Chapter I: Our visual world

Every morning we wake up to the sound of an alarm or the warmth of the sunlight through the window, lying in bed... we open our eyes and the image of the world around us begins to form. We perceive the world through many different senses such as smell, taste, and sound. Sensory neurons (nerve cells) like those found in the ear respond to pressure changes in the air while those found in retina of the eye respond to light. They transform the physical stimuli into signals and send it to the brain. Consequently, our perception is not that of the changes in air pressure or number of photons, rather it is a manifestation of the brain’s interpretation of these signals. How the brain is able to achieve this has been a matter of study and debate for several centuries amongst philosophers and scientists. To that effect, a goal of visual neuroscience research has always been to understand the mechanisms by which the brain constructs our “perceived reality”.

In humans and many other animals, it is the sense of vision that dominates how we perceive the world around us. The concept of how we ‘see’ begins with how the eye captures light. The eye is made up of millions of cells, some of which contain light-sensitive pigments. When light falls on these cells, they activate and send an electro-chemical signal to the areas of the brain that process vision. Every individual cell captures only a tiny part of the visual field, nonetheless our perception of the world is not disjointed or fragmented, but it is rather continuous. The signals generated from these millions of cells are stitched and processed together to form the perception of colour, size, depth, shapes etc. present in the visual space around us. Our brain is able to sort through all the visual information; discarding unnecessary features but also selecting for specific features that are relevant, meaningful and suitable for the immediate situation.
For example in Figure 1, if you were asked to find a ‘blue toy airplane’ among others (try to find it)…, a simple task that probably took you very little time to complete. However, in that short time, within the brain many processing steps were performed without conscious awareness: 1) Neuronal signalling cascades decoded elementary properties of the image such as lines, colours, etc. 2) Patterns of activity characterizing these elements were processed and transformed representing complex patterns for shapes and objects. 3) Based on task goals and presumably aided by previous experience or memory, activity patterns cascaded into eye-movements that eventually resulted in finding the ‘blue toy airplane’. Over the course of evolution, our brain has learnt to accomplish this extremely well. We can infer from the above example that visual perception is a constructive process supported by not only the external stimulus but also internal representations such as previous experience and knowledge.

One cognitive process that helps us successfully complete this task is called ‘selective attention’. This ability to select objects or specific features has been eloquently described by Williams James who noted in his Principles of Psychology: “Millions of items… are present to my senses which never properly enter
my experience. Why? Because they have no interest for me. My experience is what I
agree to attend to. . . Everyone knows what attention is. It is the taking possession by
the mind, in clear and vivid form, of one out of what seem several simultaneously pos-
sible objects of trains of thought. Focalizations, concentration of consciousness, are of
its essence. It implies withdrawal from some things in order to deal effectively with
others.” Selective attention is a process that allows the brain to direct its limited
resources to the task at hand. However, the mechanisms and pathways involved in
selective attention are not completely understood. Selective attention also inter-
acts with other cognitive processes such as object-perception\textsuperscript{3–10}, conscious aware-
ness\textsuperscript{11–13}, learning\textsuperscript{14–18}, and memory\textsuperscript{19–26}.

In this thesis, I make an effort to understand the mechanisms involved in
selective attention, its signature on a molecular level and explore its interaction
with other cognitive processes. In the next sections, I will explain the anatomical
and functional architecture of the brain, beginning with how the brain perceives
objects present in the environment.

\subsection*{1.1a Perception & Gestalt Psychology}

Before advances in electrophysiology and imaging techniques, the percep-
tion of objects was studied by analysing behavioural responses of participants to
different stimuli. Psychophysics is described as “the analysis of perceptual processes
by studying the effect on a subject’s experience or behaviour by systematically
varying the properties of a stimulus along one or more physical dimensions”\textsuperscript{27} i.e. it
investigates the relationship between the physical stimuli and the perceptions they
produce. Psychophysical studies measure the behavioural markers (\textit{e.g.} accuracy,
latency etc.) of subjects to the stimuli. This allows for a systematic and qualitative
measure of perception in the absence of neurophysiological data. (\textit{See} 1.4, 1.5,
Chapter \textit{III} \& IV).
Psychophysical studies on the perception of objects have lent insightful observations: 1) the visual system groups simple elements to construct complex representations and in some cases ‘fills the gaps’ with information not actually present. For example, in Figure 2a, can you spot the Dalmatian dog? Initially, it is difficult as only disorganised black patches are seen, but after some time has passed or once the individual elements (black patches) are selected (highlighted in red) the perception of the dog emerges easily. The black patches and surrounding white spaces are seen as part of an object with boundaries. More importantly, this perception now becomes robust and stable, even after switching back to the unlabelled picture, the dog motif is perceived. 2) In the event where a stimulus is bi-stable i.e. it can be interpreted in two distinct ways, the visual system is able to

**Figure 2:** Ambiguous stimuli help understand assumptions and limitations of the visual system.
dynamically select (enhance or suppress) specific neuronal representations. In the Rubin’s vase/face\textsuperscript{28} illusion (Fig. 2b -left) attending to the white space a ‘vase’ is seen while focusing or shifting attention onto the black space reveals two faces opposite to each other. A postulated mechanism is that the stimulus is equivalent but the motif/pattern of the shape, edges, contrast etc. lead to different high level neural representations during visual processing, thus creating either perception (vase or face). However, this balance between multiple high level representations is vulnerable to processes such as attentional selection\textsuperscript{29,30} that if directed to one or the other, shifts this balance resulting in the switch in perception. Another example can be seen in the Necker’s cube\textsuperscript{31} illusion (Fig. 2b -right); focusing on lower-left or top-right vertices causes the percept of the cube protruding in or out, the percept also switches back and forth at intervals.

