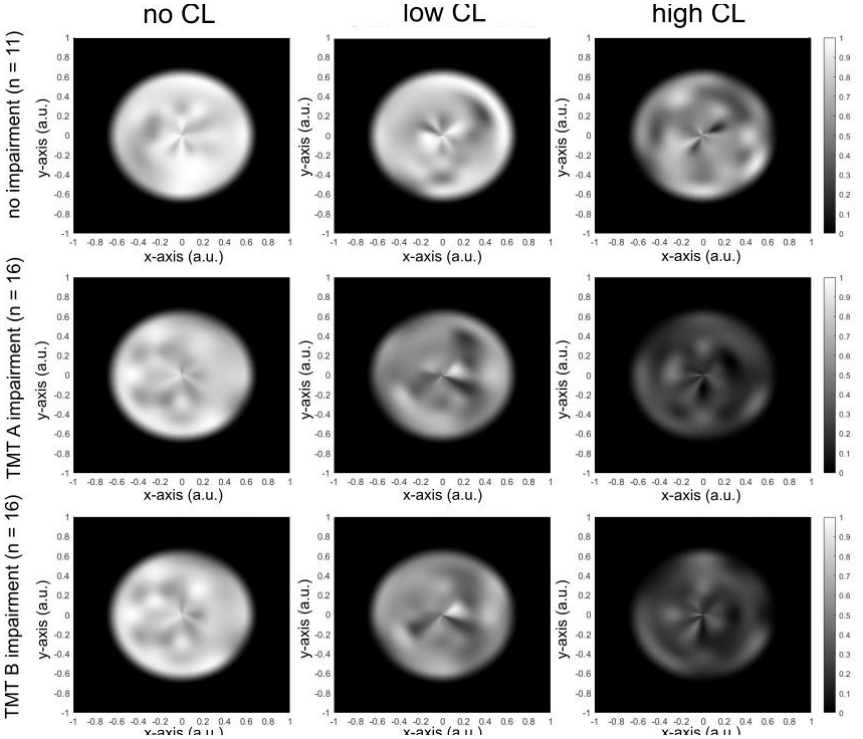


Figure 8. *Probability of explicit detection in each cognitive load condition for the three groups of cognitive impairment.* The probability was calculated for the binned data on the displacement's relative angle and magnitude.

We first verified that the inclusion of a random parameter would improve performance. The likelihood ratio test (comparison 1 and 2, Table 5), as well as an observed reduction in AIC, confirmed that the random effect parameter accounting for individual variability significantly improved the model's performance. Next, we tested if the inclusion of TMT A or TMT B as fixed effects would further improve the model, again quantified by their AIC. When tested

separately, both TMT A and TMT B reduced AIC (AIC of GLME 3a and GLME 3b respectively, Table 5), showing that their isolated addition was beneficial to the model with TMT B having a slightly greater effect. As a consequence, adding both TMT A and TMT B to the model did not further improve its performance, compared to including TMT B alone (comparison 3 and 4, Table 5). This demonstrates how explicit response probability was directly modulated by displacement magnitude, TMT B, and cognitive load, while taking into account individual differences.

Table 4. *Overview of generalised mixed-effects models (GLME) used in comparisons.* The parameters are described in Eq. 1.

Model	Formula
GLME 0	P(response) ~ 1 + Rhos + Loads
GLME 1	P(response) ~ 1 + Rhos + Loads + (1 | Participants)
GLME 2	P(response) ~ 1 + Rhos + Loads + (1 + Rhos | Participants)
GLME 3a	P(response) ~ 1 + Rhos + TMTA + Loads + (1 + Rhos | Participants)
GLME 3b	P(response) ~ 1 + Rhos + TMTB + Loads + (1 + Rhos | Participants)
GLME 3c	P(response) ~ 1 + Rhos + TMTA + TMTB + Loads + (1 + Rhos | Participants)
GLME 4a	P(response) ~ 1 + Rhos * Loads + (1 + Rhos | Participants)
GLME 4b	P(response) ~ 1 + Rhos + TMTB * Loads + (1 + Rhos | Participants)
GLME 5	P(response) ~ 1 + Rhos + TMTB * Loads + (1 + Rhos | Depression)

Lastly, we tested whether some of the fixed effects interacted with each other. Based on how they modulated AIC and/or likelihood, we found that displacement magnitude did not interact with cognitive load (comparison 5, Table 5), but TMT B did (comparison 6, Table 5). This showed that the impairment level in TMT B reduced individual explicit response probability more in the high load than in the low load condition. The final model estimates are shown in Table 6.

The model was evaluated using its residuals and the generated fitted values, as a better fit would lead to the distribution of the residuals being closer to uniform. The Pearson's residuals of the fit are slightly shifted, possibly pointing to heteroscedasticity; comparing them to the fitted values of the model confirms that there were signs of nonconstant variance amongst residuals, with most of the bias being

at extreme performance values. This implies that the model's ability to predict performance follows an inverse U-shaped curve depending on the participant's actual performance.

Table 5. *Generalised mixed-effects model (GLME) comparisons executed.* DF: degrees-of-freedom; AIC: Akaike Information Criterion; BIC: Bayesian Information riterion; LogLik: log-likelihood; LRstat: likelihood-ratio test statistic.

	DF	AIC	BIC	LogLik	LRstat	p
Comparison 1						
GLME 0 vs	4	3283.03	3302.07	-1637.51		
GLME 1	5	2169.95	2193.77	-1079.98	1115.08	<0.001
Comparison 2						
GLME 1 vs	5	2169.95	2193.77	-1079.98		
GLME 2	7	2165.88	2199.22	-1075.94	8.07	<0.05
Comparison 3						
GLME 3a vs	8	2153.30	2191.41	-1068.65		
GLME 3c	9	2153.21	2196.08	-1067.61	2.09	>0.05
Comparison 4						
GLME 3b vs	8	2152.53	2190.63	-1068.27		
GLME 3c	9	2153.21	2196.08	-1067.61	1.32	>0.05
Comparison 5						
GLME 2 vs	7	2165.88	2199.22	-1075.94		
GLME 4a	9	2166.30	2209.16	-1074.15	3.58	>0.05
Comparison 6						
GLME 3b vs	8	2152.53	2190.63	-1068.27		
GLME 4b	10	2141.81	2189.44	-1060.91	14.72	<0.001

Table 6. *Final estimates of the model.*

Name	Estimate	SE	t-Statistic	DF	p	Lower	Upper
Intercept	1.21	0.11	10.73	858	<0.001	0.99	1.43
Rhos	0.21	0.11	2.02	858	<0.05	0.01	0.42
TMT B	-0.003	0.001	-3.60	858	<0.001	-0.005	-0.001
Low CL	-0.80	0.06	-14.18	858	<0.001	-0.91	-0.96
High CL	-1.66	0.06	-28.48	858	<0.001	-1.78	-1.55
TMT B : low CL	-0.001	0.00	-1.21	858	>0.05	-0.001	0.000
TMT B : high CL	-0.002	0.00	-3.65	858	<0.001	-0.003	-0.001

To find the source of the unexplained variance that might aid to improve the model, we checked the contribution of Corsi F, Corsi B and WAIS C. The contribution of WAIS C appeared to be comparable to TMT B, but it did not reduce the heteroscedasticity. Similarly, adding the angle of the displacement (theta) to the model

did not improve its predictive power. Analysing the model only for the patients with HAM-D data and adding the depression scale as a random effect (i.e. treating it as a systematic trait, model GLME 5 in Table 4), rendered the residuals for poorer performance more constant, possibly explaining part of the variance observed in low performers and reiterating the nuanced relationship between depression, cognitive impairment, and performance in the displacement detection task.

2.3 Conclusions

We found that participants' bottom-up sensory processing (measured by implicit detections) in an attentional task was unaffected by cognitive load or levels of depression, while their top-down sensory processing (measured by explicit detections) was. The effect on explicit detections was negative for increasing levels of depression and poorer performance in TMT A and TMT B. Both TMT tests are timed and require online visual tracking ability; TMT A is considered to measure visuo-perceptual speed whereas TMT B is considered to be a measure for attentional switching and working memory (Sánchez-Cubillo et al., 2009). They both require psychomotor coordination and higher-order functions that have been shown to be affected by depression, which correlates not only with poor psychomotor speed but nonverbal problem solving, verbal and visual memory, and attention as well (Kauhanen et al., 1999). Our analyses revealed a correlation between performance in the TMT tests and depression, confirming these links between depression and cognitive performance.

The partial-correlation analyses aid in understanding the relationship between explicit responses (key presses), the participant's level of depression (HAM-D) and specific cognitive disability (TMT A and TMT B). Although all three variables correlated significantly, it appears that the relationship between depression and explicit responses can be explained mainly by a deficit in visuo-perceptual speed and working memory ability. Also, even though correlations do not explain directionality, it is unlikely that a deficit in visuo-perceptual speed and working memory would cause depression. Rather, if there is a direct causation effect between the two, the

opposite is more likely where impairments in visuo-perceptual speed and working memory result from depression. The correlation between depression and poorer perceptual speed is known (Blackburn, 1975; Fitzgerald, 2013), but disentangling their relationship has proven challenging. For instance, another study has also found that performance in the Stroop task, although affected by depression, is predicted by psychomotor speed rather than impairment of selective attention per se (Kertzman et al., 2010). Exploring the directionality of this relationship is a potential avenue for subsequent research.

The generalised linear mixed-effects model further sheds light on the exact contribution of the factors that modulate explicit response probability. Confirming previous work with healthy participants, a model that includes the displacement magnitude while accounting for inter-subject variability was shown to reasonably describe the explicit response probability, which decreases with increased cognitive load. In addition, we demonstrate here that a deficit in attentional switching and working memory (as measured by the TMT B) not only further exacerbates poorer performance but also interacts with cognitive load. The model shows that a specific cognitive impairment can have a direct additional effect on performance. These findings give predictive power to individuals' performance in the task, allowing it to be used as a potential diagnostic tool that is objective and independent of factors unrelated to cognitive ability, such as educational background.

However, the model appears to poorly predict performance at the extreme ends (extremely good or extremely poor) in the task. While we were able to improve the model's predictive performance of poor performers by including a measure for depression, more work needs to be done to find the source for the unexplained variance in high performing individuals.

Through the displacement detection task, we distinguished between the responses that are primarily driven by bottom-up input and those that result from integration of top-down processes. That the participants' implicit responses, proxies for bottom-up attention, are unaffected by cognitive load and further do not correlate with any psychoneurological tests suggests that their bottom-up attentional mechanism remained intact. Nonetheless, it was found that

impairment severity of a specific cognitive ability, in this case measured by TMT B, led to a decrease in explicit responses, especially when there was a cognitive load (Figure 8 and Table 3). In contrast, a more general measure of cognitive deficit like MoCA was uncorrelated with explicit responses, despite all participants included in this study being considered to be cognitively impaired according to their MoCA score. This suggests that the TMT B specifically measures cognitive functioning that overlaps with top-down attentional processes while MoCA does not. Standard clinical tests typically are unable to give insight as to the precise neural underpinnings of the impairments that they aim to measure. By using psychomotor tasks such as the displacement detection task, we can infer that poor performance in TMT B could be caused by a lesion that specifically disrupts frontoparietal networks known to contribute to top-down attentional processes. This notion resonates well with the view that impairments from stroke are symptoms of disruptions in distributed neuronal networks, which otherwise sustain sensorimotor and cognitive functioning in healthy individuals (Guggisberg et al., 2019).

3. SACCADIC DYNAMICS ARE MODULATED BY WORKING MEMORY LOAD

As elaborated in the earlier chapters and sections, findings from psychophysical studies have supported both the proposal that top-down attentional modulation can be inhibitive in certain situations and that top-down attentional processes dynamically modulate sensitivity to task-specific stimulus features depending on working memory load. However, the studies directly addressing the Validation Gate hypothesis have kept to a spatial domain. If the hypothesis is generalisable to the mechanism of selective attention, this effect should extend to other domains and modalities as well.

Within the visual modality, information is processed hierarchically (Liu et al., 2020; Maunsell and Newsome, 1987; Treisman, 1986) and splits into the dorsal and ventral streams (Schneider, 1969; Ungerleider and Mishkin, 1982). Also called the 'where' and 'what' streams respectively, the dorsal stream involves sensory processing more related with localisation of stimuli while the ventral stream is more associated with the identification of stimuli features. Although the two streams are unlikely to be functionally independent and their specialisation is relative rather than absolute (Schenk and McIntosh, 2010), the task-features of interest in the displacement detection task used to test the Validation Gate hypothesis are not particularly complex and should therefore have not involved significant collaboration between the two visual streams.

Adapting the displacement detection task paradigm from the spatial domain in the dorsal stream to one considered to be part of the ventral stream would allow for a comparison to determine the Validation Gate hypothesis' generalisability while remaining in the visual modality. For this reason, we designed a task that retained the key characteristics of the displacement detection task while translating it to the colour domain. Our first hypothesis in this study is that in this task, top-down attentional processes will again be modulated by working memory load except in this case the decrease in sensitivity will be towards colour instead of space.

In addition, saccadic eye movements have been demonstrated to modulate the phase and amplitude of theta oscillations which are

elicited during working memory tasks. For example, during visual exploration, theta's phase was predictive of the following fixation-location's novelty (Kragel et al., 2020). Reiterating earlier visual exploration studies (e.g. Ito et al., 2011), the phase reset was marked by saccade onset rather than fixation. Furthermore, the relationship between eye movements and neural oscillations for memory was also quantified in primates during visual exploration, where the reliability of the elicited phase reset was predictive of subsequent accuracy (Jutras et al., 2013). Likewise, the observed phase resets were shown to be distinct from evoked response potentials due to stimuli onset (Katz et al., 2020), supporting the hypothesis that oscillatory phases contribute to mnemonic processing.

At the same time, it has been proposed that multiple-item representation in working memory is expressed by gamma cycles within theta oscillations, called the theta-gamma code (Jensen and Lisman, 1998). The theoretical and computational implications of such a cross-frequency code favour a framework in which the brain optimises the number of items held in memory (Lisman, 2010; Lisman and Jensen, 2013).

Indeed, recent experimental work has shown that increases in the number of working memory items are accompanied by decreases in peak hippocampal theta frequency, and that the robustness of this shift predicts individuals' recall accuracy (Axmacher et al., 2010). Additionally, the theta's phase was shown to be modulated during working memory recall periods (Rizzuto et al., 2003). Specifically, neocortical theta displayed phase alignment to the task's probe event, suggesting that coupled oscillations not only store information about items held in memory but also that accessing this information requires a phase-tuned carrying frequency.

Consequently, these studies show that eye movements modulate the phase and amplitude of theta oscillations, and that theta oscillations enable the encoding and maintenance of items during working memory. Together, they raise the question of whether eye-movements are functionally relevant for working memory.

The second hypothesis of this study is thus that increases in working memory load (i.e. more items in memory) lead to increases in eye-movements. Mechanistically, this behavioural-neuronal interplay

would serve the purpose of strengthening both the spectral amplitude and the phasic profile of theta oscillations to cope with the demands of recalling internal representations in the absence of external stimuli. To test the hypothesis, we recorded the eye movements of human participants in a dual-task paradigm. One of the tasks was perceptual and made saccadic eye movements counter-productive while the other, a variation of the Sternberg task (Sternberg, 1969), manipulated the participants' working memory. Potential changes in saccadic activity would most likely be manifested as microsaccades—involuntary eye movements produced specifically during attempted fixation—rather than saccades, more common during visual exploration, as the experimental design requires constant fixation (Martinez-Conde et al., 2009; Rolfs, 2009). We thus show how increases in working memory load lead to memory recall-specific changes in saccadic eye movements, particularly microsaccades.

3.1 Methods

Participants were seated in a chair at approximately 65 cm in front of a 15" LED screen with a refresh rate of 60 Hz (Figure 9-A). The experiment was implemented using Unity software (Unity Technologies, San Francisco, California, U.S.) and the in-game resolution was 150 × 266 pixels. A keyboard was placed in front of the participant to log their button presses and a Tobii Pro X3-120 eye tracker (Tobii Technology AB, Sweden) with an external processing unit was placed directly below the screen. The sampling rate of the eye tracker was 120 Hz and an example of the collected eye position data can be found in Figure 9. The participants listened to auditory cues through over-ear headphones. Before the experiment started, participants were able to freely adjust the volume of the speakers.

Participants could also adjust the chair's height and back-tilting configuration for comfort and head movements during experimental blocks were limited. They were oriented to face blank, white walls behind the screen. The experiment was conducted in a dedicated empty room at the institute's facilities and the experimenter remained in the same room throughout the session.

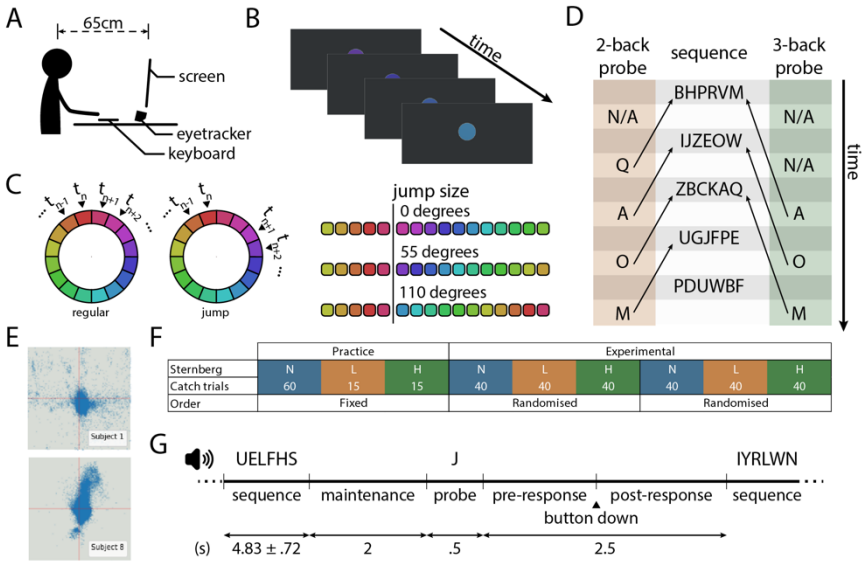


Figure 9. *Experimental equipment and protocol.* A) Illustration of the physical set-up. B) Example of what a participant would see on-screen over time. The only visible change is the colour in the circle for the catch detection task, which happens at 60 Hz except during a hue-jump. C) (left) An illustration of the colour transitions in the catch detection task, with the t's representing timesteps before and after the current timestep, t_n. (right) The squares in each row show the colour for consecutive timesteps from left to right, with a jump occurring at the vertical line. The rows show different sizes of jumps. D) Example of the Sternberg task protocol. The rows are chronological, starting from a block's beginning. The orange and green columns show how the same probe with the same sequence history needs to be compared to different sequences, as indicated by the arrows, depending on the n-back condition. E) Normalised raw gaze positions during the entire experiment from two randomly selected participants. The red lines represent 0.5 on each axis. F) Experimental design: there is one practice block per condition at the start of the experiment, after which there are two experimental blocks per condition that are pseudorandomised. G) The timeline for a Sternberg task trial with examples of what a participant would hear in the top row, the nomenclature describing the trial's stages in the middle row, and the corresponding duration of each stage in the last row.

a) Experimental protocol

Twenty naïve participants (9 females; aged 29.5±4.84 years) with normal or corrected-to-normal vision were recruited from the university campus. Before the experiment began, participants were briefed on the two tasks they were required to perform concurrently—a catch detection task (Figure 9-B and Figure 9-C) and a Sternberg auditory task (Figure 9-D and Figure 9-G)—and gave

their written informed consent. Their dominant eye was ascertained using the Porta test (Roth et al., 2002). The research protocol was approved by the ethical committee (CEIC) of Hospital San Joan de Deu (Barcelona, Spain).

The screen's brightness was adjusted to each participant's preference, and a version of the Farnsworth-Munsell 100 hue test was completed to determine the participants' colour perception acuity (Farnsworth, 1943). The best possible score in the Farnsworth-Munsell hue test is 0, while every error made contributes to the final score the sum of the differences between the erroneous hue and that of its two adjacent hues. The eye tracker was calibrated (mean error<0.5 °) using both the Tobii Pro Eye Tracker Manager (Tobii Technology AB, Sweden) as well as a 5-point routine across both x- and y-axes built into the experiment. Previous work studying eye movements have used similar screen-based eye trackers with comparable accuracy and precision (e.g. Faber et al., 2018; Lloyd et al., 2017; Manning et al., 2014).

The experiment was carried out in English or Spanish, depending on the language the participant was more comfortable with. Participants read on-screen written instructions at the start of the experiment as well as block-specific instructions before each block, reiterating the earlier verbal instructions. Before each block, interaction with the experiment was suspended for 1.5 s after which a button press would start the block. The participants carried out the catch detection task in all the blocks, and each block's working memory load condition corresponded to one of three levels as modulated by the n-back Sternberg task (Sternberg, 1969). The three load conditions were: without the Sternberg task (no load, N), 2-back Sternberg task (low load, L), and 3-back Sternberg task (high load, H).

The experiment consisted of nine blocks, the first three of which were practice blocks, one per load condition and in incremental order. The order of the remaining six blocks, the experimental blocks, was pseudorandomised in groups of three, each group consisting of one block per load condition (Figure 9-F). As the common task in all blocks, the duration of the blocks was dependent on the number of trials in the catch detection task, also referred to as hue-jumps. There were 60 hue-jumps in the first practice block and 15 hue-jumps in the remaining practice blocks. Each experimental block summed 40 hue-

jumps. The experiment lasted approximately 30 minutes and subjects received neither financial nor material remuneration.

b) Catch detection task

An off-white circle (external diameter 2 cm (1.8 °), white border thickness 0.5 mm (0.044 °)) was shown in the centre of the screen on a grey background (HSV: 210, 12.2, 19.2). It was filled with a solid colour (Figure 9-B) that moved progressively across the 360 degrees in the hue space while retaining its colour saturation and value levels (70 and 77, respectively). The saturation and value parameters were chosen to maintain the screen's luminosity at 50% for reduced eye strain. The hue changed at a rate of 60 Hz (in steps of 1 °, 6 s per revolution), and the direction of the change and starting hue were randomised at the start of each block.

Each trial consisted of a sudden shift in the hue progression, i.e. a hue-jump, occurring within an inter-trial interval of 2.5–4 s to prevent habituation. The direction of the hue-jump was chosen randomly (uniformly distributed), clockwise or anti-clockwise, and the hue was changed randomly (uniformly distributed) to 3–30 ° away from the current hue. The distance of the change in hue is referred to as the jump size (Figure 9-D, right), and hue-jumps only occurred when the participant's eyes were being tracked and the eye's gaze position was over the circle. The participants were instructed to press the spacebar on the keyboard when they were at least 75% certain that they had seen a hue-jump.

Prior to the actual experiment, we piloted the experimental setup (6 participants, 2 females, aged 29.6±4.22 years) and observed mean detection rates below 0.5, suggesting that the task was sufficiently challenging. The detection rate of yellow-green jumps was significantly lower than jumps in other hue ranges (0.39±0.19 (red-yellow), 0.12±0.15 (yellow-green), 0.46±0.20 (green-blue), 0.33±0.23 (blue-red), N=6; Friedman, $p=0.0067$, $\chi^2(3)=12$). Hence, to optimise data collection, jumps did not occur within the yellow-green hues (90–180°) in the presented experiment. Although this limited the range of sampled hues, there was still sufficient variability to prevent participants from expecting the occurrence of jumps

simply from the current hue. During the experiment, as all the conditions had to be met before a jump occurred, some trials overshot the 4 s inter-trial interval upper limit. Nonetheless, the overall mean inter-trial interval was 3.4±0.82 s. The catch detection task was carried out for all the blocks.

c) Sternberg auditory n-back task

The basic Sternberg task requires participants to remember a sequence of items, and then, when presented with a probe of a single item, report if the probe was part of the remembered sequence. In this experiment, the items were letters from the Latin alphabet, each sequence consisted of six letters, and a probe was a single letter. The letters were randomly selected and there were no repetitions of letters within a sequence. All stimuli for the Sternberg task were auditory, and participants delivered their responses through keyboard button presses. For example, if the remembered sequence was F-B-Y-E-P-H and the probe was G, the correct answer would be "no".

There was also an n-back component. Instead of comparing the probe with the most recent sequence, participants were asked to compare probes with the nth-last sequence (i.e. the 2nd-last sequence in the 2-back condition and the 3rd-last sequence in the 3-back condition). Sequences and probes were presented alternately, and there was a 50% probability of a probe being correct (i.e. the probe was part of the nth-last sequence) (Figure 9-C). Participants were not required to remember sequences between blocks.

The sequences and probes were read to the participants through headphones (Microsoft Speech API 5.3, U.S.; voice rate=-2). There was a 2 s interval between a sequence and a probe, and a 2.5 s interval after a probe before the next sequence (Figure 9-G). Responses were not accepted while a sequence was being delivered. No sequences or probes would be delivered if the catch detection task was in its final trial.

d) Data analysis

The labelling of eye movements was carried out using a velocity-based algorithm for microsaccade detection (Engbert and Kliegl, 2003) implemented as a package (von der Malsburg, 2015) in R (Team, 2019), and has been used in previous eye movement studies (e.g. Badde et al., 2020; Duchowski, 2007; Yuval-Greenberg et al., 2008). The algorithm eliminates random movements associated with drifts during fixation by using a moving average of velocity vectors over 5 data samples. Surviving vectors are then compared with a threshold for the labelling of saccade and fixation events. The detection thresholds for horizontal and vertical components are calculated independently as a multiple of the velocity time series' standard deviation (λ=6 as recommended by Engbert and Kliegl (2003)). In addition, as the sampling rate was not high (120 Hz), we followed the package developer's recommendations for low-frequency data (<100 Hz), setting the parameters 'smooth.saccades' was to False and 'smooth.coordinates' to True. This applied a linear recursive filter on the gaze positions along each axis, using a kernel [0.333, 0.333, 0.333], smoothing the coordinates of the positions with a moving average of size 3 prior to saccade detection ('smooth.coordinates'). It also prevented consecutive saccades separated by only a few samples from being merged together ('smooth.saccades'). At higher sampling rates, this avoids swing-backs from larger saccades being incorrectly classified as a separate saccade, which is not an effect often found in lower-frequency data.