It is important to understand that these examples not only highlight the remarkable ability of the brain in trying to build a consistent experience but it also shows that the brain essentially tries to infer distinct predictable patterns from the environment. These inferences are products of an individual’s experience that can often bias perception by grouping together even randomly arranged discrete elements. For example in Figure 2c, we would recognize the arrangement of shapes when turned 90° clockwise as the brain invariably groups all the elements together to form an object resembling a ‘face’ with eyes, nose, and a smile.

An important contribution in understanding how the visual system processes objects was made by Gestalt (meaning, Shape or form) psychologists in the early 20\textsuperscript{th} century\textsuperscript{32–38}. They presumed that humans “perceived” objects as a whole and not as individual components such as lines, edges or blobs of colour. Accordingly, they proposed principles by which the visual system grouped discrete stimuli together. Gestalt psychologists theorised that grouping of stimuli was based on certain underlying properties such as similarity, proximity, continuity, symmetry, closure etc. (Fig. 3)\textsuperscript{32,33,35–42}. However, a criticism of gestalt principles was that the explanation was descriptive and the intrinsic neural mechanisms remained unexplained.
In later studies\(^{43,44}\), it became apparent that these grouping principles are inherently structured in the anatomical and functional organisation of the brain with attentional-feedback processes playing a crucial role in imparting flexibility depending on the task requirements (see 1.1\(e\)). An influential model called the ‘feature integration theory’\(^{8,45}\) posits that the brain first analyses location bound features/elements of the stimuli in a rapid and parallel process creating a ‘feature map’. At a later stage, attention (like a spotlight) selectively binds the mapped features to construct object percepts. Evidence from search tasks\(^{46}\) has illustrated that attention has limitations of capacity such that it can process a limited number of objects at a time. In a large array with multiple stimuli having similar features, attention binds features serially constructing one object percept and then the next. Consequently, searching for a target amongst many distractors is a time consuming process (as seen in blue toy example in figure 1).

**Figure 3:** Gestalt principles for perceptual grouping.
To truly appreciate how the brain processes visual information and builds our perceived reality we must first learn about its anatomical and functional architecture (sections 1.1b-c), beginning with how neurons respond to stimuli and how the information flows to different areas of the visual system.

1.1b Neuronal Response & Receptive Fields

Neurons transmit information from one neuron to the next, by generating electrical pulses called action potentials or spikes and by secreting chemicals called neurotransmitters or neuromodulators (e.g. Acetylcholine, Dopamine, Glutamate etc.) that bind to specific receptors at specialised structures called synapses (Fig. 4). Typically, the frequency of the action potentials corresponds to the strength of the stimulus.

![Figure 4: Illustration of Neuromodulator-Receptor interaction at the synapse. Action potential results in the secretion of neuromodulators by the pre-synaptic neuron in the synapse. Neuromodulators bind to specific receptors on the post-synaptic neuron, triggering a signalling cascade, and transmitting the signal further along.](image-url)
An important concept in understanding how the brain processes stimuli is that of a receptive field\(^{49,50}\). The classical receptive field (RF)\(^{49,50}\) of a neuron is an area in sensory space, where the presentation of a stimulus causes a neuronal response. Additionally, it also encapsulates its response properties such as tuning\(^{51}\) i.e. cells respond only to specific stimuli/features such as orientation, colour etc. presented in their RF and not to others. The tuning properties of cells vary across brain areas and stimulus types (Fig. 7) and they are modulated due to learning\(^{52,53}\) effects or by attention\(^{54,55}\) or input from nearby cells\(^{56}\).

In the visual system the RFs of retinal ganglion cells (RGCs) in the retina\(^{50}\) are a result of the structural and functional arrangement of photo-sensitive cells, bipolar cells, horizontal cells and amacrine cells (Fig. 5 -top). The RFs consist of a central disk region called ‘centre’ and concentric ring region called ‘surround’. For

---

**Figure 5:** Top panel- Illustrates the contribution of photoreceptor cells in forming receptive fields with centre-surround configuration. (Orange) – Photoreceptors, (Blue) – Retinal Ganglion Cells, and (Green) - horizontal, bipolar and amacrine cells, that help RGCs in processing. **Bottom panel**- Each vertical line represents an action potential. The response of RGC ‘ON’ cells to a spot of light, as a series of action potentials when it falls on the centre, surround, and completely over the receptive field. RGC - ‘OFF’ cells, have the response pattern reversed. Similar organisation is seen in the LGN. *(Inspired and redrawn from Principles of Neuroscience 5\(^{th}\) Ed., 2012)*\(^{58}\)
For some cells (Fig. 5 -bottom), when light falls on the excitatory centre region the cell’s response increases and if the light falls on the inhibitory surround region it decreases; in other cells this response pattern is reversed. When a spot of light falls equally on the ‘centre’ and ‘surround’ regions, the cells respond weakly. This receptive field structure helps in detecting edges and contrast. The RGCs form the optic nerve and project to the lateral geniculate nucleus (LGN) located in the thalamus. LGN shows a similar RF organisation as the retina. It receives segregated input from each eye. Effectively, the LGN serves as a ‘gateway’ or ‘a relay’ (see Fig. 7, 8) because it sends signals to the occipital cortex and other areas.