After the package labelled the onset and offset of saccades as well as their velocity in the x- and y-axis, we calculated their peak velocity and amplitudes. The saccades' peak velocity was the Euclidean norm of its velocity vector. Following the implementation from Engbert et al. (2015), the saccades' amplitude was calculated as the Euclidean norm of the largest positional differences in the two axes. The labelled saccades showed a clear relationship between their amplitude and velocity with considerable saturation of peak velocities at larger amplitudes, satisfying the main sequence criterion (Zuber et al., 1965). The observed velocities were similar to what was found by Martinez-Conde et al. (2009) for saccades during fixations. Although we obtained saccade amplitudes that were larger than expected for microsaccades produced during fixation, it is unlikely to

be due to hardware errors or inaccuracies in the collected data as there is no drift in the data of each participant over time. This stationarity was shown by each subject's rolled-mean gaze position over time (0.5s window, 0.167s minimum period in the event of invalid data or at the edges) in both axes being tested with the augmented Dickey-Fuller unit root test (lag determined with AIC, regression model is constant with no trend), which has a null hypothesis that there is a unit root leading to non-stationarity. All participants had p-values < .05 in both axes, disproving the null hypothesis.

In the analysis for Sternberg stage-dependent saccade rates, two participants were excluded as one had not responded to the catch detection task in some conditions and another had barely responded to the Sternberg task (response rate<25% (L) and <5.6% (H)), which we assumed to mean that they had entirely given up on the Sternberg task.

Before testing for differences between conditions for the various measures, all data was tested for normality using D'Agostino and Pearson's omnibus normality test that combines skew and kurtosis (D'Agostino, 1971). For multiple comparisons between non-parametric data, the Friedman test (Friedman, 1937) was used to determine if there were any significant differences between conditions. For data that was parametric, the default implementation of a repeated measures one-way ANOVA in a Python library (Seabold and Perktold, 2010) was used.

Post-hoc testing or comparisons consisting of only two conditions were carried out using estimation statistics based on Monte Carlo permutation (Ho et al., 2019), with the paired mean difference figures including a bootstrap 95% confidence interval to illustrate the effect size. The confidence interval accounts for skewness in the distributions with a bias-corrected and accelerated bootstrap (Efron, 1987). Each test was based on 5000 resamples. The distribution of all data is reported either as 'condition (mean±std)' when inline or as 'mean±std (condition)' when nested among other statistical information.

The circular-linear correlation (Berens, 2009) is carried out by first defining the correlation coefficients r_{sx} = c(sin α, x), r_{cx} = c(cos α, x)

and r_{cs} = c(sin α, cos α), where *c(x, y)* is the Pearson correlation coefficient. The correlation ρ_{cl} is then calculated using Eq. 2.

$$\rho_{cl} = \sqrt{\frac{r_{cx}^2 + r_{sx}^2 - 2r_{cx}r_{sx}r_{cs}}{1 - r_{cs}^2}} \qquad \textit{Eq. 2}$$

A p-value for ρ_{cl} is computed by considering the test statistic $N\rho^2$, which follows a χ^2-distribution with two degrees of freedom. The linear correlation was a standard Pearson correlation (Pearson, 1895) and the linear regressions were ordinary least squares regressions in the Scikit-learn toolbox (Pedregosa et al., 2011).

3.2 Results

A total of twenty people (9 females; aged 29.5±4.84 years; 15 right-eye dominant) were recruited from the university campus. The participants scored an average of 20±25.4 in the Farnsworth-Munsell test. Their colour acuity did not affect their performance in the catch detection task (Figure 10-A; N=20; Spearman's correlation: p=0.276, r=-0.255). All participants spent at least 80% of the experimental duration fixating on the centre of the screen (examples in Figure 9-E). The distribution characteristics of all data used in analyses can be found in Table 8 and Table 9, including their sample sizes.

a) Catch detection is affected by size more than by hue

Hue-jump events in the catch detection task were characterised by the location in the colour wheel where the jump occurred (hue) and their jump-size magnitude (degrees). We first analysed detection rates, i.e. the probability of a participant reporting a hue-jump, across all experimental blocks depending on either of these jump characteristics (Figure 10-B and Figure 10-C). The hue-dependent detection rate was not uniformly distributed as its Kullback-Leibler divergence score significantly differed from that of randomly generated uniform distributions with the same means and standard deviations (Figure 10-D and Figure 10-E; Monte Carlo permutation, p<0.001, We then *performed a circular-linear regression on the hue-dependent*

*detection rates and a linear regression on the jump size-dependent detection rates per participant. We found that correlation p-values for jump size was significantly lower than that of hue (*Figure 10-F*; Monte Carlo permutation, p<0.001,* Table 7, *comparison 2). To better understand the profile of jump size-dependent detection rates, we split detection rates into three size bins: short (3–10°), medium (11–20°) and long (21–30°) (*Figure 10-G*). Detection rates were significantly different between all bins (*Figure 10-H *and* Figure 10-I*; one-way repeated measures ANOVA, p<0.001, F(2, 38)=65; Monte Carlo permutation, p<0.001,* Table 7, *comparisons 3, 4, and 5), showing a modulation of task difficulty.*

Table 7, comparison 1).

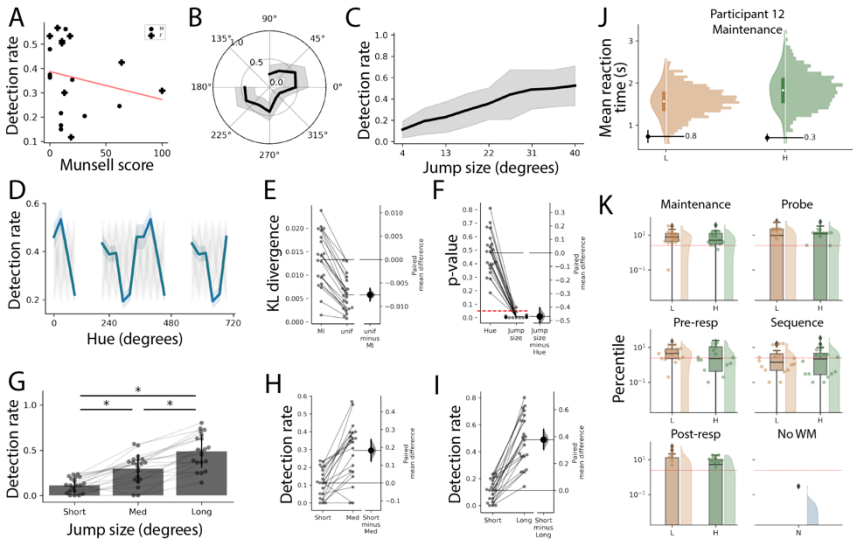


Figure 10. *Catch detection performance is modulated by task variables and not Sternberg task stage.* A) Correlation of individual participants' performance in the catch detection task and their colour acuity. B) The mean hue-dependent detection rate across all blocks (± std). C) The mean jump size-dependent detection rate across all blocks (± std). D) The mean hue-dependent detection probability in blue (± std), and random uniform distributions in grey. E) A paired comparison between the Kullback-Leibler divergence score for each participant's hue-dependent detection rate and that of random uniform distributions. F) Paired mean difference between the p-values of the regressions for each participant's hue-dependent and jump size-dependent detection rates. The red, dotted line indicates p=0.05. G) The dots indicate participants' individual mean detection rate for jumps within that bin, while the bar plots show the population's mean and standard deviation. H & I) The paired mean differences in detection rates of medium and long jumps compared to short jumps. J) An example of how a participant's percentile score per Sternberg stage is calculated. The histogram and violin plot are from the permutations' mean reaction times. The white dots show their means, while the shaded bars indicate their standard deviations. The participant's actual mean reaction time is represented by the black error bar to their left,

and the label is its resulting percentile. K) Boxplots of the participants' percentiles relative to shuffled permutation in different Sternberg stages, as described in Figure 9-G, and when there is no Sternberg task ('no WM'). The red, dotted line indicates p=0.05.

We then performed a circular-linear regression on the hue-dependent detection rates and a linear regression on the jump size-dependent detection rates per participant. We found that correlation p-values for jump size was significantly lower than that of hue (Figure 10-F; Monte Carlo permutation, $p<0.001$, Table 7, comparison 2). To better understand the profile of jump size-dependent detection rates, we split detection rates into three size bins: short (3–10°), medium (11–20°) and long (21–30°) (Figure 10-G). Detection rates were significantly different between all bins (Figure 10-H and Figure 10-I; one-way repeated measures ANOVA, $p<0.001$, $F(2, 38)=65$; Monte Carlo permutation, $p<0.001$, Table 7, comparisons 3, 4, and 5), showing a modulation of task difficulty.

Table 7. *Post-hoc Monte Carlo permutation statistics.* ID: comparison index; KL: Kullbeck-Liebler divergence score; Unif: Random sample from a uniform distribution; Corr-p: correlation p-value; Size: jump size; DR: detection rate; Med: medium jump sizes (binned); RT: reaction time Shuff: shuffled; N: no cognitive load; L: low cognitive load; H: high cognitive load; BR: blink rate; FR: fixation rate; MSR: microsaccade rate; MSR-R: microsaccade rate during recall for Sternberg task.

ID	Measure	Control	Test	Difference	CI (95%)		p-value
					lower	upper	
1	KL	Hue	Unif	−0.0092	−0.011	−0.0075	<0.001
2	Corr-p	Hue	Size	−0.47	−0.54	−0.41	<0.001
3	DR	Short	Med	0.18	0.12	0.24	<0.001
4	DR	Short	Long	0.38	0.31	0.45	<0.001
5	DR	Med	Long	0.19	0.14	0.25	<0.001
6	RT	Actual	Shuff	0.90	0.82	0.99	<0.001
7	Accuracy	L	H	−0.081	−0.15	−0.013	0.039
8	DR	N	L	−0.14	−0.18	−0.098	<0.001
9	DR	N	H	−0.13	−0.17	−0.091	<0.001
10	DR	L	H	0.01	−0.022	0.041	0.56
11	RT	N	L	0.13	0.076	0.18	0.0012
12	RT	N	H	0.092	0.035	0.12	0.0016
13	RT	L	H	−0.042	−0.078	0.0069	0.0634
14	Gradient	N	L	−0.0056	−0.011	−0.0018	0.025
15	Gradient	N	H	−0.0084	−0.013	−0.0044	0.0016
16	Gradient	L	H	−0.0028	−0.0067	−0.0013	0.19
17	BR	N	L	0.025	0.0039	0.050	0.072
18	BR	N	H	0.018	0.0030	0.036	0.076
19	BR	L	H	−0.0069	−0.035	0.0081	0.63
20	FR	N	L	0.17	0.078	0.29	0.0026
21	FR	N	H	0.093	0.033	0.20	0.014

22	FR	L	H	−0.074	−0.14	−0.23	0.023
23	MSR	N	L	0.17	0.078	0.29	0.0026
24	MSR	N	H	0.093	0.033	0.20	0.014
25	MSR	L	H	−0.074	−0.14	0.023	0.023
26	MSR-R	L	H	0.040	0.010	0.071	0.020

Table 8. *Distribution characteristics for behavioural data.* n: sample size; SD: standard deviation; K^2: D'Agostino-Pearson omnibus normality test statistic; p-value: normality test p-value; KL: Kullbeck-Leibler divergence score; Corr-p: correlation p-value; DR: detection rate; RT: reaction time; Gradient: response gradient; BR: blink rate; FR: fixation rate; MSR: microsaccade rate; MSR-R: microsaccade rate during recall for Sternberg task; Unif: Random sample from a uniform distribution; Size: jump size; Med: medium jump sizes (binned); Shuff: shuffled; R(Stern): proportion of all trials responded to in the Sternberg task; C(Stern): proportion of all trials correctly responded to in the Sternberg task; N; no cognitive load; L: low cognitive load; H: high cognitive load; Gradient: response gradient.

Measure	**Unit**	**Condition**	**n**	**Mean**	**SD**	**K^2**	**p-value**
KL	-	Hue	20	0.013	0.0058	-	-
KL	-	Unif	20	0.0041	0.0026	-	-
Corr-p	-	Size	20	0.011	0.18	-	-
Corr-p	-	Hue	20	0.48	0.15	-	-
DR	-	Short	20	0.11	0.080	4.3	0.12
DR	-	Med	20	0.30	0.16	0.50	0.78
DR	-	Long	20	0.50	0.18	1.5	0.47
RT	s	Actual	20	0.84	0.104	-	-
RT	s	Shuff	20	1.7	0.16	-	-
R(Stern)	-	L	20	0.88	0.19	20.	<0.001
R(Stern)	-	H	20	0.83	0.24	17.	<0.001
C(Stern)	-	L	20	0.55	0.15	3.2	0.21
C(Stern)	-	H	20	0.44	0.17	2.3	0.32
Accuracy	-	L	20	0.63	0.11	-	-
Accuracy	-	H	20	0.55	0.16	-	-
DR	-	N	20	0.43	0.12	6.7	0.035
DR	-	L	20	0.29	0.15	0.90	0.64
DR	-	H	20	0.30	0.15	0.84	0.66
RT	s	N	19	0.70	0.11	10.	0.0055
RT	s	L	19	0.83	0.11	3.3	0.19
RT	s	H	19	0.79	0.087	0.79	0.67
Gradient	/°	N	20	0.023	0.007	2.8	0.24
Gradient	/°	L	20	0.018	0.008	1.5	0.48
Gradient	/°	H	20	0.015	0.009	1.3	0.52

b) Participants remained engaged in both tasks

As it is crucial that participants executed both auditory and visual tasks concurrently to load their working memory while they maintained fixation, we compared their individual reaction times in the catch detection task to the same responses shuffled in time. We permutated each participant's responses for the no-load practice block and experimental blocks over time (1000 permutations per subject per block), and used the block's actual jump events to estimate the quartile ranges. The mean reaction time per block of each subject was significantly lower than the shuffled mean reaction time (Monte Carlo permutation, $p<0.001$, Table 7, comparison 6), confirming that the responses were not random at a block level.

The participant's mean reaction time per block for each Sternberg stage was calculated as a percentile in the distribution of shuffled reaction times in the same stage (Figure 10-J). We found that in all stages, the actual mean reaction times were within the 10th percentile of the shuffled distribution while in the no load condition most participants were in the 0th percentile (Figure 10-K), suggesting that they actively carried out the catch detection task in all working memory load conditions.

Next, we checked the proportion of Sternberg trials that were responded to as opposed to skipped (Figure 11-A). In both and high load conditions, participants responded to an average of about 80% of trials. As expected, participants responded to significantly fewer trials in the high load than in the low load condition (Wilcoxon signed-rank, $p=0.036$, $W=23.0$, $r=0.11$), demonstrating their engagement in the Sternberg task throughout the experiment. The number of correct responses, as a proportion of all trials, was significantly lower in the high load than the low load condition (paired Student’s t-test, $p=0.0022$, $t(19)=3.5$). Taken as a proportion of only responded trials, participants were significantly less accurate in the high than in the low load condition (Figure 11-B; Monte Carlo permutation, $p=0.039$, Table 7, comparison 7), confirming that the decrease in correct responses was not simply due to a decrease in the number of responses.

c) Performing both tasks concurrently led to task interference

Loading the working memory was significantly detrimental to both detection rate (Figure 11-C; Friedman, $p<0.001$, $\chi^2(2)=24.9$; Monte Carlo permutation, $p<0.001$, Table 7, comparisons 8 and 9) and reaction time (Figure 11-D; Friedman, $p<0.001$, $\chi^2(2)=16.5$; Monte Carlo permutation, $p<0.05$, Table 7, comparisons 11 and 12) in the catch detection task, although there were no differences between conditions with working memory load (Table 7, comparisons 10 and 13). One participant's reaction time was excluded from analyses due to a lack of response in the catch detection task.

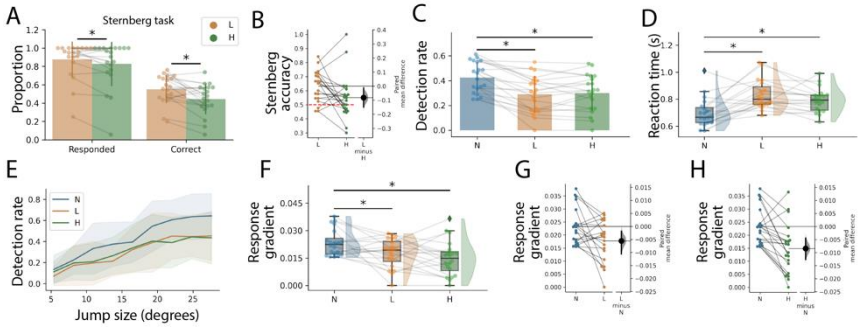


Figure 11. *Behavioural markers of task interference.* A) Number of Sternberg trials that were responded to or correct, normalised to the total number of trials. B) Paired mean difference in the proportion of responded trials that were correct between the low and high load conditions. The red, dotted line indicates chance performance. C) The detection rate in the catch detection task, and D) The mean reaction time in the catch detection task. E) The mean detection rate in the catch detection task per load condition relative to jump size (± std), from which F) the mean response gradients are obtained. G and H) The paired mean differences in response gradients with the no load condition as baseline.

Jump discretisation (9 bins) revealed that detectability of a jump evolves as a function of both its size and the working memory load condition. The averaged trend of larger jumps being detected more frequently is again seen for all load conditions (Figure 11-E). Individual participants' detection rate per cognitive load condition across jump sizes were fitted to linear regressions, with the gradient of the fit representing their response gradient. We found that response gradients were significantly larger in the no load condition than in both the low and high load conditions (Figure 11-F and Figure 11-G and Figure 11-H; one-way repeated measures ANOVA, $p=0.0016$, $F(2, 38)=7.7$; Monte Carlo permutation, $p<0.05$, Table 7, comparisons 14, 15, and 16), implying that working memory load leads to a task-specific modulation of performance instead of a generic impairment across all jump sizes.

d) Eye movements increase when working memory is loaded

We then analysed the participants' eye movements and identified periods of saccades and fixations from the continuous gaze recordings (Figure 12-A and Figure 12-B). We did not observe significant changes in gaze fixation areas across the load conditions (Figure 12-C; Friedman, p=0.91, $\chi^2(2)=0.18$), thus suggesting that subjects maintained the position of their gaze regardless of the load condition. Blink rates differed between the no load condition and the low and high load conditions, although this did not survive permutation testing (Friedman, p=0.037, $\chi^2(2)=6.6$; Monte Carlo permutation, p>0.05, Table 7, comparisons 17, 18, and 19).

Table 9. *Distribution characteristics for eye-tracking data.* n: sample size; SD: standard deviation; K^2: D'Agostino-Pearson omnibus normality test statistic; p-value: normality test p-value; Fix. area: fixation area; BR: blink rate; FR: fixation rate; MSR: microsaccade rate; MSR-R: microsaccade rate during recall for Sternberg task; N; no cognitive load; L: low cognitive load; H: high cognitive load.

Measure	Unit	Condition	n	Mean	SD	K^2	p-value
Fix. area	px^2	N	11	15.	8.1	2.3	0.32
Fix. area	px^2	L	11	13.	5.5	2.2	0.33
Fix. area	px^2	H	11	15.	9.4	14.	<0.001
BR	Hz	N	17	0.024	0.025	1.8	0.39
BR	Hz	L	17	0.049	0.046	6.5	0.039
BR	Hz	H	17	0.042	0.032	4.7	0.096
FR	Hz	N	20	0.59	0.40	1.5	0.47
FR	Hz	L	20	0.51	0.33	1.2	0.55
FR	Hz	H	20	0.42	0.22	2.1	0.34
MSR	Hz	N	20	0.42	0.21	2.9	0.24
MSR	Hz	L	20	0.59	0.39	0.052	0.97
MSR	Hz	H	20	0.51	0.32	0.24	0.89
MSR-R	Hz	L	18	0.12	0.14	-	-
MSR-R	Hz	H	18	0.081	0.095	-	-

We next measured the fixation rate (fixations per second, Hz) across conditions and found significantly higher fixation rates in both the low and high load conditions compared to the no load condition (Figure 12-D and Figure 12-E and Figure 12-F; one-way repeated measures ANOVA, p=0.0019, F(2, 38)=7.4; Monte Carlo permutation, p<0.05, Table 7, comparison 20 and 21). Moreover, the statistical trend, found in reaction times and response gradients,

reached significance, with fixation rates being significantly lower in the high load than the low load condition (Monte Carlo permutation, p=0.023, Table 7, comparison 22). Saccade rate is closely related to fixation rate, and there were also significant differences in saccade rates between all load conditions (Figure 12-G and Figure 12-H and Figure 12-I; one-way repeated measures ANOVA, p<0.001, F(2, 38)=14; Monte Carlo permutation, p<0.05, Table 7, comparisons 23, 24, and 25), suggesting that working memory load leads to increases in eye movement specific to saccades as gaze drift during fixation, measured by fixation area, was unaffected.

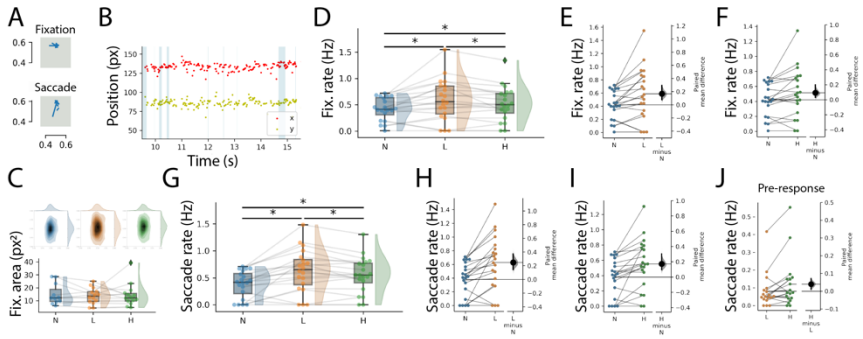


Figure 12. *Eye movements increase in conjunction with demand for working memory.* A) Examples of labelled gaze trajectories. Both examples show data from a time window of 800 ms. B) A 5 s extract from a sample participant's raw gaze positions (x & y) in pixels illustrating how it is relatively stable. The highlighted regions indicate movements that were labelled as saccades. C) The fixation area in square pixels, with the mean normalised shape of all participants' fixations above each working memory condition's column. D) The fixation rates per condition in Hz and E & F) their paired mean differences with the no load condition as baseline. G) The saccade rates per condition and H & I) their paired mean differences with the no load condition as baseline. J) The paired mean difference in saccade rate between the high and low load conditions during the Sternberg task's pre-response stage.

e) Eye movements increase specifically during memory recall

We further analysed the saccade rate depending on the stage of the working memory task, as that could provide insight into how saccades are modulated in conjunction with working memory functions. Specifically, we were interested in how the saccade rate changed during the pre-response stage (i.e. when subjects would access sequences stored in their working memory). We found that despite a decrease in overall saccade rates as reported earlier, during

this stage saccade rates were significantly higher in the high load than in the low load condition (Figure 12-J; Monte Carlo permutation, p=0.02, Table 7, comparison 26). Thus, although subjects fixated more consistently in high load compared to low load conditions, the increased demand on their working memory induced a greater occurrence of saccades when they accessed their working memory.

3.3 Conclusions

We hypothesised that saccades contribute to the retrieval of remembered information, implying that increased working memory load would require more saccadic eye movements. This is of particular interest given that oscillations in the theta-band (4–8 Hz) have been shown to carry information about working memory items through their frequency and phase (Axmacher et al., 2010; O'Keefe and Recce, 1993). Moreover, it is known that eye movements, especially saccades, aid in inducing phase-resets of theta oscillations for working memory both during encoding and recall (Hoffman et al., 2013; Jutras et al., 2013). We therefore quantified eye movements and behavioural responses of human participants performing a working memory task with distinct memory load conditions. Simultaneously, they also carried out a perceptual, catch detection task where maintaining fixation was useful for sampling visual stimuli.

Consistent with previous studies, we found task interference effects when participants carried out both tasks concurrently even though the stimuli for the tasks were from different sensory modalities (Woodman and Luck, 2004). We also showed an uptick in saccade rate which is associated with increased working memory load. Specifically, we found that participants broke fixation significantly more often when their working memory was loaded than when they were performing the catch detection task alone. Other eye movement measures were unaffected by the working memory load. An increase in saccade rate was found between memory load conditions precisely during the Sternberg task stage when participants would be accessing their working memory—the pre-response stage—suggesting that saccades are involved in memory recall.

There was behavioural evidence of task interference even though the catch detection task did not explicitly require working memory. This suggests that the catch detection task had a top-down element that competed with the Sternberg task for processing resources, as passive maintenance of working memory items was shown to be insufficient for task interference (Han and Kim, 2004). This is supported by response gradients in loaded working memory conditions being significantly smaller than in the no load condition. The difference indicates that working memory load affected larger jumps more than smaller jumps, as a uniform decrease in performance across all jump sizes would lead to negligible changes in response gradients. It suggests that top-down signals identify task features that are relevant and modulate sensitivity to them when competition for cognitive resources reduces what is available for sensory processing (Mathews et al., 2015).