The occipital cortex, located at the ‘back’ of the brain (see Fig. 4) consists of several areas that process visual information such as area V1, area V4 etc. Historically, area V1 is called the primary visual cortex because most projections from the LGN ‘first’ arrive here and area V1 then passes the output to the others areas. Experiments by Hubel and Wiesel in cats revealed that RFs in area V1 were different compared to the retina and LGN RFs. They responded to bars of light. They classified cells into simple and complex types based on their response patterns. Simple cell RFs had distinct ‘excitatory’ and ‘inhibitory’ regions similar to retina and LGN RFs, but they were arranged side-to-side (Fig. 6a). They also showed a linear summation of responses with the excitatory and inhibitory regions and it was possible to predict the type of stimuli used based on their responses. The RFs properties of complex cells were harder to define. They lacked specific excitatory and inhibitory regions. They were not sensitive to the exact position of the stimulus within the receptive field. Based on these findings Hubel and Wiesel proposed that the RFs of simple cells in the visual cortex could be constructed by the summation of LGN RFs (Fig. 6a), and that the RFs of complex cells could be the result of summing the inputs from multiple simple cells (Fig. 6b).

Hubel and Wiesel also described tuning properties of cells (simple and complex) in area V1. They discovered that area V1 cells were arranged in columnar structures (see 1.1d) and were selectively tuned for the orientation and direction of stimuli i.e. they responded only to specific orientations (Fig. 6c) and directions
Figure 6: (a) Illustrates LGN and V1 receptive fields (RFs) and the proposed mechanism by Hubel and Wiesel on how LGN receptive field properties could be summed together to form RFs of higher area V1. ‘+’ (‘-’) sign shows excitatory (inhibitory) region. (b) Shows the summation of RFs of simple cells to form RFs of complex cells. (c) Orientation selectivity and tuning properties (Broad/Narrow response) of cells. (Inspired and redrawn from Principles of Neuroscience 5th Ed., 2012)
Figure 7: A simplified illustration of serial feedforward flow of information with receptive field (RF) properties based on cortical area.
of a moving bar. Furthermore, they could be broadly tuned or narrowly tuned. Over the years, many different cells with varied tuning properties were discovered in the brain for example, it has been observed that cells in area V4 are tuned for different colours\textsuperscript{63} and hues\textsuperscript{64}, and even shapes of intermediate complexity\textsuperscript{65,66}; cells in inferior temporal cortex (IT), which have large RFs respond to complex objects and faces. For example, in a famous study by Quiroga\textsuperscript{67} and colleagues cells tuned to the faces of famous personalities such as Jennifer Aniston and Bill Clinton were discovered (Fig. 7).

Extending the idea of RFs and tuning properties (Fig. 7), Hubel and Wiesel\textsuperscript{68} proposed a serial feedforward hierarchical model in which the RFs progressively increased in size and complexity across successive levels by integrating RFs properties of the previous level. In the next section, we will see how the information is transmitted along the visual hierarchy.

1.1c Flow of information in the visual system

When a visual stimulus is presented, it rapidly propagates through the visual system from lower to higher areas of the visual brain; this is called the feedforward sweep. Information from retinal cells about the exposure (amount of light), contrast etc. is sent to the LGN in the thalamus. Cells in the LGN\textsuperscript{62} have similar receptive fields (RFs) to the retina; they pass this information to the primary visual cortex, area V1. Cells in area V1 are tuned to the orientation and direction of stimuli. From area V1, the information flows to area V2. Area V2\textsuperscript{69} has larger RFs and it transforms and integrates signals from area V1. Here cells are tuned to borders of objects, texture etc. From area V2\textsuperscript{69,70}, the information is sent to two parallel but interconnected pathways\textsuperscript{71} (Fig. 7): 1) The ventral stream and 2) the dorsal stream. In the ventral processing stream, also known as the “what” pathway, information from area V2 is sent to area V4. Area V4\textsuperscript{63–66} has large RFs and cells are tuned to colours and features like lines, curves of objects and simple shapes. From there the information is sent to areas in the interior temporal (IT) cortex such as TEO. Responses of cells in IT cortex are very specialized and tuned
to complex objects and shapes. Some of the cells are even selective for faces\textsuperscript{67}. In the dorsal processing stream, also known as the “where” pathway, area V2 projects to area MT/MST. Cells in area MT process motion, position and depth of objects\textsuperscript{72}. From there the information is sent to areas in the parietal cortex (MIP, LIP, AIP, VIP), to the premotor areas (PMd, PMv) and to the frontal eye fields and prefrontal cortex where cells are tuned to planning movement and establishing goals\textsuperscript{73,74}. Areas in the prefrontal cortex also receive signals from other brain areas\textsuperscript{75}, including limbic areas and those that process other sensory modalities\textsuperscript{76}.