While the mean reaction times of participants in the catch detection task was significantly quicker than the permuted reaction times, this difference was less clear when comparing reaction times of jumps occurring within each Sternberg task stage (Figure 10-H). However, this is in large part due to the durations of certain stages being very short and, therefore, consisting of small sample sizes. With that in mind, the low percentiles in all Sternberg task stages combined with most participants being consistently in the 0th percentile in the no load condition confirm that the participants were not randomly reporting jumps even when their working memory was loaded.

It has been shown that top-down inhibitory processes can be impaired by working memory load, resulting in less control over reflexive eye movements (Mitchell et al., 2002; Walker et al., 1998). It remains ambiguous, however, if higher working memory loads reduce the effectiveness of saccade suppression from top-down regions or if they lead to a greater need for saccadic eye movements to phase-lock theta for memory recall. Should top-down suppression apply to all eye movements, it would also affect blink rates which are known to be proportionally modulated by working memory load (Chen and Epps, 2013; Veltman and Gaillard, 1998). While fatigue has been shown to affect blink rates (Stern et al., 1994), it is an unlikely confounding factor as the load condition presentation was randomised in the present study. Indeed, participants did not exhibit any clear differences in blink rates between working memory

conditions when evidence from literature points to it increasing with load, suggesting that top-down inhibitive suppression of eye movements did not change significantly across conditions. If that is the case, it supports the hypothesis that the pre-response stage's elevated saccade rate in the high load compared to the low load condition (Figure 12-J) is due to a greater elicitation of saccades rather than a reduced ability to suppress them.

The results support the hypothesis that the Validation Gate inhibitory mechanism is generalisable to visual features other than spatial location. It also suggests that saccadic eye movement is an integral part of working memory. Mechanistically, the reported results here highlight a potential solution to a crucial aspect of oscillatory representation in working memory—that it must be sustained even in the absence of the external stimuli it represents. Eye-movements aid in eliciting oscillatory phase-locking to overcome the lack of externally-driven stimulation to maintain and recall the embedded items' information. In most related experiments, eye movements are necessary to the task, which confounds the purpose of the saccadic activity that they observed. Here, we have demonstrated that saccades do also contribute to functions separate from active sensing and that they are provoked by these functions even when eye movements are suppressed in general.

4. THE NEURAL SUBSTRATE OF SELECTIVE ATTENTION

Early research in selective attention used a spotlight metaphor to describe the top-down biasing of sensory processing, with sensory information within the moveable beam being processed while everything outside of it is not. It implies that spatial attention is limited in size and can be shifted within the visual field. This spotlight model of attention (Eriksen and Hoffman, 1972; Posner et al., 1980), along with its successor the zoom-lens model (Eriksen and James, 1986), opened experimental paradigms which have led to the understanding that top-down processes can bias sensory processing by predicting regions or objects of interest and increasing sensitivity to them (Itti and Koch, 2001; Reynolds and Heeger, 2009). Exactly how the two processes interact, however, remains a point of discussion that is relevant to multiple fields, including robotics and neuroscience.

A number of models of attention have been proposed with varying degrees of biological detail (e.g. Borji and Itti, 2013; Moore and Zirnsak, 2017). A mainstay of bottom-up attention modelling is the saliency map, which defines locations of interest in perceptual space in terms of the presence of sensory features, such as orientation, colour and brightness (Itti et al., 1998). The saliency map as a representation of bottom-up attention remains influential till today (e.g. Murray et al., 2011), and is also the main interface for top-down and bottom-up attention in most models. Models of top-down attention need to contend with the challenge of guiding perceptual processing irrespective of the dominance of the various bottom-up features. In visual search, this often takes on some form of the biased competition model, which describes how neuronal response to simultaneously-presented stimuli is a weighted average of the individual response and explains selective attention as the biasing of these weights towards attended or task-relevant stimuli (e.g. Usher and Niebur, 1996). One prominent example is the attention normalisation model where a feedforward stimulus drive is modulated by an inhibitory suppressive drive comprising the presence of non-preferred stimuli, which normalises the population response to the stimuli (Reynolds and Heeger, 2009). To counter this

averaging effect, a top-down attention field enhances activity of neurons that represent task-relevant stimuli. Although the model was proposed to account for attentional effects at the single-neuron level, it has also been extended through experiments to the neuronal-population level (Reddy et al., 2009).

Biased saliency maps have also been used to locate targets by dynamically adjusting the tuning of multi-scale salience maps by exploiting the known features of preferred stimuli (Navalpakkam and Itti, 2006). The classic Adaptive Resonance Theory describes how the brain attends and learns from stimuli, where attentional focus is a manifestation of a resonant state that emerges from the matching of expected and encountered stimuli (Grossberg, 2013, 1987). The more recent rise in accessibility and popularity of machine learning has seen its adoption in the implementation of attention models as well, especially in the field of computer vision. In one such implementation, the targets of selective attention could be learned in a so-called visual dictionary by a layered neural model through supervised training (Yang and Yang, 2017).

The guiding principle behind most models is to identify task-relevant stimulus features and to amplify activity representing them such that it overshadows activity for other features, which are assumed to be task-irrelevant. This functions well in situations where task-relevant features are known, which is the case for the typical visual search task paradigm that has been one of the most long-standing benchmarks for the performance of attention models (e.g. Mirza et al., 2019). On the flip side, there are situations where task-relevant stimuli or features are not clearly defined or predictable. In these cases, top-down inhibition of more predictable task-irrelevant stimuli or features would be more practical, as demonstrated in previously mentioned psychophysical studies (e.g. Chabris et al., 2011). Indeed, it was shown that when searching for an object based only on its identity, saccadic reaction time was significantly faster when the target object was both identity- and colour-matched than when it was only identity-matched (Foerster and Schneider, 2018), suggesting that task-irrelevant information biases top-down attention even in typical search tasks.

Neurophysiological studies have elucidated a potential model of such an inhibitory mechanism in the brain, based on the connections between the frontal cortices and the thalamic reticular nucleus (Barbas, 2000; Behrens et al., 2003; Zikopoulos and Barbas, 2012). There is also experimental evidence for a thalamocortical pathway for such a mechanism as opposed to a corticocortical one, notably through thalamocortical manipulation in mice during a cross-modal attentional task (Wimmer et al., 2015). The authors showed that activity in the mice's visual thalamic reticular nucleus was predicted by the task-relevant modality and was dependent on prefrontal cortex activity. It supports the idea that top-down attentional modulation can be inhibitory and the physiological model described by Zikopoulos and Barbas (2006) is again supported by another study that demonstrated the modulation of thalamic reticular neurons by shifts in attention (McAlonan et al., 2006). Furthermore, the prefrontal cortex's mediation of sensory processing has been found as early in the processing stream as the lateral geniculate nucleus (O'Connor et al., 2002) and top-down attention has been argued to anticipatorily modulate sensory processing, also known as signal-suppression (Gaspelin et al., 2015), as distractors elicited similar response latencies and behavioural performance in a visual attention task, regardless of their bottom-up saliency (Cosman et al., 2018). Another study also found that the chemogenetic silencing of corticostriatal neurons in rats' prefrontal cortex did not affect attentional parameters even when it reduced proactive inhibitory control (Terra et al., 2020), suggesting that while corticocortical pathways do play a role in inhibitory behavioural control it may be for a specific subset of behaviours rather than for attentional functions.

The aim of this study was thus twofold: firstly, to examine how the model performed attentional tasks which can later be extended to real-world applications and, secondly, to provide insight on the neural substrate that underlie the model as one of the mechanisms of selective attention in the brain. To do so, we analysed a model of the thalamocortical system comprising spiking neurons. We showed that the prefrontal cortex can flexibly inhibit the activity of neurons in a task-dependent manner. Through the active, anticipatory top-down

manipulation of inhibitory receptive fields (RFs) in the sensory cortices, our model selected task-relevant stimuli even in the presence of more salient distractors. In addition, the observed changes in the firing pattern of the simulated thalamic neurons reflect pertinent responses observed in the mammalian brain by Mease et al. (2014). Our results showed how the strength of the corticothalamic connection modulates the neural responses of the thalamic relay nucleus, with implications from the proposal that neuronal firing patterns are behaviourally functional (e.g. Bezdudnaya et al., 2006; Grossberg and Versace, 2008; Yu et al., 2009).

4.1 Methods

The model was implemented following the general blueprint of the thalamocortical visual processing hierarchy, with neuronal populations roughly corresponding to brain regions, but its principles can be applied to any of the sensory modalities. Information from the retina projects to the lateral geniculate nucleus followed by the primary visual cortex (van Essen and Maunsell, 1983), the lateral geniculate nucleus being the visual sensory stream's specific thalamic nucleus. The primary visual cortex has reciprocal connections to the pulvinar, which is the visual non-specific thalamic nucleus. In this study, we used a simplified version of the feed-forward connections found between the higher-order visual cortices and the pulvinar (Sherman and Guillery, 2002).

Sensory information eventually reaches the frontal cortex, which provides the top-down signal given its established part in executive functions, including selective attention (e.g. Dedoncker et al., 2016; Ridderinkhof et al., 2004; Zanto et al., 2011), and retinotopic organisation similar to the thalamic nuclei, albeit in lower resolutions (Kastner et al., 2007; Mikami et al., 1982). It has projections back into the thalamus and the thalamic reticular nucleus (Barbas, 2000; Behrens et al., 2003). The thalamic reticular nucleus has extensive inhibitory connections to the thalamus (Deleuze and Huguenard, 2006; Guillery and Harting, 2003), including the pulvinar for visual information, thus closing the loop and providing a means for top-down-driven inhibitory modulation of sensory processing.

a) Network-level description

The model was implemented in Python (Rossum, 1995) using a spiking neural network adapted from van Wijngaarden et al. (2016) and visualised in Figure 13-A. It had a population loosely representing the primary visual cortex, from which we derived the saliencies of represented features from local firing rates within the population. Other populations represented the thalamic reticular nucleus (TRN), the specific (SP) and the non-specific thalamic nuclei (NSP), with connections based on Guillery (1995) and Jones (2002). The thalamic populations corresponding with the thalamic relay nuclei consisted of 100 excitatory neurons each, and the thalamic reticular nucleus' population had 100 inhibitory neurons. The sensory cortical population had 800 excitatory (SE) and 200 inhibitory (SI) neurons, modelled as quadratic integrate-and-fire cells (Izhikevich, 2003). Spiking neurons were used as it allowed us to gain insight into the changes in spiking patterns caused by the interacting regions.

The modelled prefrontal cortex (PFC) was a mean-field approximation (Figure 13-B) from an established biophysically-based binary decision-making model (Wilson and Cowan, 1972). This reduced unnecessary complexity in the model while effectively implementing a decision-making mechanism, as the current study is interested in the thalamocortical interaction leading to selective attention rather than the decision-making mechanism of top-down attention. The two strategies for PFC→TRN excitation were chosen to be either goal-oriented or not, relative to the tasks found in **Experimental Configuration**. In one task (target detection), the prefrontal cortex received input from the non-specific nucleus. The projection had a probability of 0.8 per spike to introduce stochasticity to the model. In the other task (oddball detection), the prefrontal cortex received input from the sensory cortex as that allowed for orientation representation. Each projected spike to the prefrontal cortex contributed to the activity of a decision pool in the mean-field model, given by the linear equation Eq. 3, depending on its task-related characteristic.

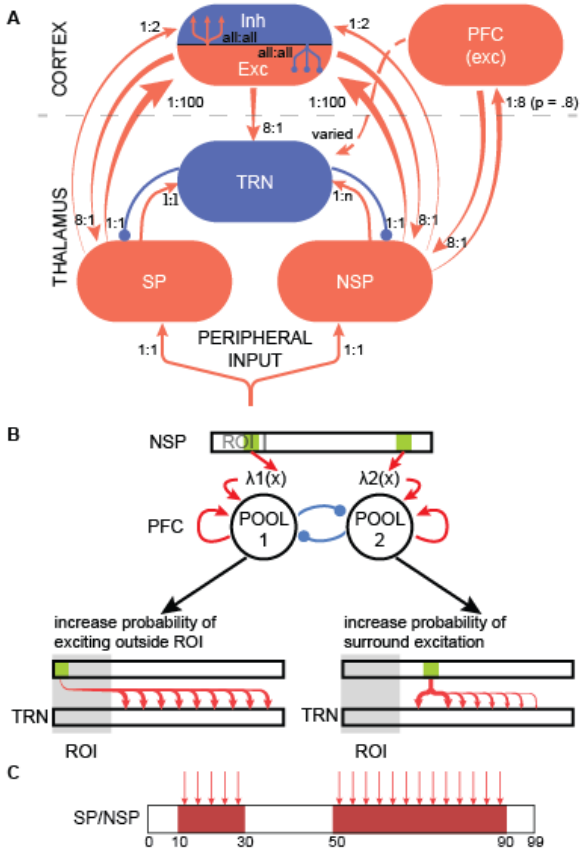


Figure 13. *Model architecture.* A) The different populations of spiking neurons used in the model, with blue representing inhibitory connections and red representing excitatory ones. SP: specific nucleus; NSP: non-specific nucleus; TRN: thalamic reticular nucleus; Exc: sensory cortex excitatory neurons; Inh: sensory cortex inhibitory neurons; PFC (exc): mean-field model of the prefrontal cortex. The numbers beside the arrows indicate the connectivity pattern of that connection. Further clarification is available in Tables 1 and 2. B) Visualisation of the connection PFC→TRN, with associated strategies based on the target detection task. The green rectangles are example NSP neurons that were active in the previous time step, contributing to the input, λ, of the pools depending on their position. The mean-field model is thus biased by the NSP's activity to select either (1) a goal-oriented pattern or (2) a pattern typical of PFC→TRN connections. C) An example of how the SP and NSP is stimulated by a peripheral current, with the red regions indicating the neurons stimulated. Here, the input is of two stimuli of widths 20 and 40 neurons each.

The decision pools represented the competing strategies for top-down excitation of the thalamic reticular nucleus.

$$\lambda_j = \frac{\phi_{upper} - \phi_{lower}}{\Phi} \times N_j + 5 \qquad \textit{Eq. 3}$$

Where λ_j is the activity of decision pool *j* in *spikes/s*, ϕ_{upper} and ϕ_{lower} are the upper and lower limits for the pool's activity for competition (Marcos et al., 2013), Φ is the experimentally determined highest spiking rate of the population (=75 spikes/s) and N_j is the sum of spikes from neurons associated with the decision pool.

Table 10. *Connection parameters for the target detection task.* The connection type was either excitatory (+) or inhibitory (−), with values based on van Wijngaarden et al. (2016) and Proske et al. (2011). Values in brackets in the 'connectivity' column represent the probability of connections. Input: peripheral input; SP: thalamic specific nucleus; NSP: thalamic non-specific nucleus; TRN: thalamic reticular nucleus; SE: excitatory neurons in sensory cortex; SI: inhibitory neurons in sensory cortex; PFC: mean-field model of prefrontal cortex.

From	To	Type	g_s (mS)	Connectivity	τ_s	Delay (ms)
Input	SP	+	0.005	1:1	10	0
Input	NSP	+	0.005	1:1	10	0
SP	TRN	+	0.018	1:1	10	3
TRN	SP	–	0.35	1:1	75	3
TRN	NSP	–	0.18	1:1	75	3
NSP	TRN	+	0.015	1:many (0.15)	10	3
SE	TRN	+	0.002	8:1	7	7
SE	SP	+	0.007	8:1	7	7
SE	NSP	+	0.002	8:1	7	7
SP	SE	+	0.002	1:8	7	7
SP	SI	+	0.002	1:2	7	7
NSP	SE	+	0.8	1:8	7	7
NSP	SI	+	0.8	1:2	7	7
SE	SE/SI	+	0.3	random	0	0
SI	SE/SI	–	1.0	random	0	0
PFC	NSP	+	0.005	1:1 (0.8)	7	7
PFC	TRN	+	0.03	varied	7	7
NSP	PFC	+	N/A	1:1 (0.8)	N/A	3

The winning decision pool engaged its associated connectivity matrix to the thalamic reticular nucleus semi-deterministically by exciting selected thalamic reticular nucleus neurons with a probability of 0.8. The prefrontal cortex also reciprocally excited the non-specific nucleus but did not connect to the specific nucleus at all (Zikopoulos and Barbas, 2006). The connectivity matrices were organised to reflect the nature of the tasks—spatially-ordered or orientation-preferring—rather than with a fixed random probability, and are further elaborated on in later sections.

Table 11. *Connection parameters for the oddball detection task.* The connection type was either excitatory (+) or inhibitory (−), with values based on van Wijngaarden et al. (2016) and Proske et al. (2011). Values in brackets in the 'connectivity' column represent the probability of connections. The weights of the connections are significantly lower than in the 1D task because of the increase in number of connections in the 2D task. Input: peripheral input; SP: thalamic specific nucleus; NSP: thalamic non-specific nucleus; TRN: thalamic reticular nucleus; SE: excitatory neurons in sensory cortex; SI: inhibitory neurons in sensory cortex; Mock PFC: model of prefrontal cortex.

From	To	Type	g_s (mS)	Connectivity	τ_s	Delay (ms)
Input	SP	+	0.005	1:1	10	0
Input	NSP	+	0.005	1:1	10	0
SP	TRN	+	0.018	1:1	10	3
TRN	SP	–	0.35	1:1	75	3
TRN	NSP	–	0.18	1:1	75	3
NSP	TRN	+	0.015	1:many (0.15)	10	3
SE	TRN	+	0.0067	19:1	7	7
SE	SP	+	0.0023	19:1	7	7
SE	NSP	+	0.0067	19:1	7	7
SP	SE	+	0.00011	1:24	7	7
SP	SI	+	0.00011	1:6	7	7
NSP	SE	+	0.039	1:24	7	7
NSP	SI	+	0.039	1:6	7	7
SE	SE/SI	+	0.3	random	0	0
SI	SE/SI	–	1.0	random	0	0
PFC	NSP	+	0.005	1:1 (0.8)	7	7
PFC	TRN	+	0.07	varied	7	7
NSP	PFC	+	N/A	1:1 (0.8)	N/A	3

When spatially organised, the excitatory connection of the NSP → SE and SP → SE was changed from a 1:many connection to a 1:24 connection. The ratio was chosen to approximate the number of connections in van Wijngaarden et al. (2016). This implies that neurons do project to overlapping areas, but maintains the average level of excitation that would have occurred in the original connection. For the same reasons, the connection of TRN → NSP was changed from 1:many to 1:1.

When the cortical receptive fields were organised such that the neurons were orientation-selective, it led to significantly more connections than in the case of van Wijngaarden et al. (2016) or the spatially organised population. Hence, the weights of the connections between the SP, NSP and SE were reduced to elicit approximately the same volume of spiking response to a stimulus in the spatially organised case.

The input to the network was held constant over the exposure duration, which could differ from the experimental duration depending on the experimental protocol, and is defined by a peripheral current which can influence specific neurons. For example, Figure 13-C illustrates the effect of a two-part stimulus exciting neurons 10–30 and 50–80. The magnitude of the current is 5.0 μ F/cm^2.

b) Neuron-level description

The thalamic neurons had characteristic membrane properties which included polarisation-dependent inactivation of T-type Ca^{2+}-channels, allowing the production of low-threshold calcium spikes under hyperpolarised conditions (burst). The neurons were single-compartment cells which changed their membrane potentials depending on an input current, I_{in}, a calcium current, I_T, and a constant conductance leak current, I_L, as described by Eq. 4.

$$C\frac{dV}{dT} = I_{in} - I_T - I_L \qquad \textit{Eq. 4}$$

The slow variable, h, modelled the release of the inactivation of these calcium currents in a variation of the classical conductance-based leaky integrate-and-fire dynamics (Smith et al., 2000). It became zero at depolarised levels when $V > V_h$ and approached unity in hyperpolarised conditions, given the time constants τ^-_h and τ^+_h (Eq. 5 and Eq. 6).

$$I_T = g_T m_\infty h(V - V_T) \qquad \textit{Eq. 5}$$

$$dh/dt = \begin{cases} \dfrac{-h}{\tau_h^-} & (V > V\_h) \\ \dfrac{1-h}{\tau_h^+} & (V < V\_h) \end{cases} \qquad \textit{Eq. 6}$$

Meanwhile, the input current, I_{in}, was dependent on the excitatory current, g_E, which was additive, and the inhibitory current, g_I, which was subtractive (Eq. 7). The summation over the multiplication of the connectivity matrix, W_{ij}, with a dichotomous spiking vector, s_j, simulated spiking neurons (Eq. 8). The connectivity matrix, W_{ij},

describes both the binary state of connectivity between a neuron and another, either within a population or in another population, and the weight of that connection (Table 10 and Table 11). For example, if neuron *i* is a neuron in the TRN population that connected to neuron *j*, a neuron in the SP population, the weight of its connection is 0.35.

$$I_{in} = g_E(V - V_E) - g_I(V - V_I) \quad \textit{Eq. 7}$$

$$\frac{dg_{E/I}}{dt} = \frac{-g_{E/I}}{\tau_{E/I}} + K_{E/I} \sum_{j=1}^{N} W_{ij}\, s_j \quad \textit{Eq. 8}$$

$$\frac{dv}{dt} = 0.04v^2 + 5v + 140 - u + I \quad \textit{Eq. 9}$$

$$\frac{du}{dt} = a(bv - u) \quad \textit{Eq. 10}$$

Once a thalamic neuron reached the threshold V_o, it was considered to have spiked and its membrane potential was immediately reset to V_{reset}. All thalamic neuron parameters can be found in Table 12. Concurrently, the membrane potential, *v*, and state variables, *u*, of the sensory cortex's integrate-and-fire neurons are updated according to Eq. 9 and Eq. 10.

They were considered to spike when the individual cell's membrane potential, *v*, reached a threshold of 30 mV. After a spike is emitted, *v* was reset to *c* and *u* was set to (*u* + *d*), thus capturing post-spike repolarisation. The parameters *a*, *b*, *c*, and *d* were chosen to produce behaviour similar to regular spiking neurons and were described in Table 13. The parameters for thalamic neurons were tuned to elicit spiking behaviour that agreed with electroencephalography data from healthy subjects, such as the dominant α-peaks, and to ensure that the thalamic nuclei at resting-state fired within physiological ranges while the SP still received predominantly from peripheral input. Without any influence from other, higher-cognitive, regions, purely bottom-up driven cortical spiking reflected the input stimuli.

Table 12. *Thalamic neuron parameters.* These parameters were used in all simulations and are based on Smith et al. (2000) and Proske et al. (2011).

Parameter	SP	NSP	TRN
Size	100	100	100
k_E	1	1	1
k_I	0.1	0.1	0.1
V_o (mV)	-35	-35	-35
V_{reset} (mV)	-50	-50	-50
V_L (mV)	-65	-65	-65
V_E (mV)	0	0	0
V_I (mV)	-85	-85	-85
V_T (mV)	-66	-66	-64
C (μF/cm^2)	2	2	2
g_L (mS/cm^2)	0.035	0.035	0.035
g_T (mS/cm^2)	0.07	0.07	0.07
τ_h^- (ms)	20	20	40
τ_h^+ (ms)	100	100	100

Table 13. *Cortical neuron parameters.* These parameters were used in all simulations and model regular spiking neurons as recommended by Izhikevich (2003).

Parameter	Excitatory	Inhibitory
Size	800	200
a (ms)	0.02	$0.02 + 0.08r$
b (ms)	0.2	$0.25–0.05r$
c (mV)	$–65+15r^2$	–65
d (mV/ms)	$8–6r^2$	2

r=random variable, uniformly distributed between 0–1 for each neuron

c) Mean-field model

Two decision pools based on the standard mean-field approximation of integrate-and-fire neurons were used in both tasks (Wilson and Cowan, 1972). Their activity is governed by Eq. 11.

$$\tau \frac{dU_j(t)}{dt} = -U_j(t) + f\left(\omega_{j+}\lambda + \lambda_j + \lambda_{CL} + \omega_+ U_j - \omega_- U_i\right) + \sigma\xi(t) \qquad \textit{Eq. 11}$$

Where U_j is the average firing rate of the pool j, ω are the weights of the connections, all λ are the external inputs to the network, except for λ_{CL} which is the additional activation due to cognitive load, and the sigmoidal function $f(.)$ is described by Eq. 12.

$$f(x) = \frac{F_{max}}{1 + e^{\frac{-(x-\theta)}{k}}} \qquad \textit{Eq. 12}$$

Where F_{max} is the saturation point of the network, regardless of external input. In the simulations, apart from λ_j, which is given by Eq. 3, the constants are as follows: F_{max}=1.5 spikes/s, τ =20 ms, ω_{-}= ω_{+}=1, ω_{j+}=0, λ =0, σ=0.1 spikes/s, k=0.4, θ=4.44 spikes/s, dt=0.1 s.

d) Experimental configuration

To test the model, we first employed a classical target detection task where participants were asked to report stimuli appearing within a defined region of interest (ROI), and ignore those appearing outside of it (distractors). In terms of performance, we considered detection successful when the activity of the target-specific neurons exceeded that of neurons driven by distractors. As the neuron populations in this task were spatially organised, injecting a constant peripheral current to subgroups of neurons was equivalent to a stimulus appearing in their receptive fields. Both the target and the distractors could take on various sizes in terms of the number of neurons' receptive fields that it spanned (e.g. Figure 13-C) and specifics for each simulation can be found in Table 14 as well as in the description of the analyses carried out. A single timestep in the simulation was 1 ms and the ROI corresponded to the first 25% of the input space.

In this task, a goal-oriented strategy would be to excite all the thalamic reticular neurons outside of the ROI such that distractors would be suppressed. The non-goal-oriented strategy was a reverse ricker wavelet around stimuli for surround-suppression (Figure 13-B), where neurons with receptive fields neighbouring the stimulus were excited as a non-linear function of their distance to the stimulus.

The second task used was an oddball detection task. The model was presented with a black line rotating at a constant rate (4 different orientations to complete a full cycle) on a white background (Figure 15-A, top row). The changes occurred at 1.3 Hz, with the presentation of each orientation lasting 750 ms. An oddball stimulus was one that was inconsistent with the established rate of change; this is referred to as a jump. For example, if rotation was clockwise at a rate of one orientation per frame and the last orientation was 90 °, a jump would

be a change to either 0 ° or 135 °. Each trial lasted 23 cycles of the 4 orientations and had 10 jumps, leading to an experimental duration of 69 s. Oddball trials were pseudorandomised to ensure that they do not occur in the first two changes of the trial and had at least two regular transitions between them to re-establish the rate of change.

Table 14. *Simulation parameters.* If the parameter is a range of values, it is summarised as R(start, stop, step size). Both start and stop values are inclusive. RN: Reinitialised neurons after each trial; Size: distractor size in terms of the number of cortical neurons' receptive field that it spans; PFC: prefrontal cortex; TRN: thalamic reticular nucleus.

Task	RN	Time (ms)	Size (cortical neurons)	Noise	Cognitive load	PFC →TRN	Figure
1D	Yes	500	R(0, 400,80)	0	0	0.05	Figure ***14***-B
1D	Yes	500	320	R(0, .9, .1)	0	R(0, .024, .004)	Figure ***14***-C
1D	Yes	1000	R(80, 400, 80)	0	R(0, 1, .2)	0.03	Figure ***14***-D
1D	Yes	500	320	R(0, .9, .1)	R(0, 1, .2)	0.05	Figure ***14***-E
1D	No	500	320	0	0	R(0, .018, .003)	Figure ***16***-B, Figure ***16***-C
2D	Yes	750	N/A	0	R(0, 1, .1)	0.005	Figure ***15***-C, Figure ***15***-D

One timestep in the simulation was 1 ms. As with the target detection task, the cortical population peak, after filtering, was indicative of the model's selection, but in this case, it was an orientation rather than a spatial location. The task required the detection of orientations that did not follow the rate of change. A goal-oriented method of inhibiting thalamic input would thus be to inhibit the neurons sensitive to orientations that were consistent with the rate of change. Conversely, a non-goal-oriented method would be to suppress neighbouring orientations for stable representation of the current orientation.

4.2 Results

a) Performance in attentional tasks

1-Dimensional target detection

In our analyses, we used the spikes of each neuron summed over time to determine the local population response to stimuli. A stimulus would accumulate spikes in the local population that encodes for its position, leading to a local peak value approximating the stimulus' location. The data was passed through a forward-backward low-pass filter with a sampling frequency of 800 Hz and a cut-off frequency of 6 Hz to obtain a population signal. The population peak is considered the stimulus that the model 'attends' to, reflecting the principle that an attended stimulus is indicated by more neural activity than unattended stimuli. Figure 14-A illustrates the excitatory cortical response to two presented stimuli over time for a representative trial. The highlighted areas in the raster plot are the receptive fields where stimuli were presented to the model, the number of covered neurons being the size, or width, of the stimulus. The areas span the entire duration of the trial as the stimuli positions were held constant over time. The number of peaks in the local population response (Figure 14-A, right) echoes the number of stimuli in the input and shows that the stimulus centred at neuron position 600 occupied a larger receptive field than the other stimulus at neuron position 180. Generally, the difference in peak magnitudes between two stimuli increased with their size difference.

As the prefrontal cortex excited both the non-specific nucleus and the thalamic reticular nucleus, it not only increased the spiking frequency for stimuli within an ROI but also indirectly inhibited task-irrelevant regions. The activity in the sensory cortex was thus modulated by interference from the prefrontal cortex such that sensitivity was reduced outside of this ROI but increased within it. This allowed top-down influence to modulate cortical response to favour target stimuli even in the presence of distractors that have more bottom-up saliency.

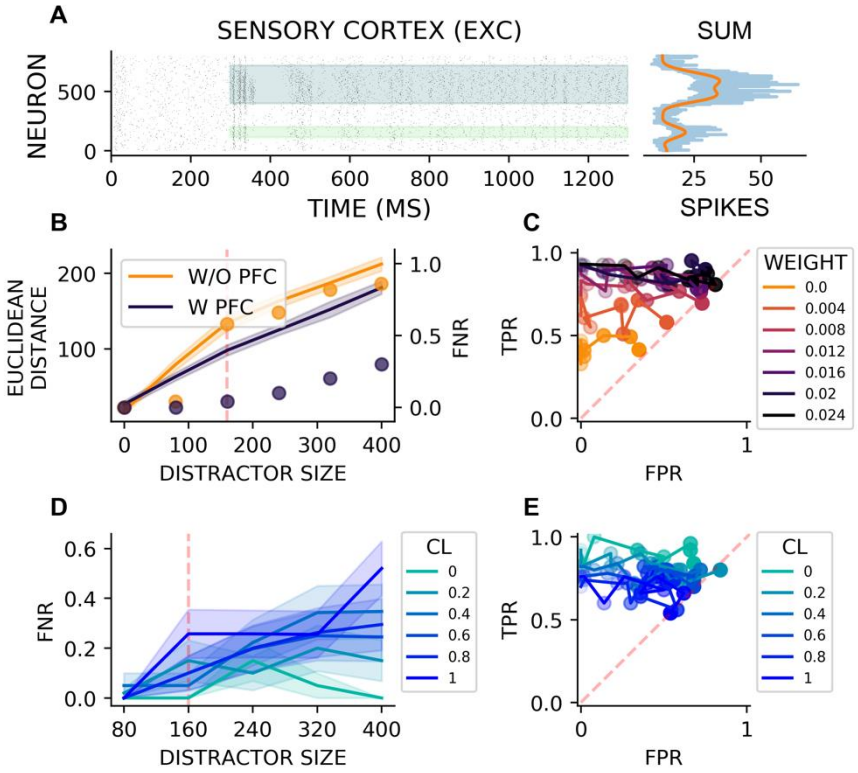


Figure 14. *Performance and neuronal responses in a target detection task.* A) An example of the sensory cortical neurons' response to external stimulation. On the left is a raster plot of each neuron spiking. The highlighted regions indicate when the peripheral stimulation is active and the sensory cortical neurons which are indirectly excited by the stimulation. On the right, the total number of spikes over the stimulation duration per neuron (blue), and its smoothed signal (orange). B) The solid lines are the mean Euclidean distance from the ideal sensory cortical response, when no distractor is present, with the shaded area being their standard deviation. The points are the corresponding false negative rates (FNR), with a slight, random offset along the x-axis so they do not overlap. C) The ROC curve obtained by varying environmental noise and PFC→TRN connection weight. Increase in noise does not modulate much the TPR, but does increase the FPR such that it approaches chance-level. The greater the weight of the PFC→TRN connection, the better the model is able to perform the task as shown by an improvement in TPR. D) The mean FNRs, each of 30 models, are shown as solid lines while their standard errors are the shaded regions. There is a trend of a higher FNR with greater CL, but there is no clear separation until the distractor size is at least double the target. E) The more opaque the point, the greater the environmental noise presented during simulations. Again, noise affects the FPR but not particularly the TPR. There is a slight trend of decreasing TPR with increased CL.

To quantify this effect, the network was exposed to two static stimuli for 50 trials. The distractor stimulus could be one of five different widths, from 80 to 400 cortical neurons, while the target stimulus, situated in the ROI, was always 160-neurons wide. The distractor

sizes simulated were restricted by the available input space, after including gaps, each equivalent to the receptive field of at least 80 sensory cortical neurons, at the edges and between both stimuli. A synaptic weight of 0.05 was chosen for the PFC→TRN connection as it was found to consistently perform the task. The mean sensory cortical filtered output to a target in the absence of distracting stimuli was used as the 'ideal' response of the network. The network's error was measured as the Euclidean distance of each trial's cortical filtered output to this ideal response. The resultant means are plotted as solid lines in Figure 14-B, while the shaded regions are their standard deviations. Distractor size is measured by the number of sensory cortical neurons it stimulated indirectly.

At the same time, the corresponding false negative rate (FNR), the probability of the sensory cortical signal producing a population peak at the distractor's position when there is a target, for each condition was calculated. The false negative rate was a direct measurement of the model's ability to perform the task, while the Euclidean distance measured how aggressively the model suppresses the distractors. We observed that the model managed to maintain a low false negative rate for all tested distractor sizes, always performing better than chance although the performance depended on the size of the distractor.

That the Euclidean distance to the ideal output increased with distractor size even with prefrontal cortex's influence reflects how the activity representing the distractor is reduced sufficiently to allow the target to win in the competition between stimuli, but is not completely silenced. Nonetheless, the prefrontal cortex's inhibitive influence is weakened by increasing bottom-up saliency of the distractor as the false negative rate approached chance level for the largest distractor sizes.

To better understand how the weight of the PFC→TRN connection affected performance and robustness to environmental noise, simulations were run on a range of values for noise and weight (Table 14) when there were both a target and a distractor present as well as when there was only a distractor. Each condition was run 50 times, with the model reinitialised each time. Noise was the probability that a neuron was excited by an externally generated input regardless of the actual peripheral input, within a range of 0 to 1. When the

probability was 0, the network received input that is purely due to the peripheral input; at a probability of 0.9, each neuron almost certainly was excited by a signal indistinguishable from the peripheral input. To be consistent with how the peripheral input was presented to the network, the noise pattern was static over time, although it was randomly re-generated at the start of each trial. For each trial, there were two ways to make an accurate classification:

- Attending to a stimulus in the ROI when there was indeed a target stimulus, a true positive
- Attending to a stimulus outside of the ROI when there was no target stimulus, a true negative

The resultant classification over multiple trials by the model on the presence of a stimulus in the ROI was recorded and the false (FPR) and true positive rates (TPR) were calculated using the standard equations Eq. 13 and Eq. 14.

$$TPR = \frac{true\ positives}{true\ positives\ +\ false\ negatives} \qquad \textit{Eq. 13}$$

$$FPR = \frac{false\ positives}{false\ positives\ +\ true\ negatives} \qquad \textit{Eq. 14}$$

These are plotted as the receiver operating characteristic (ROC) curve in Figure 14-C. Each line represents the performance of the model with a specific PFC→TRN synaptic weight, and a greater opacity of the dot represents a larger noise probability. As the probability of noise increased (0 to 1 with 0.1 increments), performance dropped almost to chance level (closer to the dotted identity line). At lower connection weights, the network barely performed the task, reaching true positive rates of chance-level or below. This was because the lower the weight of the connection, the more it relied on bottom-up representations and performed with less 'awareness' of its task.

One of the functions of selective attention was to optimally direct cognitive resources to relevant stimuli due to a mismatch in number of possibly attended stimuli and available resources to process them. Cognitive load (CL) was thus introduced as additional activity in the decision-making mean-field model (λ_{CL} in Eq. 11), affecting both decision pools equally and could range from 0 to 1. 30 models were initialised per condition (distractor size vs cognitive load) and presented with the same target and distractor combination 20 times. A weight of 0.03 was chosen for the PFC→TRN connection to

remain in the region of consistent behaviour relative to the task as shown earlier in Figure 14-C.

The distractor size was described as the number of sensory cortical neurons whose receptive fields are indirectly stimulated by the distractor. Each presentation lasted 1s and there was a 100ms interval between presentations, allowing for a false negative rate per model per condition. This provided mean false negative rates for all the models across conditions, which are shown as solid lines in Figure 14-D while their standard errors are the respective shaded areas. The dashed red line shows the size of the target used in all the simulations. Without any cognitive load, the model barely missed the presence of the target regardless of distractor size, maintaining a false negative rate of close to 0. The false negative rate increased with cognitive load, although the separation in false negative rate between conditions with some level of cognitive load and when there is absolutely no cognitive load only becomes distinct when the distractor was at least double the size of the target (distractor $\geq$ 320 neurons).

To test the model's response to cognitive load in the presence of environmental noise, we used a protocol similar to Figure 14-C, this time across a range of noise and cognitive loads instead of noise and connection weights. Again, the opacity of the point is indicative of the noise probability, which can be found in Table 14. The weight of the PFC→TRN connection was maintained at 0.05 as it was capable of consistently carrying out the task. As with connection weight, an increase in noise probability increased the false positive rate without affecting the true positive rate much, which remained relatively high regardless of cognitive load level Figure 14-E.

The results therefore suggest that noise made overall discrimination of the stimuli less reliable—either in identifying their positions or in recognising spaces between the stimuli—which was reflected by a marked increase in false positive rate. Cognitive load, on the other hand, reduced the model's ability to actively inhibit the distractor leading to an increase in false negative rate. This was further emphasised by increasing noise levels leading to a decrease in trough magnitude, as a percentage of the global peak magnitude (linear regression: $p<<0.05$, $r=-.75$). In other words, the dip in spikes

between the stimuli beame less distinct with more noise. This effect is not found for increasing CL (linear regression: p=0.96, r=-0.0031).

2-Dimensional oddball detection

The model was extended to a dynamic, 2D stimulus to demonstrate its functionality in a more complex scenario. The same concept of global peaks to determine stimulus selection was now used to select one of four orientations instead of a spatial location (Figure 15-A). The stimulus elicits spikes in the thalamic nuclei as well as the sensory cortical population at every time step. To obtain a classification of the stimulus' orientation at a time point, the sensory cortical activity was summed over a window of the most recent activity and the filtered signal provided a global peak from which an orientation could be decoded. As a buffer to noise within the system, which was separate from environmental noise and cognitive load, the 'reported' orientation from the network was the mode of the decoded orientations across a time window that ended at the current timestep.

To determine the appropriate range of window lengths, thirty simulations were run where the stimulus' orientation is fixed, each 750 ms long. A sliding window of varying lengths (Table 14) was applied across each simulation and the success rate of decoding the orientation was recorded per window per simulation. The success rates were shown in Figure 15-B, where a longer window usually allowed for more stable reporting of stimulus orientation. There were a handful of simulations where decoding success did not improve much regardless of the time window used, which may be the outcome of both a sensitivity to initialised parameters as well as the relatively small size of the simulated populations. The window length selected for subsequent simulations is 300 ms, as it was well beyond the asymptote obtained and was within the range of reaction times to changes of orientations in humans (Gilinsky and Cohen, 1972).

The network was exposed to the dynamic stimulus and was able to anticipate the orientation of the next change based on the current and previous orientations. A jump occurred when the change was inconsistent with the established rate of change (i.e. was not anticipated). A change was detected when there was a difference between the stimulus of the previous and the current timestep. To determine if a jump had occurred, the model compared the

anticipated orientation and the mode orientation, from the moment of a change to 350 ms later. Cognitive load was included in the model via additional activity in the decision pools of the prefrontal cortex, similar to the implementation in the target detection task.

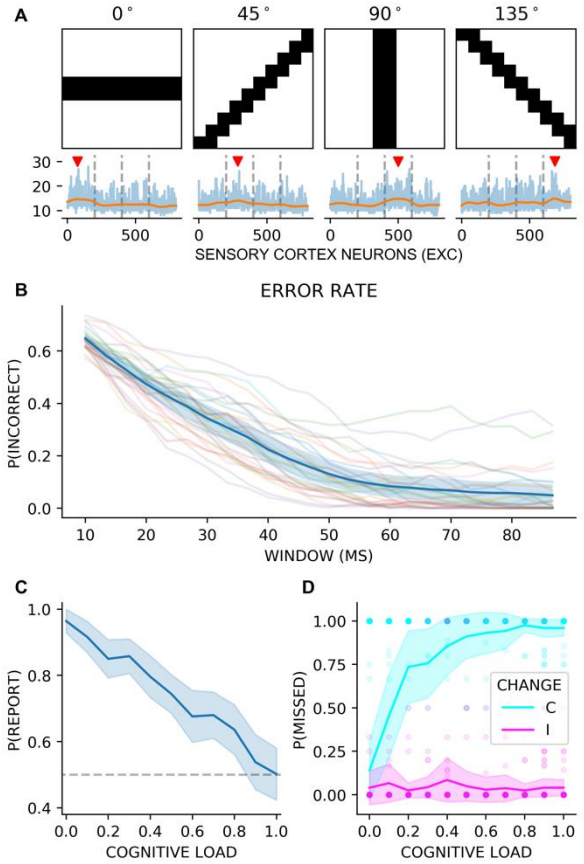


Figure 15. *Performance in a 2D oddball task.* A) The upper row shows the retinal input to the SP and NSP, while the lower row shows examples of the corresponding output of the cortical layer (excitatory). The red arrows show the positions of the population peaks, demonstrating the orientation selectivity of each quarter of the cortical population. B) 30 simulations were run with a static input orientation. A sliding window of varying lengths is applied to the simulation duration (750 ms) and tested for accuracy of representing the veridical stimulus orientation. This provides a rate of decoding failure for each window length per simulation, which is shown as translucent lines. Their mean is plotted as a solid line and the shaded region is their standard deviation. C) The performance of the model is impaired by increasing CL, reporting fewer jumps in the dynamic stimulus as it is subjected to greater CL. D) Each point, coloured to match its type (congruent/incongruent), is the detection rate of each trial while the solid lines are their means and the shaded regions their standard deviation. The decrease in number of jumps detected with increasing CL is due to a decrease in detection of jumps in a direction congruent to the stimulus' rotation (blue) and not the incongruent ones (pink). As there are many overlapping values, individual trials' detection ratios are plotted with a slight, random jitter (±0.125 in both x- and y-axes).

As cognitive load increased, the probability of anticipating two orientations instead of just one also increased as a coping mechanism for reduced cognitive resources to monitor state spaces. The anticipated orientations were always in the same direction as the established rotation.

The model detected almost all jumps when it had no cognitive load, performing the task well above chance level Figure 15-C. However, the rate of jump detection dropped to chance-level at the highest-tested level of cognitive load. This decrease in performance was further analysed by splitting the missed jumps into two types: when the jump direction was congruent with the rotation of the stimulus and when it was incongruent. The model could detect incongruent jumps regardless of the cognitive load condition Figure 15-D (pink), but it failed to detect congruent jumps more in high cognitive load conditions than low Figure 15-D (blue).

This demonstrated how, when the load on cognitive resources was increased, sensitivity to errors was lowered in a manner that is specific to the task-related anticipatory inhibition. The model's change in performance followed the same trend as behavioural outcomes from psychophysical studies with human participants (e.g. Malekshahi et al., 2016). It is important to note that the ratio is greatly affected by the sheer increase in number of missed jumps with increasing cognitive load, as there were almost no missed jumps in lower cognitive load conditions while up to 50% of jumps were missed at the upper limits of cognitive load.

b) Physiological benchmark

We next benchmarked the effect of the prefrontal cortex’s influence on the thalamic neurons' spiking pattern, which is directly addressed by Mease et al. (2014). The authors observed the thalamic response to optical stimulation of Layer 6 neurons in the barrel cortex (L6BC) of optogenetically stimulated mice while their whiskers were physically stimulated. To compare their findings with data from the model, the optically stimulated neurons in the mouse barrel cortex was loosely likened to neurons in the model's prefrontal cortex. Additionally, the associated thalamic projections are in the mouse ventromedial nucleus, which in this case was contrasted with the

model's non-specific thalamic nucleus. The physical stimulation to the mouse whiskers was analogous to the peripheral input into the model's thalamus.

A pertinent finding was how thalamic sensory responses changed with optogenetic stimulation of the L6BC. Multiple trials were run with a range of PFC→TRN weights without reinitialising the neurons so that the spiking mode of individual neurons could be compared across different connection weights Figure 16-A. To mimic the on/off nature of the optogenetic protocol from Mease et al. (2014), the simpler of the two previously described tasks (target detection) was used. Only the activity from the neurons representing the distractor stimulus was considered as they are the neurons targeted for inhibition by the prefrontal cortex. The sizes of the target and distractor stimuli were always 160- and 320-cortical neurons wide respectively. The same peripheral input was injected into the network repeatedly for seven PFC→TRN weight values, each trial lasting 500 ms, matching the stimulation duration of Mease et al. (2014), with 100 ms breaks between them.

The simulated output of the non-specific thalamic nucleus showed a decrease in spike counts in the non-specific thalamic nucleus when the prefrontal cortex was activated (Figure 16-B), reflecting what was found by Mease et al. (2014) (Figure 2C and Figure 2D in their paper). Furthermore, as the weight was increased incrementally, we found that the effect was gradual. Although the authors detected much fewer spikes than was generated in the model, the number of spikes we obtained was within the expected range for human regular-spiking neurons with the defined cell membrane potentials, which can reach bursts of up to 300–400Hz (Destexhe et al., 1998).

In addition to changes in spiking frequency, the mouse thalamic neurons also shifted to an overall tonic mode when the barrel cortex was stimulated than when it was not. The extent of this change in individual neurons was dependent on its initial firing mode, with burstier neurons at baseline having more pronounced shifts. The tonic and burst firing modes were distinguished using an inter-spike interval criterion, where any spike preceded by an inter-spike interval of less than 10 ms was considered part of a burst response. The burst probability of a neuron was the proportion of trials in which there was a response to the stimulation where a burst response was elicited.

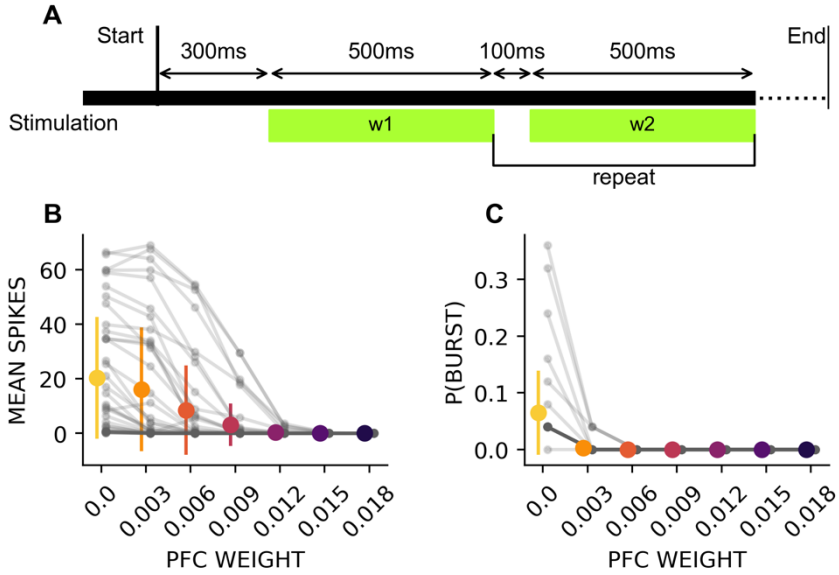


Figure 16. *Thalamic response to increasing PFC weight.* A) Protocol used to obtain spikes from the non-specific thalamic nucleus responding to the distractor. A single PFC→TRN weight is used for each trial (e.g. *w1*, *w2*) and all the spikes during stimulation are recorded for analysis. B and C) The spiking pattern of non-specific thalamic neurons receptive to the distractor's position is modulated by increased weight of the PFC→TRN connection. The grey lines show the changes for single neurons while the coloured error bars offset to their left represent their means and standard deviations.

Mease et al. (2014) had used a threshold of 10 ms, much longer than the commonly used 4 ms, as they had not observed many neurons that met that requirement. However, the membrane potentials used in the neuron model were tuned to be slightly hyperpolarised for observation of alpha oscillations, encouraging bursty behaviour. The model therefore produced more spikes than were found in mice, and 4 ms was used as the threshold for classification of neurons in the simulated data. Even with the stricter requirements, we found that the model experienced a shift in the same direction as was found in the mouse (Figure 16-C).

As in the case of the spike count, the change of the burst probability with increasing weight was not abrupt although it converged to 0 at a lower weight than for spike count. When a threshold of 10 ms was used, the convergence occurred closer to 0.015. The effect of

prefrontal cortical influence was greatest in neurons that were burstier to start with, which was also found by Mease et al. (2014).

The results showed that when the prefrontal cortex is inhibiting the non-specific thalamic nucleus through the thalamic reticular nucleus, the inhibited thalamic neurons fire more tonically and spike less often than when the network is without prefrontal cortical influence. Such behaviour was consistent with that found in mice, giving support to the plausibility of this mechanism as a part of the attentional system in mammals in general, and possibly in humans specifically.

4.3 Conclusions

We proposed a model of attention that hinged on the inhibition of task-irrelevant features rather than the excitation of task-relevant ones. It was anatomically grounded and consisted of spiking neuron populations for most of the modelled brain regions, except the prefrontal cortex which was a mean-field approximation model. Using the summed output of the model's sensory cortical population as the model's response to stimuli, we tested its performance and resistance to interference from environmental noise as well as cognitive load. The model could carry out both tasks, and performed best when there was neither environmental noise nor cognitive load.

By affecting the decision-making mechanism rather than the perceptual stream, we showed that cognitive load led to changes in performance that were similar to those observed in psychophysical studies. The network's spiking patterns and output also concurred with physiological experiments conducted on mice as well as behavioural measures from humans, further demonstrating the biological plausibility of such a mechanism. It is important to note that the model's weight of the PFC→TRN not only controlled the amount of influence of the prefrontal cortex on the TRN, but also what would be considered how committed a person is in performing the task. This is because the model performed as a purely bottom-up, reactive system when there is no top-down influence. In this sense, the performance at low PFC→TRN weights (Figure 14-C) does not purely reflect poor performance in the task but also a disregard of the task.