\textbf{Figure 8:} Depicts visual processing streams in the monkey brain along with tuning properties for different areas. Information flow begins from the retina (dashed blue line) to the LGN in the thalamus and then to Area V1 in the occipital cortex. Two parallel processing streams diverge: Ventral stream (solid red) and dorsal stream (solid blue). Double arrows indicate reciprocal connections. Other pathways between areas in solid purple, dashed red.
Over the years, it became evident that the serial feedforward visual hierarchy model\cite{51,61,68} (Fig. 7) was principally sound but incomplete. Multiple factors pointed towards a more complicated architecture for visual processing: 1) In an exhaustive study, Felleman & Van Essen\cite{60,77} and others\cite{78,79} have described approximately 10 to 14 levels of cortical processing comprising of visual areas containing several reciprocal, inter-twined parallel processing streams (Fig. 8; Fig. 9) and also connections between visual and non-visual areas such as the basal forebrain\cite{80–82} and mid-brain\cite{83,84}. Furthermore, they concluded that there were at least three types of connections: (a) feedforward connections from lower areas such as area V1 and area V4 to higher areas such as frontal eye fields (FEF), inferior temporal (IT) cortex, feedback connections from higher areas to lower areas and lateral/horizontal connections between areas at similar hierarchical levels\cite{56}. Consequently, information can flow from lower to higher areas along the visual hierarchy, but it can also flow back from higher to lower areas\cite{17,56,60,77,85–87} (Fig. 8; Fig. 9). 2) Studies\cite{56,88} measuring temporal properties of signals in different cortical areas showed that the latency of signals tends to increase at higher processing levels. However, there are a few exceptions, where higher areas receive the signal earlier than intermediate processing areas, which can be explained by the existence of multiple parallel pathways\cite{60,77,89} for the flow of information. Accordingly, both in the dorsal and ventral processing stream, many parallel and reciprocal connections between the different areas have been found. 3) The discovery of non-classical receptive fields\cite{90–93} i.e. an area surrounding the classical receptive field of a cell, where if stimulated by itself it does not evoke activity. However, stimulation can modulate responses to stimuli presented in the classical receptive field depending on the context or task goals\cite{90–93}. Studies have documented these ‘contextual effects’ in the primary visual cortex and they are thought to be a result of feedback and horizontal connections\cite{10,56,90,94,95}.
Figure 9: Depicts a few cortical areas involved in visual processing and the connections between them. (Adapted from Felleman and Van Essen, 1991)
1.1d Topographic organisation of the cortex

Visual cortical areas show a systematic and organized representation of the visual space. A property of early visual areas is that adjacent points in visual space are mapped to adjacent loci in the visual cortex, this is called retinotopy\textsuperscript{96}. Many areas of the brain responding to visual stimuli show retinotopic organisation. Retinotopic representation in the cortical areas is distorted i.e. the centre of the eye (fovea) is over-represented compared to the peripheral areas (Fig. 10a). The result of this distortion is that many more cells are committed for processing the small centre region than for the periphery, this is called cortical magnification\textsuperscript{97} (Fig. 10b).

A second organisational feature of the visual cortex is that cells are arranged in a column-like structure called cortical columns\textsuperscript{98-100}. The cortical column is not just an anatomical feature, but also a functional unit consisting of multiple neurons arranged in a vertical axis with similar receptive fields. They are classified anatomically as mini-columns\textsuperscript{101} and macro-columns\textsuperscript{102}. A mini-column contains about 80-100 neurons. Approximately 50 to 100 mini-columns form a macro-column, bound together by short-horizontal connections. An important characteristic of columnar organisation is that connections within the column are much denser than connections between columns. Hubel and Wiesel\textsuperscript{102,103} first described orientation preference columns in the area V1 of monkeys (Fig. 10c). When the orientation preferences are mapped, we see that columns are arranged together radially forming a pinwheel around a centre point. More detailed imaging studies in cats and primates have revealed columnar maps showing preference for ocular dominance\textsuperscript{102-105} (segregation of input from the two eyes) in area V1. In V1, maps for orientation selectivity\textsuperscript{64} have been observed and in V4, maps for colour/hue\textsuperscript{64}.

A third ubiquitous feature of the cortex is that in a cross-section (slice) of the tissue, a laminar or layered structure\textsuperscript{78} is seen. There are six layers; layer I is a superficial layer and layer VI is the deepest layer. Neurons in each layer differ in their morphology (shape, size) and layers have different innervation patterns.
Figure 10: Organisational principles of the visual cortex (a) Depicts the retinal eccentricity map. The centre of the concentric circles represents the fovea of the eye and the circles show the visual field in degree from the centre. (b) Cortical magnification map that shows the over-representation of the foveal region in area V1. (c) Orientation columns are arranged in pinwheel structure with different colour hues representing orientation preferences. (d) Depicts the six cortical layers (I to VI) along with the different shapes and morphologies of the neurons. A simplified laminar circuitry for area V1 showing connections within the layers and input and output connections between areas of the brain. *(Inspired and redrawn from Principles of Neuroscience 5th Ed., 2012)*
(Fig. 10d). The anatomical projections of the layers can be local i.e. within the same area or to other areas of the brain. The connections by which information from lower areas is sent to higher areas as described above are called feedforward projections\(^{78,89}\). In Area V1\(^{106,59,107}\), information arrives first from the LGN at layer IV, in addition to some at layer VI; neurons in the layer IV subsequently target neurons in layer II & III. From layer II & III the information flows to other cortical areas. Similarly, there are reciprocal connections that originate in higher areas and project to lower areas called feedback projections\(^{78,89}\). Unlike feedforward connections that target layer IV, feedback connections target superficial and deep layers such as layer I and V. Feedback connections from Layer VI project to the LGN and from layer V to superior colliculus and other sub-cortical areas.