The results demonstrated how the proposed model is a possible mechanism of selective attention in humans in situations where it is unpredictable stimulus features that are task-relevant. This contrasts with most existing models, which mostly focus on when predictable features are task-relevant, and thus offers a complementary mechanism for allocating sensory processing resources. In addition, the typical stimuli used in modelled experimental paradigms tend to be static, such as still images. The catch detection task in this study showed how the model is applicable even when changes over time are inherent to the task. The model also used the modulation of the prefrontal cortex's state to achieve its function, rather than more permanent, and potentially costlier, mechanisms like changes in weights (synaptic changes) or neuron parameters (neuronal characteristics).

The interactions modelled here are limited compared to more general models of brain functions and attention like (Grossberg, 2013; Nyamsuren and Taatgen, 2013). For one, while the model was biologically-grounded, we cannot definitively claim to show that cortical influence on the thalamus would cause a switch to a tonic firing mode, which was observed in the mouse thalamus. This is because the inter-spike interval is not the only way to quantify a neuron's burstiness. It was, however, the clearest measure available to us as the modelled neurons, once initialised, have fixed parameters that, in a real neuron, would change over time, such as the reset membrane potential of a neuron post-spike determining if it was hyper- or depolarised. Another shortcoming that arose from this same limitation was the inability to show thalamic adaptation, which Mease et al. (2014) found to decrease thalamic responsiveness over the duration of each stimulation. Nonetheless, adaptation and learning were beyond the scope of this model and avoiding the re-initialisation of neurons between conditions was sufficient to examine how a single neuron modified its behaviour according to its interactions with the other modules.

In addition, there was evidence that driving the TRN would lead to an increase in thalamocortical bursts instead of a decrease (Halassa et al., 2011), which could be due to protocol differences. In the case of both this study and Mease et al. (2014), a critical element was stimulation of the sensory network and the specific response to it; as for Halassa et al. (2011), the sensory stimulation was non-specific

and relatively uncontrolled. The current model focused on the mechanism with which top-down processes can bias selective attention towards task-relevant stimuli, and did not venture into how they determine what comprises task-relevance. Although there have been attempts to do so, like applying information theoretic formalisations of maximising information to visual search (Bruce and Tsotsos, 2010), it remains a relevant topic of research.

Future experiments can look into how different kinds of non-task-related noise affect the attention network. The only type of noise that was simulated was one that is externally generated. However, there is noise that is also internally generated, in addition to the interaction of attention on other cognitive functions, such as working memory. How these perturbations to the network affect its performance, and how the network buffers against them, is still unclear. In addition, the rules with which the prefrontal cortex modulates activity in the thalamus was hard-coded. These predictive codes could instead be learned over time, for example with reinforcement learning algorithms. The resultant 'solution(s)' for the connection pattern from the prefrontal cortex to the TRN which balances false positives and false negatives would be of great interest. The interactions between the prefrontal cortex and other subcortical regions with significant impact on selective attention, such as the amygdala (John et al., 2016) or the superior colliculus (Herman et al., 2018), can also be further detailed.

5. COMPLEMENTARY INTERACTIONS BETWEEN CLASSICAL AND TOP-DOWN DRIVEN INHIBITORY MECHANISMS OF ATTENTION

This thesis has discussed how selective attention can be implemented to prioritise task-relevant stimuli over others, as the alternative, that is to process all sensory input, is not feasible in most animals due to limited processing capabilities. Such a necessity also applies to artificial agents operating in all but the most simplistic environments, for the most simplistic purposes. The ability to prioritise stimuli is, however, not the end goal itself, but rather a means to several ends (Allport, 2016). For an agent interacting with its environment, it is ultimately behaviour that leads to survival or success in its goals. Behaviours are often associated with specific states in the sensory input or states of the agent. In the case of stimuli-driven responses, biasing sensory processing to task-relevant stimuli therefore also biases their elicited behaviours to those that advance the agent's goals and objectives. Therefore, selective attention serves several behavioural purposes; one of these is to rapidly detect and/or orient the agent towards stimuli of interest (Braun and Julesz, 1998), presumably so that action related to the object can be taken. For this reason, modelling eye movements, which measures literally the orienting of eyes, is a popular benchmark for attentional models (e.g. Adeli and Zelinsky, 2018; Marat et al., 2009; Mirza et al., 2019). Another common method to test the performance of attention models is implementing it in artificial agents, following the principle that 'one truly understands only what one can create', as eloquently summarised by Giambattista Vico.

There are two advantages to implementing models of attention in artificial agents. Firstly, and many a times this alone justifies the effort, an agent equipped with some form of selective attention often interacts better with its environment than one without, with this advantage increasing in value with environment complexity and sensor resolution. Secondly, taking a more ontological perspective, it considers the model not simply as an algorithm solving an abstract problem but something that enables a body to function, which is a key claim of embodied cognition (Anderson, 2003; Engel et al., 2013;

Wilson, 2002) and something many machine learning studies ignore (Moulin-Frier et al., 2017; Puigbò et al., 2018). Thus, there is a considerable number of implementations of selective attention in artificial agents or robots. Itti et al.'s (1998) saliency-based model of bottom-up attention has inspired many other models, from a pyramidal feature-based selection of spatial location in the visual input that most closely matches a target (Driscoll et al., 1998) to a model that uses a similar bottom-up strategy to identify landmarks as it navigates so that it is able to self-localise (Ouerhani and Hügli, 2005). Frintrop et al. (2003) added depth to the range of sensors, allowing a robot to identify more points of interest for navigation and object detection.

While deriving saliency maps from early-stage, bottom-up processes remains relevant (e.g. Dollar et al., 2014), there have also been attempts to include top-down elements to facilitate more sophisticated behaviours. One such early study that brought the two processes together oriented the gaze of a robotic agent to objects in the environment depending on its active internal drive, the dynamics of which was driven by a combination of perceptual input and internal motivational systems (Breazeal and Scassellati, 1999). Another model used interaction cues with a human partner to select salient objects in its environment, with the goal of learning and using labels for them (Haasch et al., 2005). Colombini et al. (2016) stored individual sensor information so that bottom-up attention also considers historical information and is able to engage inhibition of return, in addition to top-down enhancement of regions in the attentional map. Another study applied biologically sound principles to a functional model of attention in a robot for learning, demonstrating the effects of certain biological limitations in visual sampling such as foveation, where only the region of visual input covered by the fovea is focused and sharp. It also addressed the related indirectness problem, which refers to the fact that foveation might not always occur where effectors are causing learning feedback (Ognibene and Baldassare, 2015).

As exemplified by these studies, while varied in terms of architecture and hardware, the typical role that is given to top-down attention is

excitatory. The success of robotic implementations of selective attention, such as those cited above, supports not only its functional importance but also the physiological experimental work that has grown alongside this field (e.g. Cutrone et al., 2014; Desimone et al., 1998; Sani et al., 2017). Nonetheless, the previously described tendency of task paradigms in computational and experimental neuroscience, where they revolve around task-relevant stimuli being more predictable than task-irrelevant ones, persists in robotics and computer vision as well. This has led to the development of robotic attention models that, again, favour top-down excitation. For example, using Tsotsos' (2011) taxonomy of three classes of attentional mechanisms, suppression, selection and restriction, out of the 52 reviewed cognitive architectures, only 5 included the suppression of task-irrelevant stimuli, space or features while all of them exhibited at least one element of selection and restriction (Kotseruba and Tsotsos, 2020).

As mentioned in earlier chapters, neglecting the inhibitory top-down pathways of attention is to deny that there may be situations where the excitatory mechanism is not the most efficient. Two intuitive scenarios for this would be 1) when bottom-up activity due to distractors reach much higher levels than that from task-relevant stimuli and 2) when what is task-relevant is not precisely known and is instead defined by what is task-irrelevant (e.g. find the non-red object). There is strong support for both mechanisms in the mammalian brain, which has already been discussed in detail earlier in this thesis, and they appear to have strengths and weaknesses that are complementary rather than competitive. We therefore investigated how these two mechanisms, both separately and combined, drove the behaviour of a simulated, artificial, autonomous agent performing a controlled, foraging task. It was placed in an environment populated with two types of objects—rewards and distractors—distinguishable by their colours. By approaching and touching an object, the agent effectively demonstrated that its control dynamics had selected that object over other objects in the vicinity. In this study, the agent's control architecture was based on the biologically grounded Distributed Adaptive Control (DAC; Figure 17-right; Pfeifer and Verschure, 1992; Verschure, 2016), which has

been shown to robustly learn sequences of state-action couplets that eventually lead to a goal state (Marcos et al., 2014).

We first proved the utility of top-down attentional biasing on bottom-up sensory competition, either through excitatory or inhibitory mechanisms, in the simplest possible setting: two objects, one a reward and the other a distractor, at equal distances from the agent. We found that regardless of the exact mechanism, any top-down biasing led to a distinct increase in probability of selecting the reward when compared to conditions where there were only bottom-up processes. Nonetheless, in natural settings there are usually more than two stimuli competing for attention. In a second experiment, to assess the performance of the top-down mechanisms in a more ecological situation, we tested the agent's performance with more objects in its visual field and with different proportions of rewards and distractors among them. To determine the effect of the top-down biasing mechanisms on behaviour, the agent was tested with only bottom-up sensory competition, with either of the two top-down mechanisms in isolation, and with the two top-down mechanisms in combination. In total, this led to four conditions. We showed that although there remained certain limitations, the best performance across the widest range of target-distractor configurations is achieved when both excitatory and inhibitory mechanisms are in place.

5.1 Methods

To demonstrate the effects of top-down biasing, and their different mechanisms, on behaviour, we created an embodied agent (Figure 17) in a virtual arena performing a foraging task. The environment, physics engine and visualisation of the simulated arena were created using Python (Rossum, 1995) and the open-source libraries PyGame (Shinners et al., 2011) and Box2D (Catto, 2011). The simulated agent was a mobile, two-wheeled robot with seven sensors for proximity and vision each. Both types of sensors functioned by emitting a single projection at a fixed angle depending on its position on the agent and returning the distance and the colour of the first obstacle that it hits respectively (Figure 18-A). The objects used as rewards and

distractors are solid-coloured circles, with rewards bearing a value of 1 and distractors 0. The distractors were intentionally valued to be neutral instead of punishments (values<0) as it is important to make a distinction between a distraction and a punishment. Critically, a punishment is task-relevant, as it necessitates aversive behaviour, while a distraction itself has no benefit or loss and is therefore task-irrelevant. In natural environments, especially when they are cluttered, this classification is not always static as distractors can become punishments or rewards, for example due to movement from the agent or the objects. Nonetheless, this interpretation of distractors and rewards is sufficient for the purposes of the current study.

The top-down attentional mechanisms tested here were the excitatory model, conceptually congruent with most enhancement-based models of attention, and the inhibitory model, where activity representing task-irrelevant stimuli is suppressed. In both cases, the visual sensor that was targeted, and by extension the stimulus its activity represents, was dependent on the motor decision from the contextual layer (further elaborated in **Control architecture** below) in the previous time step. The key difference between the two top-down biasing mechanisms was that if the action was towards a reward, an excitatory signal was used whereas an action towards a distractor led to inhibition of the activity from the selected sensor for the next time step(s). This led to a biasing of subsequent sensory competition and, by extension, perceptual information that is next fed into the contextual layer. The strength of the top-down excitatory and top-down inhibitory connections were defined by their gains, G_{TDexc} and G_{TDinh} respectively (Figure 17). Thus, the top-down attentional biasing mechanism could be modulated though changing the values of these two gains.

a) Control architecture

As a cognitive architecture, DAC posits that cognition arises from the interaction between interconnected control loops operating at four increasing levels of abstraction for hierarchical control: soma, reactive, adaptive, and contextual. At the lowest level of abstraction,

the soma layer defines the body of the agent in its environment. The layer above it is the reactive control of behaviour, which associates certain sensory input with hard-wired behaviours. While an agent with only reactive behaviours is limited in its handling of sensory stimuli, it is still a necessary and important part of the control architecture (Nolfi, 2002). The reactive layer of DAC further extends other behaviour-based control architectures (e.g. Brooks, 1986) by using an allostatic process to mitigate drives that are modulated by the sensory input.

Behaviour generated by the reactive layer bootstraps learning of the agent's environment and the sensory outcomes of its actions on the environment, which takes place in the adaptive layer. This helps the agent learn behavioural responses to perceptual signals that are more noise resistant. In the adaptive layer, these pairs of perceived state, or prototype, and motor action are associated together into state-action couplets, which then are sent to the contextual layer that uses them for higher-level cognitive functions like goal selection and planning. Through Bayesian decision-making (Bayes, 1763; Beck et al., 2008), DAC's contextual layer captures a knowledge level description of intelligence and the principle of rationality by exploiting perceptual and behavioural learning through interaction with the environment (Verschure et al., 2003). This is supported by the contextual layer's short-term memory, which stores the received state-action couplets in a sequence until a goal state is reached, at which point the sequence is transferred to the contextual layer's long-term memory.

This implementation of long-term memory is modelled after how long-term memory is formed in animals, which was proposed to be through the accumulation of knowledge and abilities which is then optimised for subsequent planning and behaviour (Newell, 1990). The psychological theories of cumulative learning and scaffolding (Berk and Winsler, 1995; Boblett, 2012; Gagne, 1968) built on that proposal, suggesting that learning starts from skills with low complexity which are transformed hierarchically into increasingly complex ones for improved performance. In this study, the usage of such accumulated knowledge was manifested by the agent comparing its current state with those in its long-term memory to select valuable

states to be reached, anticipate the state-actions couplets that would bring it there, and trigger goal-oriented behaviours based on the couplets.

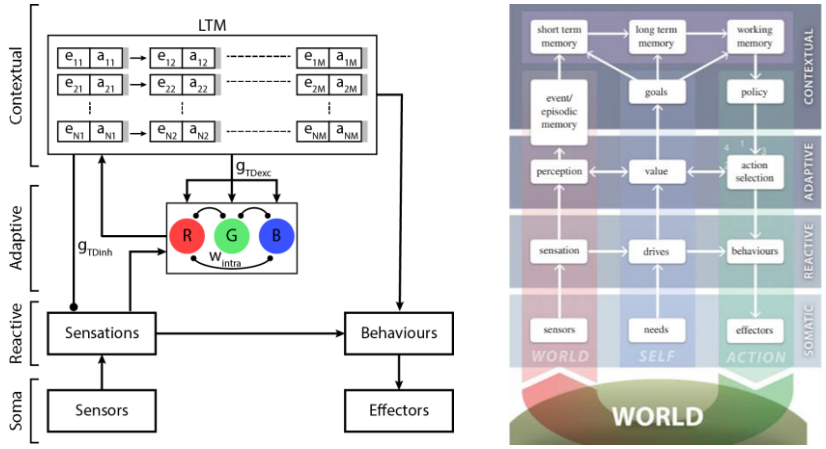


Figure 17. *Control architecture of the agent. (Left)* Lines terminating with arrowheads show excitatory connections while lines terminating with circles show inhibitory connections. The contextual layer decided on actions and change the gains of the top-down attentional biasing based on its long-term memory (LTM), which is shown here as a collection of sequences of state-action couplets. The visual channels are indicated by the nodes R, G, and B, and the connections which were modulated in the experiments are labelled G_{TDexc} and G_{TDinh}. The gain of the inhibition, representing bottom-up processes, is shown by W_{intra} and applied to all the inhibitory connections between the visual channels. (Right) The Distributed Adaptive Control architecture which the current control architecture is modelled after, reproduced with permission from (Verschure, 2016).

As the objective in this study was to demonstrate the effects of different attentional models on behavioural outcome, the agent's control architecture was designed to minimise extraneous complexity. The agent's behaviour was a combination of motor commands from the reactive and the contextual layer (Figure 17). The adaptive layer did not provide any direct motor control as we were interested in behavioural control through the contextual layer and creating that additional pathway for motor commands would create a confounding variable. The agent's reactive layer was hard-coded with a reflex to avoid colliding with walls, using the activation of proximity sensors at one side of the agent to generate a proportional motion in the wheel on the same side, effectively turning the agent away from the obstacle a la Braitenberg vehicles (Braitenberg, 1986). In addition, if none of the agent's sensors detect

sensory stimuli, the agent had a reflexive drive to randomly explore the environment. This ensured that the agent was always proactively interacting with the environment even when it was not perceiving objects.

Although the adaptive layer did not directly issue motor commands, it contributed to the architecture by integrating bottom-up and top-down attentional signals to produce the perceptual state that was used by the contextual layer. The perceptual information was driven by sensory input, which consisted of three colour channels per visual sensor. The intralayer competition observed in neuronal population by Reynolds et al. (1999) and others was simulated here with the sensors' activity contributing to the activity of each channel's own node. They then exhibited competition by inhibiting the other channels' nodes after their activity was weighted by W_{intra} (more details are available in the sections **Model formalisation** and **Bottom-up sensory competition**). After this interaction, the nodes were then decomposed into the different sensors that had contributed to them (saliency vector P in **Model formalisation**), and the modulated activity of the most active sensor was selected through a winner-takes-all mechanism as the perceptual input to the contextual layer. If there were no signals from the contextual layer, or if the gains of the top-down biasing connections were all zero, this would be equivalent to the winning stimulus from a purely bottom-up sensory competition.

However, in conditions with top-down biasing, the top-down gains (G_{TDexc} and G_{TDinh}) were not zero and the contextual layer's previous motor decision modulated activity of individual sensors prior to the calculation of the intralayer inhibition. The contextual layer matched the perceptual input to the states in the stored sequences of state-action couplets in the long-term memory. This way, it predicted the potential reward of following a sequence of motor actions and made decisions that maximised the rewards gained in the future (more details can be found in the sections **Model formalisation** and **Top-down biasing components**). In the same way, the contextual layer used the selected state-action couplet to modulate sensory competition in a goal-directed fashion, sending an inhibitory signal

if the selected state belonged to a distractor and sending an excitatory signal if it was a reward. This top-down biasing signal is applied with a specified gain to the sensory input to the adaptive layer in the next time step. As the learning process in the contextual layer has already been demonstrated (Duff et al., 2011; Marcos et al., 2014) and the scope of this study is to examine how the contextual layer can be used instead of how it develops, in this study we pre-trained the contextual layer and did not update it during experiments to control for experience between the attention models tested.

To train the contextual layer, an additional appetitive behaviour towards both red and green objects was included in the agent's reactive layer. There was no top-down biasing of sensory input during training, leading to output from the adaptive layer reflecting purely bottom-up processes. The agent was placed at a random position and with a random orientation in the arena along with a red or a green object. It explored the arena using its reactive exploration behaviour until it detected the object, and then its reactive appetitive behaviour took over to lead the agent to the chosen object. A trial ended when the agent touched the object. A total of 450 trials were used to populate the long-term memory, half of which were with a red object and the other half with a green object.

b) Task description

Foraging is an everyday, basic task for most animals and is an accepted paradigm to test the efficacy and hypotheses of control architectures in closely related fields such as evolutionary robotics (Heinerman et al., 2016; Sugawara et al., 2004; Winfield, 2009). We were primarily interested in the effectiveness of top-down attention in biasing bottom-up competition between rewards and distractors towards rewards, leading to elicited behavioural outcomes preferring rewards. Hence, to control for the presentation of both types of objects, the agent started every trial with a number of objects equidistant to it and within its field of vision, so that it could theoretically detect them all without needing to rotate. The arena that it foraged in is a square of length ten metres. In the experiments, the

agent always had intralayer inhibition (W_{intra}) for sensory competition as a form of bottom-up attentional processing. However, the value of the top-down biasing gains (G_{TDexc} and G_{TDinh}) was changed between conditions. When they were zero, the generated behaviour was a result of only bottom-up attentional processes.

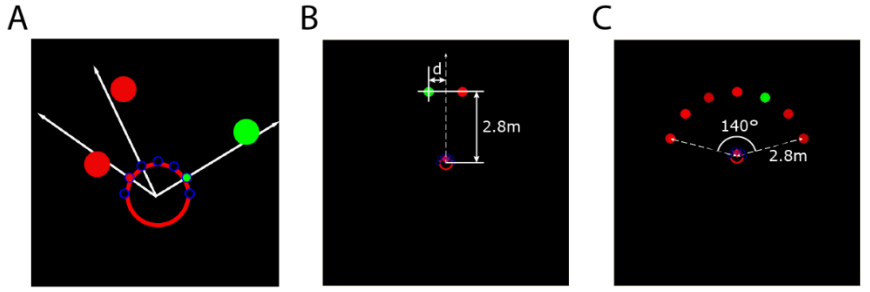


Figure 18. *Experimental setup.* A) Example of how the agent senses objects in its environment. The agent has seven visual sensors, represented by the blue circles spread across its front. The white lines show the linear projections of three visual sensors, and the individual sensor is filled with the colour of the object that it detects. A sensor filled with black indicates that is has not detected any object. B) The first experiment consisted of two stimuli presented to the agent, and the order of the reward (green) and distractor (red) was randomised. The positions of the objects were symmetric about the agent's midline, but their distance to the midline, indicated by *d*, was not fixed (0.23 m − 0.78 m). C) The second experiment consisted of seven objects presented to the agent, and the proportion of rewards was varied as an experimental condition (reward-distractor ratio). The positions at which the objects were placed were fixed to be evenly distributed, but the order was randomised.

There were two versions of the task; the first was designed to observe the effects of only the excitatory or only the inhibitory top-down attentional mechanism in the most basic problem of two objects competing for attention in the visual field. Hence, the agent was presented with two objects, one a reward and the other a distractor, located equidistant to it but with varying distances to each other (Figure 18-B). The objects were always 2.8 m away from the agent in the axis parallel to its orientation, but their position in the perpendicular axis varied 0.23 m−0.78 m from the centre (Figure 18-B, parameter *d*). As there were only two objects, one would be on the left while the other would be on the right of the agent's midline. Their values were indicated by their colour: green for a reward and red for a distractor. The exact colour of the objects was randomly drawn from a range of intensities (200–255) in the object's RGB channel to introduce some sensory noise and the order of the objects was

randomised. To prevent either mechanism from confounding the results, one of the top-down gains was kept constant at 0, eliminating top-down biasing with its associated mechanism, while the other gain was tested in the range [0–0.7] with step sizes of 0.1. 50 trials were run per gain condition. A trial ended when the agent reached one of the objects, indicating that the object had 'won' the competition for attention by manifesting appetitive behaviour towards it. The agent's performance was thus defined as the proportion of trials in which it reached a reward instead of a distractor.

In the second experiment, the agent was again presented with objects located equidistant to it. However, this time there were seven objects and the ratio between rewards and distractors was varied across experimental conditions (Figure 18-C). The number of objects chosen matched the number of visual sensors in the agent to maximise the number of objects that the agent could potentially detect. Within each reward-distractor ratio condition, the order of the objects was randomised. The number of rewards was varied, such that there is a modulation of the reward-distractor ratio. When there was only one green object, most of the objects around the agent were red and therefore distractors. The opposite occurred when there were 6 green objects. The objects were evenly distributed in a semicircle in front of the agent spanning 140 °. All of them were 2.8 m away from the agent. Just as in the first experiment, a trial ended when the agent reached an object and one of the measures of the agent's performance was the proportion of trials in which it reached a reward. Based on results from the first experiment, when active, the top-down biasing gains for both excitatory and inhibitory mechanisms were set to 0.5 in this experiment as it had yielded approximately the mean performance for top-down gains greater than 0. There were four attentional conditions: with no top-down biasing (i.e. $g_{G,TDexc}=g_{R,TDinh}=0$), with only inhibitory top-down biasing (i.e. $g_{G,TDexc}=0$, $g_{R,TDinh}=0.5$), with only excitatory top-down biasing (i.e. $g_{G,TDexc}=0.5$, $g_{R,TDinh}=0$), and with both excitatory and inhibitory biasing (i.e. $g_{G,TDexc}=g_{R,TDinh}=0.5$). In addition, there were four reward-distractor ratios: with one reward, three rewards, four rewards and six rewards. Hence, there were 16 combinations of attentional

condition and reward-distractor ratio, and for each combination 50 trials were run leading to a total of 800 trials.

c) Model formalisation

The proposed attentional and control architecture is shown in Figure 17-A, and the model is formalised as follows:

1) A set of *nvs* visual sensors $\{s_1, s_2, \ldots, s_{nvs}\} \in S$, where $nvs=7$
2) A set of *nvs* observations for each sensor $\{c_{1,t}, c_{2,t}, \ldots, c_{nvs,t}\} \in Sc$ at time *t*, with each $c_i \in Sc$ associated to sensor $s_i \in S$, each consisting of three colour channels $\{c_{R,i}, c_{G,i}, c_{B,i}\} \in c_i$ and initialised to 0
3) A 2-dimensional intra-layer weight vector $\{w_{R,intra}, w_{G,intra}, w_{B,intra}\} \in W_{intra}$, where $\{w_{j,intra,1}, w_{j,intra,2}, \ldots, w_{j,intra,nvs}\} \in w_{j,intra}$ denotes the weights for each sensor $s_i \in S$ along the channels $j \in \{R, G, B\}$
4) A 2-dimensional vector of top-down excitatory gains $\{g_{R,TDexc}, g_{G,TDexc}, g_{B,TDexc}\} \in G_{TDexc}$, where $\{g_{j,TDexc,1}, g_{j,TDexc,2}, \ldots, g_{j,TDexc,nvs}\} \in g_{j,TDexc}$ denotes the weights for each sensor $s_i \in S$ along the channels $j \in \{R, G, B\} = J$
5) A 2-dimensional vector of top-down inhibitory gains $\{g_{R,TDinh}, g_{G,TDinh}, g_{B,TDinh}\} \in G_{TDinh}$, where $\{g_{j,TDinh,1}, g_{j,TDinh,2}, \ldots, g_{j,TDinh,nvs}\} \in g_{j,TDinh}$ denotes the gain for each sensor $s_i \in S$ along the channels $j \in \{R, G, B\}$
6) A 2-dimensional saliency vector $\{p_1, p_2, \ldots, p_{nvs}\} \in P$
7) *Computing top-down attention biasing*
    a. In the current time step *t*, the observation of the selected sensor, $s_s \in S$, from the previous time step *t-1*, $c_{s,t-1} \in S_{c,t-1}$ is used to determine if the previously observed object is green or red.
    b. If it is green, the sensor's current observation $c_{s,t}$ is enhanced by summing $c_{s,t-1} \times g_{G,TDexc,s}$ to it
    c. If it is red, the sensor's current observation $c_{s,t}$ is suppressed by subtracting $c_{s,t-1} \times g_{R,TDinh,s}$ from it

8) *Computing bottom-up sensory competition*
    a. The incoming observations, *Sc*, are compressed into the three channels $\{c_R, c_G, c_B\} \in C$
    b. The intra-layer weights are calculated $W_{intra} = C * w_{intra}$
    c. All observations for a channel per sensor, $\{c_{j,1}, c_{j,2}, \ldots, c_{j,nvs}\} \in c_j$ with $j \in J = \{R, G, B\}$, is then subjected to subtractions of $W_{J\text{-}j,intra}$, where $W_{J\text{-}j,intra} \in W_{intra}$

d) Bottom-up sensory competition

The bottom-up process transforms raw, sensory input into perceptual input for the contextual layer to use for decision-making. Each visual sensor returns the intensity of the object it detects in each colour channel (red, green, and blue). If it does not detect any object, it returns zeros. The combined activity in each colour channel is first multiplied by a weight ($w_{j,intra}$) and then subtracted from the activity in each sensor for the other channels. A winner-takes-all algorithm then selects the sensor with the highest activity and uses its signal as the adaptive layer's output, represented by *P* in **Model formalisation** above. The sensory inhibitory weight was fixed in all conditions and experiments at $w_{j,intra}$=0.3 based on selectivity indices of single-cell recordings in a monkey's V4 (Reynolds et al., 1999). Of the possible selectivity indices reported, we had focused on the region where the neurons responded selectively to only one stimulus in a tested pair as that was the condition in the current experimental paradigm.

e) Top-down biasing components

The two main functions of the contextual layer are memorisation and recall, which are enabled through the structures of short-term memory and long-term memory. The short-term memory stores sensorimotor couplets which are chained together temporally and stored in the long-term memory when a goal state is reached. The long-term memory is thus able to compare current sensory

information with its contents to determine the optimal action to take to reach the goal state.