1.1e Attention & Feedback processing

The brain has a limited capacity\(^{108}\) to process visual information and must select only that which is of immediate relevance. Attention allows for allocation and reallocation of these limited resources to the task at hand\(^{54}\). It is distinguished into exogenous/bottom-up attention\(^{9,109,110}\) which is driven by external stimuli, or endogenous/top-down attention\(^{9,109,110}\) which is goal-driven. Furthermore, we can also direct attention to a stimulus with an eye-movement\(^{111}\) (overt-attention) or by attending to it without an eye-movement (covert attention)\(^{112}\). Attention can also be directed to specific features described as feature based attention\(^{8,45,46,113}\) and even to entire objects, termed as object-based attention\(^{2-6}\).

Here, we will specifically explore the relationship between feedback processing and object-based attention with respect to perceptual grouping. We learnt in section 1.1a that Gestalt psychologists described how the visual system grouped elements together but did not adequately explain the neural processes for perceptual organisation. One of the interesting problems in neuroscience is called the 'binding problem'\(^{114}\). We know that different neuronal populations of the brain process different stimuli features (see. 1.1b, c). The binding problem posits that if two or more stimuli with multiple features (e.g. shape, colour etc.) are processed...
simultaneously, how does the brain combine the discrete features corresponding to the respective stimuli? For example in Figure 11, we see that the shape and colour of the aircraft and the boat are processed separately. Speculatively, it is possible to perceive all outcomes, but to our knowledge, this does not happen. One might speculate that the brain consists of dedicated neuronal connections and neurons representing the complex feature combination(s). However, this would lead to ‘combinatorial explosion’ i.e. a need for an exponential number of dedicated neurons/neural connections to represent every possible complex feature combination.

Roelfsema and colleagues proposed the ‘incremental grouping theory (IGT)’\textsuperscript{10,43} that is based on feedforward and feedback connectivity, a more modern approach compared to the ‘feature integration model’\textsuperscript{8,45}. IGT suggests attentional-feedback signals group low-level features together to form percept(s) by labelling them with enhanced neuronal activity. Consider the following task (Fig. 12): In a jumbled up mess of wires, we must find and trace the correct wire for the plug. An easy task...?

\textbf{Figure 11:} Binding problem: Combining the correct features corresponding to the stimulus.
...Very likely, you began to trace from the arrow, at first slowly and then with quick saccadic (eye-movement) jumps to parts of the curve; slowing at points where the curves cross each other, till you eventually reached the electric plug! Keeping in mind that this arrangement of curves was novel, how were we able to tell the two curves apart and correctly trace the wire connecting the plug?

The IGT\textsuperscript{10,43} elucidates that perceptual grouping has two processing steps 1) a rapid process called ‘base-grouping’ that relies on feedforward processing and dedicated neurons coding for stimuli features (see 1.1b,c). Base-grouping corresponds to ‘pre-attentive vision’ at the psychological level, a process that occurs without needing attention or feedback. However, there are constraints to base-grouping such that it can only occur for features that are familiar and not for novel features. 2) As it is not possible for the brain to have dedicated neurons for every feature combination, a dynamic and flexible system is necessary. For grouping novel feature combinations, a time-consuming ‘incremental-grouping process’ is needed to help label grouped image elements. This labelling process is presumed to require attentional-feedback signals from higher areas to bias/select from input at lower areas that increase the firing rate of these neurons.

However, much about the source and origin of the attentional-feedback signal remains to be explored. Clarity on the mechanisms and functions comes from the study of the temporal aspect of these signals. Roelfsema and colleagues\textsuperscript{3,85,115} trained monkeys to perform (Fig. 13a) a contour grouping task called

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure12.png}
\caption{Trace the wire to the plug point. (See Chapter III – Fig. 1 for more information)}
\end{figure}
curve-tracing task\textsuperscript{3,116–118} while they recorded activity in Area V1 to study the role of feedback signals in object-based attention. The monkeys fixate on a point on screen, the fixation point, for a period of time, following which they are presented with two curves. A ‘target curve’ is connected to the fixation point while the other curve, not connected to the fixation point is the distractor curve. The monkey keeps fixating and must decide which of the two curves is the ‘target’ curve and report by making an eye-movement to the disc at the end of the curve. In Figure 12a, we can see that the neurons in area V1 show an initial visual response at approximately 50 ms associated with the feedforward sweep\textsuperscript{85,88,119}; at this time, there was no distinction between activity for target and distractor curves. After 150-250 ms, however, they observed a difference in activity with more activity elicited by the attended curve than the non-attended curve\textsuperscript{54}. This difference in activity is called attentional modulation. It is presumably the result of feedback activity from higher areas.

Wannig and colleagues\textsuperscript{44} reported that when attention is directed to an element or feature of an object, attentional modulation (increased activity) spreads across the entire object based on Gestalt principles\textsuperscript{32,33} in a serial (incremental) manner, strongly suggesting that the spread of attentional modulation over objects was not an instantaneous process but rather a time consuming process\textsuperscript{10,43,44,120}. Reaction times increase linearly with the length of the curve, with delays increasing up to several hundreds of milliseconds\textsuperscript{121} (see 1.5; Chapter IV for more information). Interestingly, a similar pattern of activity (as seen in curve-tracing) is observed in other tasks such as figure-ground perception\textsuperscript{56,122,123} (Fig. 13b) and visual search tasks\textsuperscript{124}, where the subject is asked to search for a target object among distractor.

Over the years, neuronal signatures for the attentional-feedback signal have been explored using a wide variety of techniques. Neurophysiological studies have observed increased response for attended objects in many areas such as V1\textsuperscript{125}, V4\textsuperscript{126}, frontal eye fields\textsuperscript{127} (FEF) and others\textsuperscript{128,129}. Studies measuring synchrony between large population of neurons have shown an increase in gamma\textsuperscript{130}. 
Figure 13: (a) Curve-tracing task and its response modulation for target and distractor curves. FP – Fixation Point, RF - Receptive field (green square) in area V1 (b) Figure-ground task and its response modulation. Receptive Field (green square), figure (red square), and background (texture). Both tasks show similar pattern of activity. (Adapted from Lamme & Roelfsema, 2000)\textsuperscript{85}. 
(30-100 Hz) band in area V4 and FEF and beta band\textsuperscript{131} frequencies (13-30 Hz) in FEF, which have been shown to track shifts of attention. Imaging techniques such as fMRI (functional magnetic resonance imaging) which measure changes in blood flow representing neural activity have shown large signal changes in early visual areas\textsuperscript{132,133} for attended versus non-attended stimuli. Studies have also suggested that neuromodulatory systems such as the cholinergic\textsuperscript{125,128,134}, dopaminergic\textsuperscript{128} and noradrenergic\textsuperscript{128} play an important role.

An important difference between the feedforward and feedback processes is that while feedforward processing is largely dependent on direct anatomical connectivity\textsuperscript{60,77} and fixed neuronal tuning properties of cells, feedback processes are more dynamic and flexible. Feedback processes modulate neural patterns activated during the feedforward sweep based on the context or relevance to immediate task goals. For example, contextual\textsuperscript{10,56,90,94,95} effects such as the enhancement or suppression of the representation of specific features of the visual stimulus. They also can modulate tuning properties of cell\textsuperscript{3,56}.

Computational models support that feedback processes allow the brain to select and change/bias in favor of a specific neural input based on task goals or previous experience\textsuperscript{16,20,54}. Consequently, feedback processes provide a mechanism by which cognitive processes such as selective attention\textsuperscript{3,56,120,127,135}, learning\textsuperscript{14-16}, working memory\textsuperscript{19-26} etc. can influence the selection of behaviourally relevant information arising from feedforward signals.
1.2 Research questions

Overall, it can be said that the attentional-feedback signal carries a rich amount of information that facilitates the selection and segmentation of objects and backgrounds in our visual world. However, much about the influence of the attentional feedback signal on other cognitive processes such as learning\textsuperscript{14,17,18,136–141}, working memory\textsuperscript{19–24,142,143} and conscious access to information\textsuperscript{11,12} remain to be understood. In this thesis, I explore the role of attentional-feedback signals and their interaction with different cognitive processes occurring in the brain. We ask the following questions:

1) What is the role of the cholinergic system and acetylcholine in attention? How can we measure it? \textit{(Chapter II)}

2) How do selective attention and reward interact with each other to influence trial and error learning? \textit{(Chapter III)}

3) If contrast and attentional feedback signals both enhance activity in early visual areas, how do they interact with each other? \textit{(Chapter IV)}

4) How does the glutamatergic system play a role in working memory? \textit{(Chapter V)}

5) How is information about stimuli efficiently transferred across different visual areas? What are the factors that determine if information reaches awareness or remains subliminal? \textit{(Chapter VI)}

The following sections will give necessary background information and introduce each chapter in brief.
1.3 Measuring attention’s molecular foundation

In Chapter II, we investigate the molecular foundation of attentional signalling. Numerous studies\textsuperscript{125,128,144-147} have explored the role of Acetylcholine ($ACh$) as a neuromodulator for attentional signalling. $ACh$ binds to two types of receptors present on neuronal cells, namely nicotinic and muscarinic receptors. Cholinergic signals originate in parts of the basal forebrain such as the nucleus basalis, the septum, the substantia innominata and the diagonal band of Broca\textsuperscript{80-82}, which contain several cholinergic nuclei. In addition, $ACh$ is produced in some midbrain nuclei, called the pedunculopontine nucleus and laterodorsal tegmental area\textsuperscript{83}. In contrast to its localised production, the effects of acetylcholine on brain networks are widely distributed\textsuperscript{144,148}. Almost all regions of the brain are innervated by cholinergic neurons and many neurons also express $ACh$ receptors\textsuperscript{134,149}. This is also supported by histological studies\textsuperscript{150} that observe a high density of nicotinic and muscarinic receptors in the cortex of humans, rhesus monkeys, cats, rodents and other mammals.

Consequently, $ACh$ and the cholinergic system have been implicated in several cognitive functions such as attention\textsuperscript{125,128,144-147}, learning\textsuperscript{150-153} and memory\textsuperscript{154-157}. Several studies\textsuperscript{145,158,159} have tried to measure the amount of $ACh$ that is released during such shifts of attention. However, due to technical and practical reasons several of these techniques such as microdialysis\textsuperscript{160} can only measure the average level of $ACh$ across a period of several seconds to minutes. However, recent advances in biosensors have allowed the construction of sensors that can measure levels of acetylcholine present in the brain at a very fast temporal resolution (<1 second). In the present study, we tried to adapt the technique developed for the rodent by Parikh\textsuperscript{158} and colleagues for use in the non-human primates. We aimed to measure acetylcholine levels in an awake behaving monkey while the monkey performed a selective attention task. We report that our attempts to measure acetylcholine levels using amperometry proved difficult and tedious. Our results were inconsistent and prone to noise. In this chapter, we have outlined the challenges that need to be addressed before this technique can be utilized robustly. 
1.4 Attention and learning

In our everyday life, we encounter many tasks that require us to learn by trial and error. When humans or animals encounter a novel problem, they will try an action and if that action results in a positive outcome such as a treat or another reward, the same action will be repeated again. However, if the outcome is negative such as a punishment, they will be hesitant to try the same action again. In short, we often learn from positive and negative experiences.