In this study, the long-term memory was populated beforehand with sensorimotor couplets of appetitive behaviour towards both red and green objects. During the experiment, to ensure that experience and memory did not confound the results, the initialised long-term memory for all trials was the same and was not updated during the experiment. Sensory information in couplets in the long-term memory were compared with the current perceived information, and matches led to probability distributions for the possible actions to take. The quality of the match between the actual prototype, *x*, and the predicted state in the long-term memory sequences, *e*, in a long-term memory with *N* sequences is defined by the discrepancy measure D which is obtained through Eq. 15.

$$D(t+1) = \propto_D D(t) + (1 - \propto_D) d(x, e) \qquad \textit{Eq. 15}$$

Where α_D is the integration time constant and the distance *d(x, e)* between the states is calculated as in Eq. 16.

$$d(x, e) = \frac{1}{N} \sum_{j=1}^{N} \left| \frac{x_i}{max(x_j)} - \frac{e_j}{max(e_j)} \right| \qquad \textit{Eq. 16}$$

During recall, all the prototypes in memory, *e*, is compared with the generated prototype, *e*. The degree of matching of segment *l* in a sequence *q* determines its input to its collector, *c* (Eq. 17).

$$c = (1 - d(e, e_{lq}))t_{lq} \qquad \textit{Eq. 17}$$

$$t_{lq}(t+1) = \propto_t + (1 - \propto_t) t_{lq}(t) \qquad \textit{Eq. 18}$$

The collector determines the segment's contribution to action selection depending on the distance *d(.)* and a trigger value, t_{lq}. This trigger value biases the matching process and allows chaining through the sequence. The default value of triggers is 1, which does not bias the collector value, and subsequently its value is determined

by Eq. 18, where $\alpha_t \in [0; 1]$. The trigger value of a selected sequence is reset to 1. This trigger value implements a decision inertia, which biases the sequence that was previously selected (Figure 19-A).

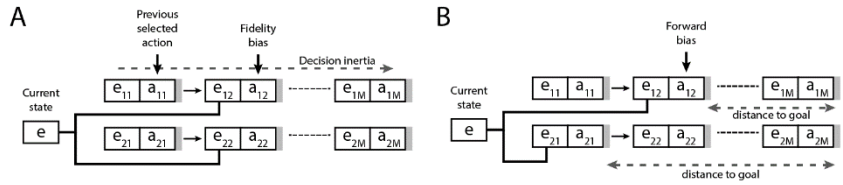


Figure 19. *Selected concepts in the contextual layer.* The current state, *e*, was compared with the states in the contextual memory's long-term memory (e_{11}, $e_{12, \ldots}$, e_{MN}), shown by the lines extending from the current state. The remembered states were given probability values and the action associated to the state with the highest probability was selected. The assignment of probabilities was calculated based on several state characteristics, two of which were A) decision inertia, where the state belonging to the previously selected sequence was preferred, and B) distance to goal, where the state that is closest to the goal in its sequence was preferred.

Furthermore, action selection in the contextual layer also took into consideration, distance to goal, relative reward and discounted reward per couplet. There was also a forward bias which gave priority to couplets according to how close they were to the end of their goal state (Figure 19-B). The selected action, a_c, is forward biased by distance to goal based on Eq. 19, where δ_{lq} is the distance measured in couplets from the selected couplet to the last couplet in the sequence. The sign of the fraction depends on the sequence's associated reward; positive for appetitive and negative for aversive. Discounted rewards, where the reward at the end of a sequence is propagated backwards to earlier couplets in the sequence, and relative rewards, where reward values are normalised to the available range of values in the long-term memory, were also calculated when transferring sequences from the short-term memory to the long-term memory, but played a smaller role in this task due to the consistent value of rewards and the relatively limited field of vision and physical exploration of the agent. The implementation of the contextual layer was adopted from Marcos et al. (2014).

$$a_c = \sum_{l,q \in LTM} \pm \frac{c_{lq} H\left(c_{lq} - \theta^C\right)}{\delta_{lq}} a_{lq} \qquad \textit{Eq. 19}$$

Once the sensorimotor couplet is selected, the contextual layer triggered the associated motor action. At the same time, it determined if the selected action was due to a reward or a distractor. If it was due to a reward, the contextual layer sent to the sensor that had provided the perceptual input an excitatory signal for the next timestep. If it was due to a distractor, the contextual layer sent an inhibitory signal instead. These excitatory and inhibitory signals were calculated with fixed gains $g_{G,TDexc}$ and $g_{R,TDinh}$ respectively. A gain of 0 would eliminate the corresponding top-down biasing signal while a gain of 1 would close to eliminate the sensory signal for inhibited sensors and double the sensory signal from excited sensors.

f) Data analysis

All grouped data was first tested for normality using D'Agostino and Pearson's omnibus normality test that combines skew and kurtosis (D'Agostino, 1971). For multiple comparisons between non-parametric data, the Friedman test (Friedman, 1937) was used to determine if there were any significant differences between groups.

Post-hoc testing or comparisons consisting were carried out using estimation statistics based on Monte Carlo permutation (Ho et al., 2019), with the paired mean difference figures including a bootstrap 95% confidence interval from 5000 resamples to illustrate the effect size. This avoids any assumption as to the data's normality, focuses on effect sizes and avoids the pitfalls of traditional significance testing (Halsey et al., 2015). The confidence interval accounts for skewness in the distributions with a bias-corrected and accelerated bootstrap (Efron, 1987). Each test was based on 5000 resamples. The distribution of all data is reported either as 'condition (mean±std)' when inline or as 'mean±std (condition)' when nested among other statistical information. Conditions where the gains of top-down biasing signals are set to 0 are referred to as 'no TD'; conditions

where only excitatory top-down biasing gains are nonzero are ‘exc’ while when only inhibitory top-down biasing gains are nonzero the label ‘inh’ is used. When both excitatory and inhibitory gains are nonzero, it is labelled as ‘both’.

5.2 Results

a) Top-down attentional biasing increased probability of choosing a reward over a distractor

In the first experiment, to determine the viability of excitatory and inhibitory mechanisms to bias selective attention, we ran simulations of an agent choosing between two objects, one reward and one distractor, with varying gains $g_{G,TDexc}$ and $g_{R,TDinh}$. Chance level performance was 0.5, and it was found that both mechanisms improved performance to a mean of at least 0.7 as long as the top-down excitatory or inhibitory gain was at least 0.1 (Figure 20-A).

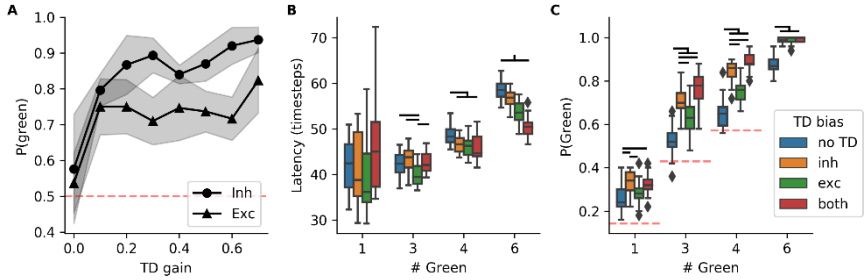


Figure 20. *Performance of the agent in the two experiments.* A) In the first experiment, the probability of a successful trial, determined by the agent approaching and touching a reward (green), across a range of top-down (TD) gains when the top-down biasing mechanism was either only inhibitory or only excitatory. The red line indicates chance level. B) The mean durations per trial in each condition in the second experiment. Within each column, the top-down biasing groups are ordered, from left to right, as follows: no top-down biasing (no TD), only inhibitory top-down biasing (inh), only excitatory top-down biasing (exc), and both inhibitory and excitatory top-down biasing (both). Stand-alone horizontal lines indicate significance between the distributions located below their ends. Horizontal lines connected to another horizontal line below indicate significance between the distribution below the single free end and all the distributions covered by the lower line. A horizontal line with a single vertical line over it indicates significance between all the distributions under the horizontal line. C) The probability of the agent touching a green reward in the second experiment. The order and labelling are the same as in the previous plot. The red dotted lines indicate chance level for that reward proportion.

The trials were then separated into three groups and their durations, in time steps, were analysed: without top-down biasing (68±32 steps), with nonzero inhibitory gains (80±38 steps), and with nonzero excitatory gains (65±33 steps). Trials with nonzero excitatory weights were significantly faster than those without top-down biasing and with nonzero inhibitory gains (Table 15).

b) Attentional mechanisms led to quicker trials when there were more rewards than distractors

We next analysed the latencies of trials in the second experiment. In general, the differences in mean latency between attentional conditions increased with the number of rewards, with the clearest differences found when a large majority of the objects were rewards. When there was one reward, the top-down attentional mechanism did not significantly affect trial duration (Figure 20-B; 42.±6.0 steps (no TD), 41.±7.8 steps (inh), 40.±8.2 (exc), 46.±9.3 steps (both); Friedman, $p=0.077$, $\chi^2(3)=6.9$). However, when there were three rewards, we found that the duration of trials with only excitatory top-down biasing was significantly quicker than the other conditions (Friedman, $p=0.0021$, $\chi^2(3)=14.$; Table 15).

With four rewards, all the conditions with some form of attentional biasing had similar durations and were all significantly faster than when there was no attentional biasing (49.±2.1 steps (no TD), 47.±1.8 steps (inh), 46.±2.3 steps (exc), 46.±2.9 steps (both); Friedman, $p=0.0039$, $\chi^2(3)=13.$). When there were six rewards, there were significant differences between all the attentional conditions (59.±2.0 steps (no TD), 57. ±1.9 steps (inh), 54.±2.4 (exc), 50.±2.2 steps (both); Friedman, $p<0.001$, $\chi^2(3)=47.$).

Table 15. *Post-hoc Monte Carlo permutation testing statistics for mean differences in latencies of trials.* Exp: experiment; N(G): number of green objects, which serve as rewards; CI (95%): the 95% confidence interval.

Exp	N(G)	Control	Test	Difference	CI (95%)		p-value
					lower	upper	
1	1	No TD	Exc	-3.1	–6.0	–0.25	0.043
	1	No TD	Inh	13.	9.8	16.	<0.001
	1	Exc	Inh	16.	14.	18.	<0.001
2	3	No TD	Inh	1.2	–0.44	2.9	0.15
	3	No TD	Exc	-2.2	–3.6	–0.57	0.0096
	3	No TD	Both	0.55	–0.93	2.2	0.50
	3	Exc	Both	2.7	1.3	4.1	0.0012
	3	Inh	Exc	–3.4	–4.8	–1.8	<0.001
	4	No TD	Inh	–2.0	–3.2	–0.88	0.0024
	4	No TD	Exc	–2.4	–3.8	–1.1	0.0018
	4	No TD	Both	–3.0	–4.5	–1.5	<0.001
	4	Inh	Exc	–0.42	–1.7	0.81	0.53
	6	No TD	Inh	–1.8	–3.0	–0.71	0.006
	6	No TD	Exc	–4.9	–6.3	–3.6	<0.001
	6	No TD	Both	–8.2	–9.4	–6.9	<0.001
	6	Exc	Both	–3.3	–4.6	–1.8	<0.001
	6	Inh	Exc	–3.1	–4.4	–1.8	<0.001

c) Combining excitatory and inhibitory attentional mechanisms improved performance regardless of reward proportion

The agent's performance is calculated similarly to the first experiment. A trial is considered a success if the agent reaches any reward, and a failure when it reaches a distractor. This gives a probability of reaching a green object for each condition (Figure 20-C). When there was one reward, the excitation model did not perform significantly different from the model without top-down biasing (Table 16). The models with only inhibitory biasing and both excitatory and inhibitory biasing performed similarly, and at a level significantly better than without top-down biasing and with only excitatory biasing (0.33±0.054 (inh), 0.32 ±0.049 (both); Friedman, $p=0.001$, $\chi^2(3)=16.$; Table 16)

The models with some form of top-down attention all performed better than the one without any top-down biasing once the reward-distractor ratio increased. Performance was significantly better with

only inhibitory than with only excitatory top-down biasing, and performance was best when both inhibitory and excitatory mechanisms were used. With three rewards, this effect was significant (0.52±0.078 (no TD), 0.70 ±0.061 (inh), 0.62±0.084 (exc), 0.77 ±0.076 (both); Friedman, $p<0.001$, $\chi^2(3)=43.$) and with four rewards as well (0.65±0.073 (no TD), 0.84 ±0.051 (inh), 0.76±0.056 (exc), 0.88 ±0.040 (both); Friedman, $p<0.001$, $\chi^2(3)=43.$). The statistics from the individual post-hoc Monte Carlo permutation testing can all be found in Table 16.

When there was a large majority of rewards, six out of seven, the performance of the models with top-down attention reached a ceiling but the model without top-down biasing remained significantly poorer in performance (0.87±0.042 (no TD), 0.99 ±0.012 (inh), 0.99±0.018 (exc), 0.99 ±0.0098 (both); Friedman, $p<0.001$, $\chi^2(3)=45.$; Table 16).

Table 16. *Post-hoc Monte Carlo permutation testing statistics for mean differences in probability of success.* Exp: experiment; N(G): number of green objects, which serve as rewards; CI (95%): the 95% confidence interval.

Exp	N(G)	Control	Test	Difference	CI (95%)		p-value
					lower	upper	
2	1	No TD	Inh	0.076	0.040	0.11	<0.001
	1	No TD	Exc	0.033	–0.0030	0.066	0.071
	1	no TD	Both	0.073	0.040	0.10	<0.001
	1	Inh	Exc	–0.043	–0.074	–0.008	0.015
	1	Inh	Both	–0.003	–0.033	0.028	0.85
	3	No TD	Inh	0.19	0.14	0.23	<0.001
	3	No TD	Exc	0.10	0.053	0.15	<0.001
	3	no TD	Both	0.25	0.20	0.29	<0.001
	3	Inh	Exc	–0.082	–0.13	–0.039	0.001
	3	Inh	Both	0.064	0.020	0.10	0.006
	4	No TD	Inh	0.19	0.15	0.23	<0.001
	4	No TD	Exc	0.11	0.062	0.14	<0.001
	4	no TD	Both	0.23	0.19	0.27	<0.001
	4	Inh	Exc	–0.088	–0.12	–0.053	<0.001
	4	Inh	Both	0.039	0.013	0.068	0.013
	6	No TD	Inh	0.12	0.097	0.13	<0.001
	6	No TD	Exc	0.11	0.094	0.13	<0.001
	6	no TD	Both	0.12	0.10	0.14	<0.001
	6	Inh	Exc	–0.002	–0.013	0.006	0.63
	6	Inh	Both	0.004	–0.003	0.010	0.18

5.3 Conclusions

For agents acting in dynamic environments, it is important to attend to stimuli that would better inform the selection of behaviour to achieve a goal or maintain survival. Selective attention facilitates this in humans, and typically it is thought to do so via interlayer excitation of neurons representing selected stimuli or features while intralayer inhibition of close competitors sharpens the contrast in neural activity between what is selected and what is not (Itti and Koch, 2001). However, there is increasing interest in an attentional mechanism that instead selects stimuli and features to inhibit. We investigated the effectiveness of the two mechanisms of attention in simulation, where an artificial agent performed a simple foraging task. In this task, the environment consisted of rewards and distractors and the goal of the agent was to reach a reward. The evaluation of the different attentional models was based on the agent's performance (proportion of successful trials) and the time needed to reach an object. Furthermore, the models were evaluated with a variety of reward-to-distractor ratios.

The agent was controlled using a hierarchical control architecture, namely the Distributed Adaptive Control (DAC) architecture, and its reflexive behaviour was directly controlled with a reactive layer that prevented collisions with the walls of the arena and produced explorative behaviour when the agent did not perceive any objects. When it did perceive objects, the adaptive layer would process the representations of the objects and, through a winner-takes-all mechanism, send a perceptual state to the contextual layer. The contextual layer would compare this state with those in its long-term memory, which was pre-trained to exhibit appetitive behaviour towards red and green objects, to decide on an action to take. In addition, it also biased the activity of the agent's sensors in the next time step through excitatory or inhibitory mechanisms. By changing the values of the top-down connection gains $g_{G,TDexc}$ and $g_{R,TDinh}$, the agent was tested with four types of top-down attentional models: without top-down biasing, with only excitatory biasing, with only inhibitory biasing, and with both excitatory and inhibitory biasing.

We found that all combinations of attentional biasing led to faster trials with increasing proportions of rewards to distractors. This is reminiscent of the fable of Buridan's ass, which finds itself both starving and dehydrating because it is unable to choose between a pile of hay on one side and a trough of water on the other. At first glance, it seems that when there are more rewards the agent might have had difficulty choosing one reward and carrying out actions to reach it without getting distracted by other rewards. Top down attentional mechanisms would then help the agent choose an object to approach and maintain that decision till the end of the trial. However, the contextual layer was the main driver of behaviour even in conditions without top-down attention, as all motor actions that were decided through perception take place via the contextual layer. As the contextual layer was implemented with decision inertia, where previously-selected sequences were prioritised over others, decisions made through purely bottom-up interactions would still exhibit decision inertia. Decision inertia alone is thus insufficient to explain the reduced trial durations in conditions with top-down biasing, suggesting that top-down attentional mechanisms do more than simply ensuring that a decision is carried out till its goal is met.

It was also found that top-down excitation-based biasing consistently performs faster than top-down inhibition-based biasing, with this effect being clearer with greater proportions of rewards. This might be because although all the objects were located equidistant to the agent, the low-resolution of the visual sensors meant that it did not detect all of them at the same time. In this task, the excitatory mechanism biased sensory input to the contextual layer to select directions that the agent should go towards, while the inhibitory mechanism did the same to select directions that the agent should not go towards. When there were fewer 'correct' directions, the inhibitory mechanism would allow for faster convergence than when there are fewer 'incorrect' directions. The differences in trial durations between the inhibitory and excitatory mechanisms would therefore be indicators that they do approach a search task from distinct perspectives. This was further supported by the shorter trial durations when both mechanisms were in effect and the reward proportion is high.

The agent performed better with top-down attentional biasing than without in all the reward proportion conditions, demonstrating the usefulness of top-down attention even in a task as simple as this. In addition, it was found that inhibitive biasing consistently led to better performance in terms of probability of a successful trial than excitatory biasing. This was somewhat unexpected, but, as with trial durations, it could be a sign that the two mechanisms do occupy separate niches in a search paradigm. Most notably, the attentional model that consistently performed the best was one that combined both excitatory and inhibitory mechanisms. This supports the hypothesis that the two mechanisms are complementary.

The current study used a simple foraging task to illustrate the effects of the two mechanisms of top-down attentional biasing on behaviour. While it was sufficient for that purpose, it still revolved around a classic visual search task. Future work could test the models in a paradigm that is closer to an oddball task, where what is predictable is not task-relevant. Also, the visual sensors used in the agent had very low resolution. While they approximate vision in humans, it is nowhere close to the complexity of visual input from human eyes. Its simplicity, and that of the environment and objects, allowed the sensory competition to be simulated cleanly through representations of three colour channels.

As future work, it would be interesting to test the models' performance in more ecological environments, such as when the perceptual states it deals with is of higher resolution or is noisier, for example from an autoencoder's compressed representation of a two-dimensional camera in an environment with more variation of objects. A wider range of stimulus features can also be used as perceptual states in future experiments, such as shape, orientation, or, at a higher level of abstraction, object type. Additionally, the calculation for bottom-up saliency can be more detailed in terms of biological-grounding, for example including a parameter for contrast. Also, learning and attention are closely coupled (e.g. Eldar et al., 2013; Grossberg, 1999), and a potentially interesting field of research would be on how the excitatory and inhibitory models of attention

could lead to differences in the learning process either through changes in neuronal synapses or through higher-level learning strategies.

6. CONCLUSIONS

Except for the most basic scenarios and behaviours, the interaction between higher-order top-down processes and sensory-driven bottom-up processes is critical for selective attention, which is the selection of stimuli for further sensory processing. This selection is evidenced by the fact that < 1% of visual information from the retina reaches conscious perception (Anderson et al., 2005). Over the years, most research has focused on how top-down processes influence selective attention by exciting representations of the selected stimuli. In this excitatory mechanism, inhibition most often takes the form of surround suppression or overlay suppression (Petrov et al., 2005). Hence, using the spatial domain as an example, top-down attentional processes shift the spatial centre and surround of the classical receptive field of sensory neurons to the attended location (Anton-Erxleben et al., 2009). Experimental findings have, however, suggested that this is not a complete mechanistic description of selective attention in the human brain. In response, an alternative, inhibitory mechanism of selective attention was proposed (e.g. Houghton and Tipper, 1994), where top-down processes target representations of task-irrelevant features for inhibition. This thesis focused on exploring the biological and computational plausibility of such an inhibitory mechanism. It formalised the mechanism using the Validation Gate hypothesis, which extended the initial proposal to include the known interplay between working memory capacity and sensory processing (e.g. Sörqvist et al., 2012). This allowed the thesis to systematically analyse the proposed mechanism, with each study validating a hypothesis from a previous one and using the emerging new insights to continue the chain of investigation.

The first chapter used psychophysiological methods to dissociate the contributions of top-down and bottom-up processes on selective attention, allowing us to further examine the impact of both cognitive impairment and working memory load on top-down attentional biasing. We found that, as predicted by the Validation Gate hypothesis, an increase in working memory load led to a decrease in sensitivity to task-relevant stimulus features, both in the spatial domain and in the colour domain. This supports the theory that top-down attentional processes form inhibitory anticipatory fields that are task-specific, and that they dynamically increase their threshold

when working memory capacity is taxed by other cognitive processes, thereby reducing the amount of sensory input to the system to a volume that is manageable at the given time.

In addition, the first study used a generalised linear mixed-effects model to show that cognitive impairment interacts with explicit detections rather than implicit detections in the displacement detection task, implying that the behavioural measures of standard clinical scales is an outcome of impairment in top-down processes rather than bottom-up ones. This raises the issue of a lack of specificity present in many standard clinical scales for cognitive impairment. The scales may diagnose patients based on impairments in different aspects of cognitive function, leading to cognitive impairment being an ineffective diagnosis for rehabilitation. The second study of the same chapter also found an increase in microsaccadic activity during retrieval of working memory items with higher working memory load. This comes on the heels of earlier findings that saccadic onset lead to phase reset of neocortical and hippocampal theta oscillations (Jutras et al., 2013), and that the phase that the oscillations are reset to is indicative of the working memory function used at the moment of the triggering saccade (Kragel et al., 2020). Together, they imply that saccadic eye movements have a role beyond that of active sampling and, as increased movement was elicited even when participants tried to maintain fixation, it is a necessary part of the working memory system.

Having established the functional validity of the inhibitory mechanism of selective attention, this thesis elucidated and tested a computational spiking model of the described mechanism in the second chapter. Specific cortico-thalamocortical pathways that map to the functional inhibitory model of attention have been experimentally shown in animal models to modulate performance in attentional tasks (e.g. Wimmer et al., 2015). However, the human counterpart of this network has yet to be directly tested. The brain regions modelled were chosen based on brain physiology and aimed to produce responses that were indicative of selective attention while minimising the complexity of the model. The model successfully carried out two attentional tasks and showed reasonable robustness against external noise. It also illustrated the differential effects of cognitive load and external noise on performance. As it was implemented using spiking neurons, the neural dynamics could be

analysed and benchmarked against. Top-down signals in the computational model was shown to produce changes in neural spiking patterns in the thalamic nucleus that mirrored what was found in mice, lending further support to the plausibility of the model's thalamocortical substrate.

For agents that interact with their environment, the purpose of most cognitive processes is to guide behavioural choices that help the agent achieve a goal, even if that goal is as general as surviving for as long as possible. Selective attention is no exception; indeed, it is arguable that one key purpose of selective attention is to maintain performance in the face of conflicting signals through the biased selection of information (Posner, 1988). In chapter 3, this thesis therefore put several mechanisms of selective attention to the test by measuring an artificial agent's ability to select rewards among distractors through action. Using a simple foraging task to prevent confounding elements, such as visual occlusion, this thesis demonstrated that any kind of top-down biasing mechanism is, unsurprisingly, useful for improving performance. Avoiding a false dichotomy between the top-down excitatory mechanism and the top-down inhibitory mechanism, the simulation also included an attentional model where both top-down excitatory and inhibitory mechanisms contributed to the biasing of sensory processing. The artificial agent's performance showed that while the excitatory mechanism consistently led to quicker physical selection of an object, in many reward-distractor ratios the inhibitory mechanism was better at maximising rewards. Nonetheless, especially at low reward-distractor ratios, it was the combination of both top-down excitation and inhibition that yielded the best results in terms of accuracy reward selection. This reiterates the importance of both mechanisms in selective attention, each complementing the shortcomings of the other.

Currently, most experimental tasks are biased towards paradigms where task-relevant stimuli are predictable, leading to an inherent bias towards evidence for the excitatory mechanism of selective attention. On top of this, it was found that inhibition, unlike excitation, evoked no measurable blood oxygenation changes, implying that brain activity as measured by imaging methods using blood oxygenation is implicitly biased towards detection of changes

due to excitatory connections (Waldvogel et al., 2000). These methodological biases may explain the overshadowing of the inhibitory mechanism by the more popular, excitatory one; increasing physiological evidence for the inhibitory model is, however, changing this. The presented results point to the inhibitory mechanism being, in certain circumstances, not only computationally more efficient but also behaviourally preferred over excitatory mechanisms. The utility of the approach as a diagnostic for deficits in executive function, the generalization to non-spatial perceptual processing, the coupling between eye movements and mnemonic processing, the distinct signature in bursting rate of thalamic neurons, and the balancing of the inhibitory top-down channel and an excitatory one in a speed-accuracy trade-off are all novel outcomes that push the boundaries of the field. They also indicate a pathway that is separate from the established excitatory top-down biasing mechanisms, specifically through the thalamic reticular nucleus, which offers an experimental method to test the existence of both mechanisms in the brain. Indeed, animal models have shown the existence of the inhibitory (Wimmer et al., 2015) and excitatory (Saalmann et al., 2007) pathways, although there remains discussion as to whether the excitatory mechanism is thalamocortical, corticocortical, or both (Anderson et al., 2011; Van Essen, 2005). The next step would be to bring the methods together to test the importance of either pathways in tasks that vary the predictability of either task-relevant or task-irrelevant stimuli.

While opening new research questions through the validation of a new pathway of attentional processing, the results presented here also implicate other cognitive functions that often work hand-in-hand with selective attention, perhaps most notably memory and learning. This thesis has focused entirely on the use of selective attention to elicit an immediate, or close to immediate, goal-oriented behaviour, and investigating the longer-term effects of the different mechanisms of attention on learning would be interesting. In addition, there are several psychological disorders and developmental disabilities that have been associated with impaired or dysfunctional attentional control, such as schizophrenia, depression and autism. Research into these topics would benefit from examining their attentional dysfunctions through the integration of both top-down excitatory and inhibitory mechanisms.

Bibliography

Adeli, H., Zelinsky, G., 2018. Deep-BCN: Deep networks meet biased competition to create a brain-inspired model of attention control, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. https://doi.org/10.1109/CVPRW.2018.00259

Allport, A., 2016. Selection for action: Some behaviorial and neurophysiological considerations of attention and action, in: Perspectives on Perception and Action. pp. 395–419. https://doi.org/10.4324/978131562799

Anderson, C.H., Van Essen, D.C., Olshausen, B.A., 2005. Directed visual attention and the dynamic control of information flow, in: Neurobiology of Attention. https://doi.org/10.1016/B978-012375731-9/50007-0

Anderson, J.C., Kennedy, H., Martin, K.A.C., 2011. Pathways of attention: Synaptic relationships of frontal eye field to V4, lateral intraparietal cortex, and area 46 in macaque monkey. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.0622-11.2011

Anderson, M.L., 2003. Embodied Cognition: A field guide. Artif. Intell. 149, 91–130. https://doi.org/10.1016/S0004-3702(03)00054-7

Anton-Erxleben, K., Stephan, V.M., Treue, S., 2009. Attention reshapes center-surround receptive field structure in macaque cortical area MT. Cereb. Cortex. https://doi.org/10.1093/cercor/bhp002

Armstrong, K.M., Moore, T., 2007. Rapid enhancement of visual cortical response discriminability by microstimulation of the frontal eye field. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.0701104104

Axmacher, N., Henseler, M.M., Jensen, O., Weinreich, I., Elger, C.E., Fell, J., 2010. Cross-frequency coupling supports multi-item working memory in the human hippocampus. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.0911531107

Badde, S., Myers, C.F., Yuval-Greenberg, S., Carrasco, M., 2020. Oculomotor freezing reflects tactile temporal expectation and aids tactile perception. bioRxiv. https://doi.org/10.1101/2020.04.27.064899

Barbas, H., 2000. Connections underlying the synthesis of cognition, memory, and emotion in primate prefrontal cortices. Brain Res. Bull. 52, 319–330. https://doi.org/10.1016/S0361-

9230(99)00245-2
Bayes, T., 1763. LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S. Philos. Trans. R. Soc. London 370–418.
Beck, J.M., Ma, W.J., Kiani, R., Hanks, T., Churchland, A.K., Roitman, J., Shadlen, M.N., Latham, P.E., Pouget, A., 2008. Probabilistic Population Codes for Bayesian Decision Making. Neuron. https://doi.org/10.1016/j.neuron.2008.09.021
Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C. a M., Boulby, P. a, Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, a J., Brady, J.M., Matthews, P.M., 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. 6, 750–7. https://doi.org/10.1038/nn1075
Belmonte, M.K., Yurgelun-Todd, D.A., 2003. Functional anatomy of impaired selective attention and compensatory processing in autism. Cogn. Brain Res. https://doi.org/10.1016/S0926-6410(03)00189-7
Berens, P., 2009. CircStat : A MATLAB Toolbox for Circular Statistics . J. Stat. Softw. https://doi.org/10.18637/jss.v031.i10
Berk, L.E., Winsler, A., 1995. Scaffolding Children's Learning: Vygotsky and Early Childhood Education. NAEYC Research into Practice Series. Volume 7. ERIC.
Bezdudnaya, T., Cano, M., Bereshpolova, Y., Stoelzel, C.R., Alonso, J.M., Swadlow, H.A., 2006. Thalamic burst mode and inattention in the awake LGNd. Neuron. https://doi.org/10.1016/j.neuron.2006.01.010
Blackburn, I.M., 1975. Mental and psychomotor speed in depression and mania. Br. J. Psychiatry. https://doi.org/10.1192/bjp.126.4.329
Boblett, N., 2012. Scaffolding: Defining the metaphor. TESOL Appl. Linguist.
Borji, A., Itti, L., 2013. State-of-the-art in visual attention modeling. IEEE Trans. Pattern Anal. Mach. Intell. 35, 185–207. https://doi.org/10.1109/TPAMI.2012.89
Braitenberg, V., 1986. Vehicles: Experiments in Synthetic Psychology. Philos. Rev. https://doi.org/10.2307/2185146
Braun, J., Julesz, B., 1998. Withdrawing attention at little or no cost: Detection and discrimination tasks. Percept. Psychophys. https://doi.org/10.3758/BF03211915

Breazeal, C., Scassellati, B., 1999. A context-dependent attention system for a social robot, in: IJCAI International Joint Conference on Artificial Intelligence.
Broadbent, D.E., 1958. Perception and communication.
Brooks, R.A., 1986. A Robust Layered Control System For A Mobile Robot. IEEE J. Robot. Autom. https://doi.org/10.1109/JRA.1986.1087032
Bruce, N., Tsotsos, J., 2010. Attention based on information maximization. J. Vis. https://doi.org/10.1167/7.9.950
Catto, E., 2011. Box2d: A 2d physics engine for games.
Chabris, C.F., Weinberger, A., Fontaine, M., Simons, D.J., 2011. You do not talk about fight club if you do not notice fight club: Inattentional blindness for a simulated real-world assault. Iperception. https://doi.org/10.1068/i0436
Chalk, M., Herrero, J.L., Gieselmann, M.A., Delicato, L.S., Gotthardt, S., Thiele, A., 2010. Attention Reduces Stimulus-Driven Gamma Frequency Oscillations and Spike Field Coherence in V1. Neuron. https://doi.org/10.1016/j.neuron.2010.03.013
Chen, S., Epps, J., 2013. Automatic classification of eye activity for cognitive load measurement with emotion interference. Comput. Methods Programs Biomed. https://doi.org/10.1016/j.cmpb.2012.10.021
Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. https://doi.org/10.1017/S0140525X12000477
Colombini, E.L., da Silva Simões, A., Ribeiro, C.H.C., 2016. An attentional model for autonomous mobile robots. IEEE Syst. J. 11, 1308–1319.
Conway, A.R.A., Kane, M.J., Bunting, M.F., Hambrick, D.Z., Wilhelm, O., Engle, R.W., 2005. Working memory span tasks: A methodological review and user's guide. Psychon. Bull. Rev. 12, 769–786. https://doi.org/10.3758/BF03196772
Corbetta, M., Shulman, G.L., 2011. Spatial neglect and attention networks. Annu. Rev. Neurosci. https://doi.org/10.1146/annurev-neuro-061010-113731
Corsi, P.M., 1973. Human memory and the medial temporal region of the brain. Diss. Abstr. Int.
Cosman, J.D., Lowe, K.A., Woodman, G.F., Schall, J.D., 2018. Prefrontal Control of Visual Distraction. Curr. Biol. 28, 414-420.e3. https://doi.org/10.1016/j.cub.2017.12.023

Cutrone, E.K., Heeger, D.J., Carrasco, M., 2014. Attention enhances contrast appearance via increased input baseline of neural responses. J. Vis. 14, 16–16. https://doi.org/10.1167/14.14.16

D'Agostino, R.B., 1971. An omnibus test of normality for moderate and large size samples. Biometrika. https://doi.org/10.1093/biomet/58.2.341

Daneman, M., Carpenter, P. a., 1980. Individual differences in working memory and reading. J. Verbal Learning Verbal Behav. 19, 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Debener, S., Kranczioch, C., Herrmann, C.S., Engel, A.K., 2002. Auditory novelty oddball allows reliable distinction of top-down and bottom-up processes of attention. Int. J. Psychophysiol. https://doi.org/10.1016/S0167-8760(02)00072-7

Dedoncker, J., Brunoni, A.R., Baeken, C., Vanderhasselt, M.A., 2016. The effect of the interval-between-sessions on prefrontal transcranial direct current stimulation (tDCS) on cognitive outcomes: a systematic review and meta-analysis. J. Neural Transm. https://doi.org/10.1007/s00702-016-1558-x

Deleuze, C., Huguenard, J.R., 2006. Distinct electrical and chemical connectivity maps in the thalamic reticular nucleus: Potential roles in synchronization and sensation. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.2333-06.2006

Desimone, R., 1998. Visual attention mediated by biased competition in extrastriate visual cortex Visual attention mediated by biased competition in extrastriate visual cortex. Philos. Trans. R. Soc. London 1245–1255. https://doi.org/10.1098/rstb.1998.0280

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. https://doi.org/10.1146/annurev.ne.18.030195.001205

Destexhe, A., Contreras, D., Steriade, M., 1998. Mechanisms underlying the synchronizing action of corticothalamic feedback through inhibition of thalamic relay cells. J. Neurophysiol. https://doi.org/10.1152/jn.1998.79.2.999

Dollar, P., Appel, R., Belongie, S., Perona, P., 2014. Fast feature pyramids for object detection. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2014.2300479

Drew, T., Võ, M.L.H., Wolfe, J.M., 2013. The Invisible Gorilla

Strikes Again: Sustained Inattentional Blindness in Expert Observers. Psychol. Sci. https://doi.org/10.1177/0956797613479386
Driscoll, J.A., Peters, R.A., Cave, K.R., 1998. Visual attention network for a humanoid robot, in: IEEE International Conference on Intelligent Robots and Systems. pp. 1968–1974. https://doi.org/10.1109/iros.1998.724894
Dubois, B., Slachevsky, A., Litvan, I., Pillon, B., 2000. The FAB: A frontal assessment battery at bedside. Neurology. https://doi.org/10.1212/WNL.55.11.1621
Duchowski, A., 2007. Eye tracking methodology: Theory and practice, Eye Tracking Methodology: Theory and Practice. https://doi.org/10.1007/978-1-84628-609-4
Duff, A., Sanchez Fibla, M., Verschure, P.F.M.J., 2011. A biologically based model for the integration of sensory-motor contingencies in rules and plans: A prefrontal cortex based extension of the Distributed Adaptive Control architecture. Brain Res. Bull. https://doi.org/10.1016/j.brainresbull.2010.11.008
Duncan, J., 1984. Selective attention and the organization of visual information. J. Exp. Psychol. Gen. https://doi.org/10.1037/0096-3445.113.4.501
Efron, B., 1987. Better bootstrap confidence intervals. J. Am. Stat. Assoc. https://doi.org/10.1080/01621459.1987.10478410
Eldar, E., Cohen, J.D., Niv, Y., 2013. The effects of neural gain on attention and learning. Nat. Neurosci. https://doi.org/10.1038/nn.3428
Engbert, R., Kliegl, R., 2003. Microsaccades uncover the orientation of covert attention. Vision Res. https://doi.org/10.1016/S0042-6989(03)00084-1
Engbert, R., Sinn, P., Mergenthaler, K., Trukenbrod, H., 2015. Microsaccade Toolbox for R. URL http//read. psych. uni-potsdam. de/attachments/article/140/MS_Toolbox_R. zip.
Engel, A.K., Maye, A., Kurthen, M., König, P., 2013. Where's the action? The pragmatic turn in cognitive science. Trends Cogn. Sci. https://doi.org/10.1016/j.tics.2013.03.006
Engle, R.W., 2002. Working memory capacity as executive attention. Curr. Dir. Psychol. Sci. https://doi.org/10.1111/1467-8721.00160
Eriksen, C.W., Hoffman, J.E., 1972. Temporal and spatial characteristics of selective encoding from visual displays.

Percept. Psychophys. 12, 201–204.
Eriksen, C.W., James, J.D.S., 1986. Visual attention within and around the field of focal attention: A zoom lens model. Percept. Psychophys. 40, 225–240.
Esterman, M., Yantis, S., 2010. Perceptual expectation evokes category-selective cortical activity. Cereb. Cortex. https://doi.org/10.1093/cercor/bhp188
Faber, L.G., Maurits, N.M., Lorist, M.M., 2012. Mental Fatigue Affects Visual Selective Attention. PLoS One. https://doi.org/10.1371/journal.pone.0048073
Faber, M., Bixler, R., D'Mello, S.K., 2018. An automated behavioral measure of mind wandering during computerized reading. Behav. Res. Methods. https://doi.org/10.3758/s13428-017-0857-y
Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex. https://doi.org/10.1093/cercor/1.1.1
Fitzgerald, P.J., 2013. Gray colored glasses: Is major depression partially a sensory perceptual disorder? J. Affect. Disord. https://doi.org/10.1016/j.jad.2013.06.045
Foerster, R.M., Schneider, W.X., 2018. Involuntary top-down control by search-irrelevant features: Visual working memory biases attention in an object-based manner. Cognition. https://doi.org/10.1016/j.cognition.2017.12.002
Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. https://doi.org/10.1016/0022-3956(75)90026-6
Friedman, M., 1937. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. J. Am. Stat. Assoc. https://doi.org/10.1080/01621459.1937.10503522
Frintrop, S., Rome, E., Nüchter, A., Surmann, H., 2003. An attentive, multi-modal laser "eye," in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/3-540-36592-3_20
Friston, K., 2012. The history of the future of the Bayesian brain. Neuroimage. https://doi.org/10.1016/j.neuroimage.2011.10.004
Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. B Biol. Sci. https://doi.org/10.1098/rstb.2005.1622
Gagne, R.M., 1968. Learning hierarchies. Educ. Psychol.

https://doi.org/10.1080/0022027750070205
Gaspelin, N., Leonard, C.J., Luck, S.J., 2015. Direct Evidence for Active Suppression of Salient-but-Irrelevant Sensory Inputs. Psychol. Sci. https://doi.org/10.1177/0956797615597913
Gazzaley, A., Nobre, A.C., 2012. Top-down modulation: Bridging selective attention and working memory. Trends Cogn. Sci. https://doi.org/10.1016/j.tics.2011.11.014
Giesbrecht, B., Weissman, D.H., Woldorff, M.G., Mangun, G.R., 2006. Pre-target activity in visual cortex predicts behavioral performance on spatial and feature attention tasks, in: Brain Research. https://doi.org/10.1016/j.brainres.2005.09.068
Gilinsky, A.S., Cohen, H.H., 1972. Reaction time to change in visual orientation. Percept. Psychophys. https://doi.org/10.3758/BF03210358
Gregoriou, G.G., Gotts, S.J., Zhou, H., Desimone, R., 2009. High-Frequency, long-range coupling between prefrontal and visual cortex during attention. Science (80-.). https://doi.org/10.1126/science.1171402
Grossberg, S., 2013. Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. Neural Networks. https://doi.org/10.1016/j.neunet.2012.09.017
Grossberg, S., 1999. The Link between Brain Learning, Attention, and Consciousness. Conscious. Cogn. https://doi.org/10.1006/ccog.1998.0372
Grossberg, S., 1987. Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. Adv. Psychol. https://doi.org/10.1016/S0166-4115(08)60909-7
Grossberg, S., Versace, M., 2008. Spikes, synchrony, and attentive learning by laminar thalamocortical circuits. Brain Res. https://doi.org/10.1016/j.brainres.2008.04.024
Guggisberg, A.G., Koch, P.J., Hummel, F.C., Buetefisch, C.M., 2019. Brain networks and their relevance for stroke rehabilitation. Clin. Neurophysiol. https://doi.org/10.1016/j.clinph.2019.04.004
Guillery, R.W., 1995. Anatomical evidence concerning the role of the thalamus in corticocortical communication: a brief review. J. Anat.
Guillery, R.W., Harting, J.K., 2003. Structure and connections of the thalamic reticular nucleus: Advancing views over half a

century. J. Comp. Neurol. https://doi.org/10.1002/cne.10738
Haasch, A., Hofemann, N., Fritsch, J., Sagerer, G., 2005. A multi-modal object attention system for a mobile robot, in: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS. https://doi.org/10.1109/IROS.2005.1545191
Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Van Wedeen, J., Sporns, O., 2008. Mapping the structural core of human cerebral cortex. PLoS Biol. https://doi.org/10.1371/journal.pbio.0060159
Halassa, M.M., Siegle, J.H., Ritt, J.T., Ting, J.T., Feng, G., Moore, C.I., 2011. Selective optical drive of thalamic reticular nucleus generates thalamic bursts and cortical spindles. Nat. Neurosci. https://doi.org/10.1038/nn.2880
Hall, J., O'Carroll, R.E., Frith, C.D., 2010. 7 - Neuropsychology, in: Johnstone, E.C., Owens, D.C., Lawrie, S.M., McIntosh, A.M., Sharpe, M.B.T.-C. to P.S. (Eighth E. (Eds.), . Churchill Livingstone, St. Louis, pp. 121–140. https://doi.org/https://doi.org/10.1016/B978-0-7020-3137-3.00007-3
Halsey, L.G., Curran-Everett, D., Vowler, S.L., Drummond, G.B., 2015. The fickle P value generates irreproducible results. Nat. Methods. https://doi.org/10.1038/nmeth.3288
Han, S.H., Kim, M.S., 2004. Visual search does not remain efficient when executive working memory is working. Psychol. Sci. https://doi.org/10.1111/j.0956-7976.2004.00730.x
Hastie, T., Tibshirani, R., Friedman, J., 2009. Elements of Statistical Learning 2nd ed., Elements. https://doi.org/10.1007/978-0-387-84858-7
Heinerman, J., Zonta, A., Haasdijk, E., Eiben, A.E., 2016. On-line evolution of foraging behaviour in a population of real robots, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-31153-1_14
Helmholtz, H. von, 1962. Helmholtz's treatise on physiological optics (JPC Southall, Trans.). Handb. der Physiol. Opt.
Herman, J.P., Katz, L.N., Krauzlis, R.J., 2018. Midbrain activity can explain perceptual decisions during an attention task. Nat. Neurosci. https://doi.org/10.1038/s41593-018-0271-5
Ho, J., Tumkaya, T., Aryal, S., Choi, H., Claridge-Chang, A., 2019. Moving beyond P values: data analysis with estimation

graphics. Nat. Methods 16, 565–566. https://doi.org/10.1038/s41592-019-0470-3
Hoffman, K.L., Dragan, M.C., Leonard, T.K., Micheli, C., Montefusco-Siegmund, R., Valiante, T.A., 2013. Saccades during visual exploration align hippocampal 3–8 Hz rhythms in human and non-human primates. Front. Syst. Neurosci. https://doi.org/10.3389/fnsys.2013.00043
Hopfinger, J.B., Buonocore, M.H., Mangun, G.R., 2000. The neural mechanisms of top-down attentional control. Nat. Neurosci. https://doi.org/10.1038/72999
Houghton, G., Tipper, S.P., 1994. A model of inhibitory mechanisms in selective attention., in: Inhibitory Processes in Attention.
Hutmacher, F., 2019. Why Is There So Much More Research on Vision Than on Any Other Sensory Modality? Front. Psychol. https://doi.org/10.3389/fpsyg.2019.02246
Isaacs, E.B., Vargha-Khadem, F., 1989. Differential course of development of spatial and verbal memory span: A normative study. Br. J. Dev. Psychol. https://doi.org/10.1111/j.2044-835x.1989.tb00814.x
Ito, J., Maldonado, P., Singer, W., Grün, S., 2011. Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. Cereb. Cortex. https://doi.org/10.1093/cercor/bhr020
Itti, L., Koch, C., 2001. Computational modelling of visual attention. Nat. Rev. Neurosci. https://doi.org/10.1038/35058500
Itti, L., Koch, C., Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. 20, 1254–1259. https://doi.org/10.1109/34.730558
Izhikevich, E.M., 2003. Simple model of spiking neurons. IEEE Trans. Neural Networks 14, 1569–1572. https://doi.org/10.1109/TNN.2003.820440
Jensen, O., Lisman, J.E., 1998. An oscillatory short-term memory buffer model can account for data on the Sternberg task. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.18-24-10688.1998
John, Y.J., Zikopoulos, B., Bullock, D., Barbas, H., 2016. The Emotional Gatekeeper: A Computational Model of Attentional Selection and Suppression through the Pathway from the

Amygdala to the Inhibitory Thalamic Reticular Nucleus. PLOS Comput. Biol. 12, 1–36. https://doi.org/10.1371/journal.pcbi.1004722
Johnston, W.A., Dark, V.J., 1986. Selective Attention. Annu. Rev. Psychol. https://doi.org/10.1146/annurev.ps.37.020186.000355
Jones, E.G., 2002. Thalamic circuitry and thalamocortical synchrony, in: Philosophical Transactions of the Royal Society B: Biological Sciences. https://doi.org/10.1098/rstb.2002.1168
Jutras, M.J., Fries, P., Buffalo, E.A., 2013. Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1302351110
Kahneman, D., 1973. Attention and effort. Citeseer.
Kastner, S., DeSimone, K., Konen, C.S., Szczepanski, S.M., Weiner, K.S., Schneider, K.A., 2007. Topographic maps in human frontal cortex revealed in memory-guided saccade and spatial working-memory tasks. J. Neurophysiol. https://doi.org/10.1152/jn.00010.2007
Kastner, S., Ungerleider, L.G., 2001. The neural basis of biased competition in human visual cortex. Neuropsychologia 39, 1263–1276. https://doi.org/10.1016/S0028-3932(01)00116-6
Kastner, S., Ungerleider, L.G., 2000. Mechanisms of Visual Attention in the Human Cortex. Annu. Rev. Neurosci. 23, 315–341.
Katsuki, F., Constantinidis, C., 2012. Early involvement of prefrontal cortex in visual bottom-up attention. Nat. Neurosci. https://doi.org/10.1038/nn.3164
Katz, C.N., Patel, K., Talakoub, O., Groppe, D., Hoffman, K., Valiante, T.A., 2020. Differential Generation of Saccade, Fixation, and Image-Onset Event-Related Potentials in the Human Mesial Temporal Lobe. Cereb. Cortex. https://doi.org/10.1093/cercor/bhaa132
Kaufman, A.S., Lichtenberger, E.O., 2002. Assessing adolescent and adult intelligence, Wiley.
Kauhanen, M.L., Korpelainen, J.T., Hiltunen, P., Brusin, E., Mononen, H., Määttä, R., Nieminen, P., Sotaniemi, K.A., Myllylä, V. V., 1999. Poststroke depression correlates with cognitive impairment and neurological deficits. Stroke. https://doi.org/10.1161/01.STR.30.9.1875
Kertzman, S., Reznik, I., Hornik-Lurie, T., Weizman, A., Kotler, M., Amital, D., 2010. Stroop performance in major depression:

Selective attention impairment or psychomotor slowness? J. Affect. Disord. https://doi.org/10.1016/j.jad.2009.08.009
Kidd, G., Arbogast, T.L., Mason, C.R., Gallun, F.J., 2005. The advantage of knowing where to listen. J. Acoust. Soc. Am. https://doi.org/10.1121/1.2109187
Koch, C., Ullman, S., 1985. Shifts in selective visual attention: Towards the underlying neural circuitry. Hum. Neurobiol. https://doi.org/10.1007/978-94-009-3833-5_5
Kotseruba, I., Tsotsos, J.K., 2020. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artif. Intell. Rev. https://doi.org/10.1007/s10462-018-9646-y
Kragel, J.E., Vanhaerents, S., Templer, J.W., Schuele, S., Rosenow, J.M., Nilakantan, A.S., Bridge, D.J., 2020. Hippocampal theta coordinates memory processing during visual exploration. Elife. https://doi.org/10.7554/eLife.52108
Leśniak, M., Bak, T., Czepiel, W., Seniów, J., Członkowska, A., 2008. Frequency and prognostic value of cognitive disorders in stroke patients. Dement. Geriatr. Cogn. Disord. https://doi.org/10.1159/000162262
Lisman, J., 2010. Working memory: The importance of theta and gamma oscillations. Curr. Biol. https://doi.org/10.1016/j.cub.2010.04.011
Lisman, J.E., Jensen, O., 2013. The Theta-Gamma Neural Code. Neuron. https://doi.org/10.1016/j.neuron.2013.03.007
Liu, Y., Li, M., Zhang, X., Lu, Y., Gong, H., Yin, J., Chen, Z., Qian, L., Yang, Y., Andolina, I.M., Shipp, S., Mcloughlin, N., Tang, S., Wang, W., 2020. Hierarchical Representation for Chromatic Processing across Macaque V1, V2, and V4. Neuron. https://doi.org/10.1016/j.neuron.2020.07.037
Lloyd, E.P., Hugenberg, K., McConnell, A.R., Kunstman, J.W., Deska, J.C., 2017. Black and White Lies: Race-Based Biases in Deception Judgments. Psychol. Sci. https://doi.org/10.1177/0956797617705399
Loboda, T.D., 2012. Reading Span (RSPAN) Task [WWW Document]. URL https://ubiq-x.github.io/rspan (accessed 2.5.16).
Maier, M., Ballester, B.R., Leiva Bañuelos, N., Duarte Oller, E., Verschure, P.F.M.J., 2020. Adaptive conjunctive cognitive training (ACCT) in virtual reality for chronic stroke patients: A randomized controlled pilot trial. J. Neuroeng. Rehabil.

https://doi.org/10.1186/s12984-020-0652-3
Malekshahi, R., Seth, A., Papanikolaou, A., Mathews, Z., Birbaumer, N., Verschure, P.F.M.J., Caria, A., 2016. Differential neural mechanisms for early and late prediction error detection. Sci. Rep. https://doi.org/10.1038/srep24350
Manning, C., Dakin, S.C., Tibber, M.S., Pellicano, E., 2014. Averaging, not internal noise, limits the development of coherent motion processing. Dev. Cogn. Neurosci. https://doi.org/10.1016/j.dcn.2014.07.004
Marat, S., Ho Phuoc, T., Granjon, L., Guyader, N., Pellerin, D., Guérin-Dugué, A., 2009. Modelling spatio-temporal saliency to predict gaze direction for short videos. Int. J. Comput. Vis. https://doi.org/10.1007/s11263-009-0215-3
Marcos, E., Pani, P., Brunamonti, E., Deco, G., Ferraina, S., Verschure, P., 2013. Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. Neuron 78, 249–255. https://doi.org/10.1016/j.neuron.2013.02.006
Marcos, E., Ringwald, M., Duff, A., Sánchez-Fibla, M., Verschure, P.F.M.J., 2014. The hierarchical accumulation of knowledge in the distributed adaptive control architecture, in: Computational and Robotic Models of the Hierarchical Organization of Behavior. https://doi.org/10.1007/978-3-642-39875-9_10
Martinez-Conde, S., Macknik, S.L., Troncoso, X.G., Hubel, D.H., 2009. Microsaccades: a neurophysiological analysis. Trends Neurosci. https://doi.org/10.1016/j.tins.2009.05.006
Martinez-Trujillo, J.C., Treue, S., 2004. Feature-based attention increases the selectivity of population responses in primate visual cortex. Curr. Biol. https://doi.org/10.1016/j.cub.2004.04.028
Mathews, Z., Cetnarski, R., Verschure, P.F.M.J., 2015. Visual anticipation biases conscious decision making but not bottom-up visual processing. Front. Psychol. 6. https://doi.org/10.3389/fpsyg.2015.00443
Mathews, Z., i Badia, S.B., Verschure, P.F.M.J., 2012. PASAR: An integrated model of prediction, anticipation, sensation, attention and response for artificial sensorimotor systems. Inf. Sci. (Ny). 186, 1–19. https://doi.org/10.1016/j.ins.2011.09.042
Maunsell, J.H.R., Newsome, W.T., 1987. Visual processing in monkey extrastriate cortex. Annu. Rev. Neurosci. https://doi.org/10.1146/annurev.ne.10.030187.002051

McAdams, C.J., Maunsell, J.H.R., 1999. Effects of attention on the reliability of individual neurons in monkey visual cortex. Neuron. https://doi.org/10.1016/S0896-6273(01)80034-9
McAlonan, K., Cavanaugh, J., Wurtz, R.H., 2006. Attentional modulation of thalamic reticular neurons. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.5602-05.2006
Mease, R. a, Krieger, P., Groh, A., 2014. Cortical control of adaptation and sensory relay mode in the thalamus. Proc. Natl. Acad. Sci. U. S. A. 111, 6798–803. https://doi.org/10.1073/pnas.1318665111
Mengotti, P., Boers, F., Dombert, P.L., Fink, G.R., Vossel, S., 2018. Integrating modality-specific expectancies for the deployment of spatial attention. Sci. Rep. https://doi.org/10.1038/s41598-018-19593-7
Mikami, A., Ito, S., Kubota, K., 1982. Visual response properties of dorsolateral prefrontal neurons during visual fixation task. J. Neurophysiol. https://doi.org/10.1152/jn.1982.47.4.593
Mirza, M.B., Adams, R.A., Friston, K., Parr, T., 2019. Introducing a Bayesian model of selective attention based on active inference. Sci. Rep. https://doi.org/10.1038/s41598-019-50138-8
Mitchell, J.P., Neil MacRae, C., Gilchrist, I.D., 2002. Working memory and the suppression of reflexive saccades. J. Cogn. Neurosci. https://doi.org/10.1162/089892902317205357
Moore, T., Zirnsak, M., 2017. Neural Mechanisms of Selective Visual Attention. Annu. Rev. Psychol. https://doi.org/10.1146/annurev-psych-122414-033400
Moran, J., Desimone, R., 1985. Selective attention gates visual processing in the extrastriate cortex. Science 229, 782–784. https://doi.org/10.1126/science.4023713
Moscatelli, A., Mezzetti, M., Lacquaniti, F., 2012. Modeling psychophysical data at the population-level: The generalized linear mixed model. J. Vis. https://doi.org/10.1167/12.11.26
Moulin-Frier, C., Puigbò, J.Y., Arsiwalla, X.D., Sanchez-Fibla, M., Verschure, P.F., 2017. Embodied artificial intelligence through distributed adaptive control: An integrated framework, in: 2017 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob). https://doi.org/10.1109/DEVLRN.2017.8329825
Murray, N., Vanrell, M., Otazu, X., Parraga, C.A., 2011. Saliency estimation using a non-parametric low-level vision model, in:

Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2011.5995506
Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J.L., Chertkow, H., 2005. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. J. Am. Geriatr. Soc. https://doi.org/10.1111/j.1532-5415.2005.53221.x
Nassi, J.J., Lomber, S.G., Born, R.T., 2013. Corticocortical feedback contributes to surround suppression in V1 of the alert primate. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.5124-12.2013
Navalpakkam, V., Itti, L., 2006. An integrated model of top-down and bottom-up attention for optimizing detection speed, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2006.54
Newell, A., 1990. Unified theories of cognition. Harvard University Press.
Nolfi, S., 2002. Power and limits of reactive agents. Neurocomputing. https://doi.org/10.1016/S0925-2312(01)00598-7
Nyamsuren, E., Taatgen, N.A., 2013. Pre-attentive and attentive vision module. Cogn. Syst. Res. https://doi.org/10.1016/j.cogsys.2012.12.010
O'Connor, D.H., Fukui, M.M., Pinsk, M.A., Kastner, S., 2002. Attention modulates responses in the human lateral geniculate nucleus. Nat. Neurosci. https://doi.org/10.1038/nn957
O'Keefe, J., Recce, M.L., 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. Hippocampus. https://doi.org/10.1002/hipo.450030307
Ognibene, D., Baldassare, G., 2015. Ecological active vision: Four bioinspired principles to integrate bottom-up and adaptive top-down attention tested with a simple camera-arm robot. IEEE Trans. Auton. Ment. Dev. https://doi.org/10.1109/TAMD.2014.2341351
Olsen, A., 2012. The Tobii IVT Fixation Filter Algorithm description.
Otazu, G.H., Tai, L.H., Yang, Y., Zador, A.M., 2009. Engaging in an auditory task suppresses responses in auditory cortex. Nat. Neurosci. https://doi.org/10.1038/nn.2306

Ouerhani, N., Hügli, H., 2005. Robot self-localization using visual attention, in: Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation, CIRA. https://doi.org/10.1109/cira.2005.1554295
Paneri, S., Gregoriou, G.G., 2017. Top-down control of visual attention by the prefrontal cortex. Functional specialization and long-range interactions. Front. Neurosci. https://doi.org/10.3389/fnins.2017.00545
Paolucci, S., Antonucci, G., Gialloreti, L.E., Traballesi, M., Lubich, S., Pratesi, L., Palombi, L., 1996. Predicting stroke inpatient rehabilitation outcome: The prominent role of neuropsychological disorders. Eur. Neurol. https://doi.org/10.1159/000117298
Park, J.H., Kim, B.J., Bae, H., Lee, Jisung, Lee, Juneyoung, Han, M., O, K.Y., Park, S.H., Kang, Y., Yu, K., Lee, B., 2013. Impact of Post-Stroke Cognitive Impairment with No Dementia on Health-Related Quality of Life. J. Stroke 15, 49–56. https://doi.org/10.5853/jos.2013.15.1.49
Pearson, K., 1895. VII. Note on regression and inheritance in the case of two parents. Proc. R. Soc. London. https://doi.org/10.1098/rspl.1895.0041
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine learning in Python. J. Mach. Learn. Res.
Peña-Casanova, J., Blesa, R., Aguilar, M., Gramunt-Fombuena, N., Gómez-Ansón, B., Oliva, R., Molinuevo, J.L., Robles, A., Barquero, M.S., Antúnez, C., 2009. Spanish multicenter normative studies (NEURONORMA project): Methods and sample characteristics. Arch. Clin. Neuropsychol. 24, 307–319.
Petrov, Y., Carandini, M., McKee, S., 2005. Two distinct mechanisms of suppression in human vision. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.2871-05.2005
Pfeifer, R., Verschure, P., 1992. Distributed adaptive control: A paradigm for designing autonomous agents, in: Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Lifee; MIT Press: Cambridgee, MA, USA. pp. 21–30.
Posner, M.I., 1988. Structures and function of selective attention. American Psychological Association.

Posner, M.I., Snyder, C.R.R., Davidson, B.J., 1980. Attention and the detection of signals. J. Exp. Psychol. Gen. 109, 160–174. https://doi.org/10.1037/0096-3445.109.2.160
Proske, J.H., Jeanmonod, D., Verschure, P.F.M.J., 2011. A computational model of thalamocortical dysrhythmia. Eur. J. Neurosci. 33, 1281–1290. https://doi.org/10.1111/j.1460-9568.2010.07588.x
Puigbò, J.Y., Arsiwalla, X.D., Verschure, P.F.M.J., 2018. Challenges of machine learning for living machines, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-95972-6_41
Reddy, L., Kanwisher, N.G., VanRullen, R., 2009. Attention and biased competition in multi-voxel object representations. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.0907330106
Reynolds, J.H., Chelazzi, L., Desimone, R., 1999. Competitive mechanisms subserve attention in macaque areas V2 and V4. J. Neurosci. 19, 1736–1753. https://doi.org/10.1523/jneurosci.19-05-01736.1999
Reynolds, J.H., Desimone, R., 2003. Interacting roles of attention and visual salience in V4. Neuron. https://doi.org/10.1016/S0896-6273(03)00097-7
Reynolds, J.H., Heeger, D.J., 2009. The Normalization Model of Attention. Neuron. https://doi.org/10.1016/j.neuron.2009.01.002
Ridderinkhof, K.R., Van Den Wildenberg, W.P.M., Segalowitz, S.J., Carter, C.S., 2004. Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cogn. https://doi.org/10.1016/j.bandc.2004.09.016
Rizzuto, D.S., Madsen, J.R., Bromfield, E.B., Schulze-Bonhage, A., Seelig, D., Aschenbrenner-Scheibe, R., Kahana, M.J., 2003. Reset of human neocortical oscillations during a working memory task. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.0732061100
Rolfs, M., 2009. Microsaccades: Small steps on a long way. Vision Res. https://doi.org/10.1016/j.visres.2009.08.010
Rossum, G., 1995. Python reference manual.

Roth, H.L., Lora, A.N., Heilman, K.M., 2002. Effects of monocular viewing and eye dominance on spatial attention. Brain. https://doi.org/10.1093/brain/awf210
Saalmann, Y.B., Pigarev, I.N., Vidyasagar, T.R., 2007. Neural mechanisms of visual attention: How top-down feedback highlights relevant locations. Science (80-.). https://doi.org/10.1126/science.1139140
Saalmann, Y.B., Pinsk, M.A., Wang, L., Li, X., Kastner, S., 2012. The pulvinar regulates information transmission between cortical areas based on attention demands. Science 337, 753–6. https://doi.org/10.1126/science.1223082
Sanbonmatsu, D., Strayer, D., 2013. Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. PLoS One 8, e54402. https://doi.org/10.1371/journal.pone.0054402
Sanbonmatsu, D.M., Strayer, D.L., Biondi, F., Behrends, A.A., Moore, S.M., 2016. Cell-phone use diminishes self-awareness of impaired driving. Psychon. Bull. Rev. https://doi.org/10.3758/s13423-015-0922-4
Sánchez-Cubillo, I., Periáñez, J.A., Adrover-Roig, D., Rodríguez-Sánchez, J.M., Ríos-Lago, M., Tirapu, J., Barceló, F., 2009. Construct validity of the Trail Making Test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. J. Int. Neuropsychol. Soc. https://doi.org/10.1017/S1355617709090626
Sani, I., Santandrea, E., Morrone, M.C., Chelazzi, L., 2017. Temporally evolving gain mechanisms of attention in macaque area V4. J. Neurophysiol. 118, 964–985. https://doi.org/10.1152/jn.00522.2016
SanMiguel, I., Corral, M.J., Escera, C., 2008. When loading working memory reduces distraction: Behavioral and electrophysiological evidence from an auditory-visual distraction paradigm. J. Cogn. Neurosci. https://doi.org/10.1162/jocn.2008.20078
Schenk, T., McIntosh, R.D., 2010. Do we have independent visual streams for perception and action? Cogn. Neurosci. https://doi.org/10.1080/17588920903388950
Schneider, G.E., 1969. Two visual systems. Science (80-.). https://doi.org/10.1126/science.163.3870.895
Schneider, W., Shiffrin, R.M., 1977. Controlled and automatic human information processing: I. Detection, search, and

attention. Psychol. Rev. 84, 1.
Schwartz, Z.P., David, S. V., 2018. Focal suppression of distractor sounds by selective attention in auditory cortex. Cereb. Cortex. https://doi.org/10.1093/cercor/bhx288
Seabold, S., Perktold, J., 2010. Statsmodels: Econometric and Statistical Modeling with Python. Proc. 9th Python Sci. Conf.
Serences, J.T., Schwarzbach, J., Courtney, S.M., Golay, X., Yantis, S., 2004. Control of object-based attention in human cortex. Cereb. Cortex. https://doi.org/10.1093/cercor/bhh095
Sherman, S.M., 2016. Thalamus plays a central role in ongoing cortical functioning. Nat. Neurosci. 16, 533–41. https://doi.org/10.1038/nn.4269
Sherman, S.M., Guillery, R.W., 2002. The role of the thalamus in the flow of information to the cortex. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 357, 1695–708. https://doi.org/10.1098/rstb.2002.1161
Shinners, P., 2011. PyGame.
Simons, D.J., Levin, D.T., 1997. Change blindness. Trends Cogn. Sci. https://doi.org/10.1016/S1364-6613(97)01080-2
Simpson, G. V., Weber, D.L., Dale, C.L., Pantazis, D., Bressler, S.L., Leahy, R.M., Luks, T.L., 2011. Dynamic activation of frontal, parietal, and sensory regions underlying anticipatory visual spatial attention. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.1519-10.2011
Smith, G.D., Cox, C.L., Sherman, S.M., Rinzel, J., 2000. Fourier analysis of sinusoidally driven thalamocortical relay neurons and a minimal integrate-and-fire-or-burst model. J. Neurophysiol. 83, 588–610.
Sörqvist, P., Stenfelt, S., Rönnberg, J., 2012. Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: toward a unified view of attention. J. Cogn. Neurosci. https://doi.org/10.1162/jocn_a_00275
Stern, J.A., Boyer, D., Schroeder, D., 1994. Blink rate: A possible measure of fatigue. Hum. Factors. https://doi.org/10.1177/001872089403600209
Sternberg, S., 1969. Memory-scanning: mental processes revealed by reaction-time experiments. Am. Sci.
Sugawara, K., Kazama, T., Watanabe, T., 2004. Foraging behavior of interacting robots with virtual pheromone, in: 2004 IEEE/RSJ International Conference on Intelligent Robots and

Systems (IROS). https://doi.org/10.1109/iros.2004.1389878
Taylor, K., Mandon, S., Freiwald, W.A., Kreiter, A.K., 2005. Coherent oscillatory activity in monkey area V4 predicts successful allocation of attention. Cereb. Cortex. https://doi.org/10.1093/cercor/bhi023
Team, R.C., 2019. R: A Language and Environment for Statistical Computing. Vienna, Austria.
Terra, H., Bruinsma, B., de Kloet, S.F., van der Roest, M., Pattij, T., Mansvelder, H.D., 2020. Prefrontal Cortical Projection Neurons Targeting Dorsomedial Striatum Control Behavioral Inhibition. Curr. Biol.
Tishby, N., Pereira, F.C., Bialek, W., 2000. The information bottleneck method. arXiv Prepr. physics/0004057.
Tombaugh, T.N., 2004. Trail Making Test A and B: Normative data stratified by age and education. Arch. Clin. Neuropsychol. https://doi.org/10.1016/S0887-6177(03)00039-8
Treisman, A., 1986. Features and Objects in Visual Processing. Sci. Am. https://doi.org/10.1038/scientificamerican1186-114b
Tsotsos, J.K., 2011. A computational perspective on visual attention. MIT Press.
Ungerleider, L.G., Mishkin, M., 1982. Two cortical visual systems. Anal. Vis. Behav. https://doi.org/10.2139/ssrn.1353746
Usher, M., Niebur, E., 1996. Modeling the temporal dynamics of IT neurons in visual search: A mechanism for top-down selective attention. J. Cogn. Neurosci. https://doi.org/10.1162/jocn.1996.8.4.311
Van Essen, D.C., 2005. Corticocortical and thalamocortical information flow in the primate visual system, in: Progress in Brain Research. https://doi.org/10.1016/S0079-6123(05)49013-5
van Essen, D.C., Maunsell, J.H.R., 1983. Hierarchical organization and functional streams in the visual cortex. Trends Neurosci. 6, 370–375. https://doi.org/10.1016/0166-2236(83)90167-4
van Wijngaarden, J.B.G., Zucca, R., Finnigan, S., Verschure, P.F.M.J., 2016. The Impact of Cortical Lesions on Thalamo-Cortical Network Dynamics after Acute Ischaemic Stroke: A Combined Experimental and Theoretical Study. PLOS Comput. Biol. 12, e1005048. https://doi.org/10.1371/journal.pcbi.1005048
Veltman, J.A., Gaillard, A.W.K., 1998. Physiological workload reactions to increasing levels of task difficulty. Ergonomics.

https://doi.org/10.1080/001401398186829
Verschure, P.F.M.J., 2016. Synthetic consciousness: The distributed adaptive control perspective. Philos. Trans. R. Soc. B Biol. Sci. https://doi.org/10.1098/rstb.2015.0448
Verschure, P.F.M.J., Voegtlin, T., Douglas, R.J., 2003. Environmentally mediated synergy between perception and behaviour in mobile robots. Nature 425. https://doi.org/10.1038/nature02024
von der Malsburg, T., 2015. saccades: Detection of Fixations in Eye-Tracking Data.
Waldvogel, D., Van Gelderen, P., Muellbacher, W., Ziemann, U., Immisch, I., Hallett, M., 2000. The relative metabolic demand of inhibition and excitation. Nature. https://doi.org/10.1038/35023171
Walker, R., Husain, M., Hodgson, T.L., Harrison, J., Kennard, C., 1998. Saccadic eye movement and working memory deficits following damage to human prefrontal cortex. Neuropsychologia. https://doi.org/10.1016/S0028-3932(98)00004-9
Wass, S. V., Forssman, L., Leppänen, J., 2014. Robustness and precision: How data quality may influence key dependent variables in infant eye-tracker analyses. Infancy. https://doi.org/10.1111/infa.12055
Wass, S. V., Smith, T.J., Johnson, M.H., 2013. Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. Behav. Res. Methods. https://doi.org/10.3758/s13428-012-0245-6
Wells, M.F., Wimmer, R.D., Schmitt, L.I., Feng, G., Halassa, M.M., 2016. Thalamic reticular impairment underlies attention deficit in Ptchd1 Y'mice. Nature 532, 58–63. https://doi.org/10.1038/nature17427
Wilson, B., Cockburn, J., Halligan, P., 1987. Development of a behavioral test of visuospatial neglect. Arch. Phys. Med. Rehabil.
Wilson, H.R., Cowan, J.D., 1972. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. Biophys. J. https://doi.org/10.1016/S0006-3495(72)86068-5
Wilson, M., 2002. Six views of embodied cognition. Psychon. Bull. Rev. 9, 625–636. https://doi.org/10.3758/BF03196322
Wimmer, R.D., Schmitt, L.I., Davidson, T.J., Nakajima, M., Deisseroth, K., Halassa, M.M., 2015. Thalamic control of

sensory selection in divided attention. Nature 526, 705–709. https://doi.org/10.1038/nature15398
Winfield, A.F., 2009. Foraging Robots, in: Encyclopedia of Complexity and Systems Science. https://doi.org/10.1007/978-0-387-30440-3_217
Woodman, G.F., Luck, S.J., 2004. Visual search is slowed when visuospatial working memory is occupied. Psychon. Bull. Rev. https://doi.org/10.3758/BF03196569
Yang, J., Yang, M.H., 2017. Top-down visual saliency via joint CRF and dictionary learning. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2016.2547384
Yu, X.J., Xu, X.X., Chen, X., He, S., He, J., 2009. Slow recovery from excitation of thalamic reticular nucleus neurons. J. Neurophysiol. https://doi.org/10.1152/jn.91130.2008
Yuval-Greenberg, S., Tomer, O., Keren, A.S., Nelken, I., Deouell, L.Y., 2008. Transient Induced Gamma-Band Response in EEG as a Manifestation of Miniature Saccades. Neuron. https://doi.org/10.1016/j.neuron.2008.03.027
Zanto, T.P., Rubens, M.T., Thangavel, A., Gazzaley, A., 2011. Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. Nat. Neurosci. https://doi.org/10.1038/nn.2773
Zhang, Y., Meyers, E.M., Bichot, N.P., Serre, T., Poggio, T.A., Desimone, R., 2011. Object decoding with attention in inferior temporal cortex. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1100999108
Zikopoulos, B., Barbas, H., 2012. Pathways for Emotions and Attention Converge on the Thalamic Reticular Nucleus in Primates. J. Neurosci. 32, 5338–5350. https://doi.org/10.1523/JNEUROSCI.4793-11.2012
Zikopoulos, B., Barbas, H., 2006. Prefrontal Projections to the Thalamic Reticular Nucleus form a Unique Circuit for Attentional Mechanisms. J. Neurosci. 26, 7348–7361. https://doi.org/10.1523/JNEUROSCI.5511-05.2006
Zuber, B.L., Stark, L., Cook, G., 1965. Microsaccades and the velocity-amplitude relationship for saccadic eye movements. Science (80-.). https://doi.org/10.1126/science.150.3702.1459