Learning induces connectivity changes\textsuperscript{17,52,53,138,139,161,162} in the brain. On a psychophysical and behavioural level, learning is reflected by improved task performance\textsuperscript{163–165} such as a lower response (reaction) time, increased accuracy, and enhanced sensitivity post learning. Two factors are known to influence the ability to learn and induce connectivity changes: attention\textsuperscript{14,17,18,136–139,141,166} and reward\textsuperscript{167–174}. Their role in learning has garnered interest and has lent support for reinforcement learning\textsuperscript{175,176} models that describe learning rules for biological/artificial systems. Roelfsema\textsuperscript{16,20,166} and colleagues conceived a computational model known as ‘attention-gated reinforcement learning’ describing the role of attentional-feedback signal in learning. The model proposes that feedback signals corresponding to attention highlight, tag, and strengthen relevant patterns of activity generated during the input phase possibly via cortico-cortical connections (see 1.1e). Additionally, ‘neuromodulators’ are released that signal whether the outcome of the action is better or worse than expected (i.e. a ‘reward prediction error’, \textit{RPE}). The idea that neuromodulators such as dopamine\textsuperscript{170,173} and acetylcholine\textsuperscript{147} code for the \textit{RPE} has been supported by studies in animals. Schultz and colleagues\textsuperscript{170,173} demonstrated that dopaminergic neurons (\textit{that release dopamine}) code for \textit{RPEs}\textsuperscript{177}. During a learning task neurons are initially activated by unexpected rewards, but as the learning progresses they are activated by stimuli that predict that reward will be obtained later in the trial\textsuperscript{170,173}. Furthermore, Hangya\textsuperscript{147} and colleagues have observed that cholinergic neurons are sensitive to \textit{RPE}.\textsuperscript{147}
Accordingly, reinforcement learning theories\textsuperscript{171,174,175} and models\textsuperscript{16,20,166} of learning have proposed that selective attention and reward prediction errors can help encode powerful learning rules\textsuperscript{16,20,166}. However, it is not yet clear how these factors together contribute to the learning of new stimulus-response associations. In Chapter III, we address the conjoint influence of attention and reward on the learning of new stimulus response associations with novel relevant-redundant cue-learning paradigm. The paradigm tested attended and unattended shapes associated with different quantities of reward, allowing us to study the effect attention and reward and their interaction during learning. (See Chapter III for more information)

### 1.5 Spread of attention

Roelfsema\textsuperscript{3,10,43,44,120} and colleagues have suggested that an ‘incremental-grouping process’ is required to help label grouped image elements with object-based attention. This labelling process allows for grouping of novel features albeit at the cost of processing time (see 1.1e). They have investigated the spread of object-based attention using the curve-tracing task in the early visual cortex and their findings strongly support that curve tracing depends on the propagation of enhanced neuronal activity resulting from attentional feedback processes. However, attention is not the only factor that influences neuronal activity in the early visual cortex; almost all neurons in the visual cortex strongly respond to changes in the luminance contrast (i.e. difference in the amount of light reflected by two or more objects) of stimuli, with higher contrast stimuli generating stronger activity compared to low contrast stimuli.

In that case, is it possible that the luminance contrast and perceptual grouping interact with each other? For example, would it be more difficult to trace a low-contrast curve in presence of a high-contrast curve? Previously, three models (Fig. 14) have described the interaction between neuronal activity evoked by luminance contrast and by attention. This can be observed in the ‘Contrast response function’ (Fig. 14) which depicts the relationship between changes in
contrast and neuronal activity. The ‘contrast gain’\textsuperscript{178,179} model posits that attention increases the neuronal response to the stimuli by increasing its apparent contrast and that this effect is strongest at intermediate contrast levels and weaker at high contrast levels (Fig. 14 –left). Another model called the ‘response gain’\textsuperscript{180} model posits that attention and contrast interact multiplicatively (Fig. 14 –centre). Consequently, attention has a stronger effect at higher-contrast levels. A third model called the ‘additive’ model\textsuperscript{182} (Fig. 14 –right) suggests that attention and luminance contrast influence the neuronal response to the stimuli in a fairly independent manner.

![Figure 14: Contrast response function of three models describing contrast and attention interactions. (from Chapter IV, Fig. 2 published as Watson et. al., 2013)](image-url)

In Chapter IV, we asked which model best described the relationship between attentional modulation and contrast-based modulation during a contour grouping task. First, we measured the effect of luminance contrast on the speed of contour grouping using different contrast levels. Second, we tested if subjects would be able to efficiently trace a low contrast target curve in presence of a high contrast distractor. Third, we tested if contrast polarity affects the speed of contour tracing. We found that the speed of the spread of attention over the curve did not depend on contrast and it also did not depend on the contrast polarity. Moreover, tracing of the low contrast curve was not hampered by the presence of a
high contrast distractor. Overall, our results support the ‘additive model’, which implies that object-based attention and contrast have largely independent effects on visual processing. (See Chapter IV for more information)

1.6 Working memory processing

In Chapter V, we investigate how glutamate receptors contribute to working memory. Working memory refers to the ability to store, albeit with a limited capacity, and manipulate information over short periods of time, generally lasting only a few seconds. Working memory has been strongly linked to attention, it is postulated that attentional selection acts as a ‘gatekeeper’ for working memory, which ensures that only relevant information is encoded into working memory.

Working memory processing is thought to be due to ‘persistent activity’ of cells (i.e. the cells continue to respond even long after the visual stimuli has disappeared). The memory remains active only as long as these cells maintain their activity. Consequently, persistent activity of these cells is thought be the neural substrate for working memory. Interestingly, studies have found cells that show persistent activity in areas such as the dorsolateral prefrontal cortex.

Although the exact mechanism for working memory is only partially understood, a variety of neuromodulatory systems—dopaminergic, cholinergic and norepinephrinergic—have been implicated in working memory processing.

We investigated the role of the glutamatergic system, the major excitatory neurotransmitter of the brain. Previous work on the function of glutamatergic receptors has suggested that ionotropic glutamate receptors: NMDA and AMPA receptors contribute differentially to feedforward and feedback processing in tasks such as figure-ground perception i.e. distinguishing between figure and background. The pattern of differential activity observed between figure and background (figure-ground modulation) resembles the extra activity observed in contour grouping (attentional modulation) (see 1.1e, Fig. 13). Self and colleagues report-
ed that in area V1, AMPA receptors contribute during early (feedforward) phase of the response whereas NMDA receptors contribute during a later, sustained phase of the response in which figures elicited more activity than backgrounds. In a similar fashion, it is hypothesised that AMPA and NMDA receptors may contribute differentially to sustained activity seen during working memory processing.

Such a different role would also be in accordance with the intrinsic cellular dynamics and receptor kinetics of AMPA and NMDA. When glutamate binds to AMPA receptors, they activate rapidly in contrast to NMDA receptors whose activation is slower. Modelling studies postulate the slow activation and decay of NMDA allows for a mechanism to maintain persistent activity. In this study, we compared the contribution of AMPA and NMDA receptors to persistent activity in the dorsolateral prefrontal cortex of macaque monkeys performing a working memory task. We ejected small amounts of selective antagonists for AMPA and NMDA allowing us to differentiate between contributions of the receptors during the different stages of the task. (See Chapter V for more information)

### 1.7 Recurrent processing in conscious awareness

Finally, in Chapter VI, we explore the role of recurrent (feedforward-feedback) processing in conscious awareness and perception of objects. An influential computational and theoretical framework called the Global Neuronal Workspace (GNW) aims to link psychophysical and neurophysiological findings on conscious awareness.

Dehaene and colleagues have modelled the GNW framework with computer simulations. They ascribe an important role to recurrent (feedforward-feedback) processing as seen between lower and higher areas of the visual hierarchy. The framework describes that the conscious access to the stimulus relies on two phases: the first phase is that the stimulus propagates along the visual hierarchy similar to a feedforward sweep (bottom up), possibly enabled by AMPA receptors. In the second phase, if the stimulus is selected with respect to the attentional state or goal, it is further amplified and maintained with sustained activity.
(similar to working memory). This sustained activity is called ‘conscious ignition’ or ‘global ignition’ where the representation of the stimuli is spread across cortical processors and networks. By enabling the selection of relevant information from the input signals during second phase selective attention gates access to the conscious/global workspace, which enables ‘global ignition’. However, if the stimulus processing is interrupted either during the feedforward sweep or during feedback the stimulus fails to evoke ‘ignition’, and remains subconscious\(^{200}\).

Previous studies\(^{119}\) demonstrated that the attentional blink and masking (where a discernible ‘visible’ stimuli consistently misses detection by participants) can interfere with gaining access to the global workspace. However, the factors that determine efficient activity propagation towards the global workspace and the level where activity propagation fails for those stimuli that stay subliminal still remains to be explored\(^{12,201}\). An effective approach to investigate this and study the neuronal correlates of awareness, is to hold the sensory input constant while perception is allowed to fluctuate and to then compare neuronal activity between perceived stimuli and identical but non-perceived stimuli. In our study, we investigated neuronal activity elicited by weak stimuli at different stages along the visual cortical hierarchy (area V1, area V4 and the dorsolateral prefrontal cortex) to determine where the information pertaining to subliminal stimuli is lost. *(See Chapter VI for more information)*
References


75. Miller, E., Erickson, C. & Desimone, R. Neural mechanisms of visual working


82. Mesulam, M. M. *Structure and function of cholinergic pathways in the cerebral cortex, limbic system, basal ganglia, and thalamus of the human brain.* (Raven Press, 1995).


temporal visual areas of the macaque monkey. *J. Neurosci.* 6, 44 (1986).


Buschman, T. J. & Miller, E. K. Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron* 63, 386–396 (2009).
149. Van Der Zee, E. A. & Luiten, P. G. M. Muscarinic acetylcholine receptors in the hippocampus, neocortex and amygdala: A review of immunocytochemi-


187. Sawaguchi, T. Attenuation of Delay-Period Activity of Monkey Prefrontal Neurons by an α 2-Adrenergic Antagonist During an Oculomotor Delayed-