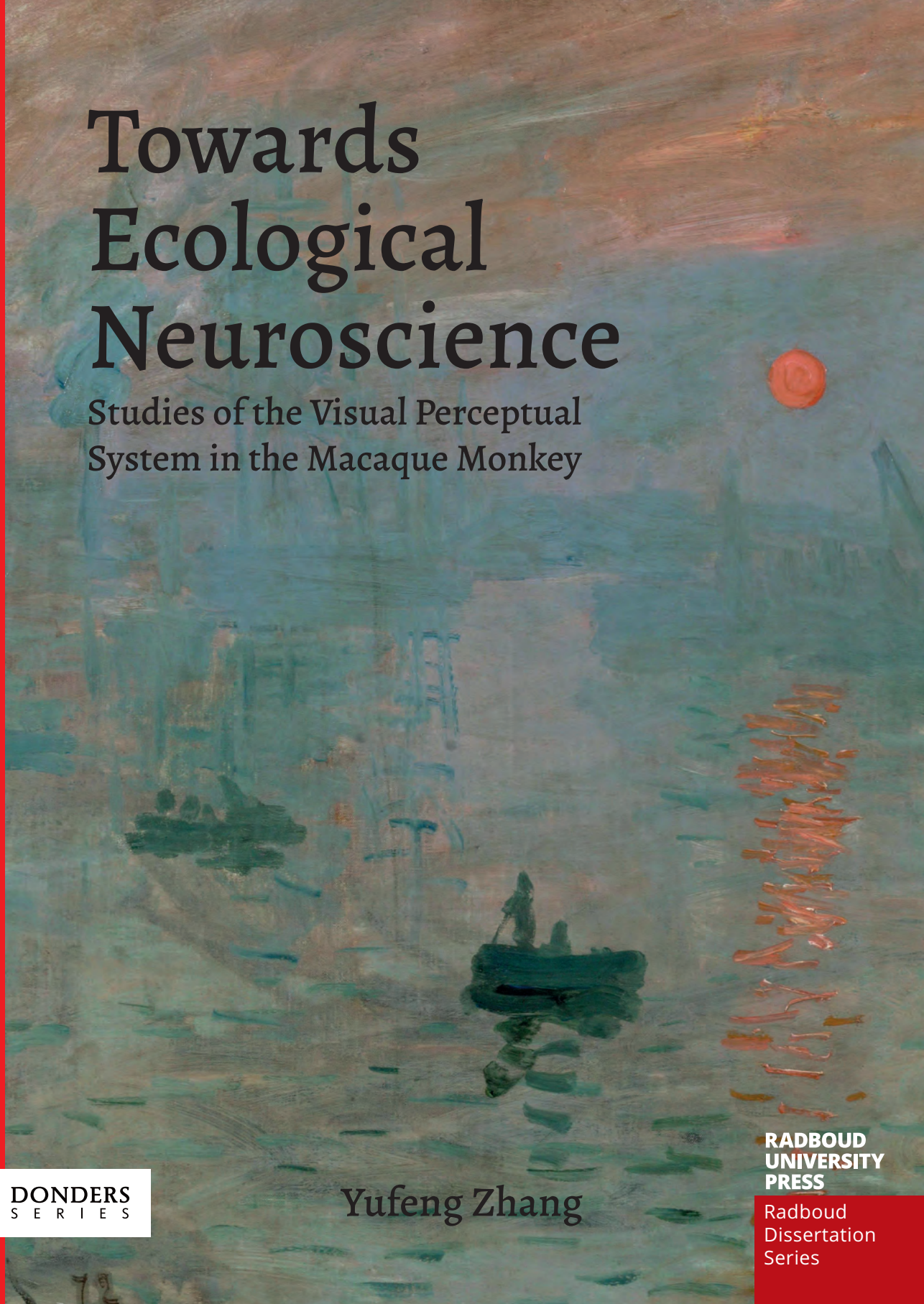
Towards Ecological Neuroscience

Studies of the Visual Perceptual System in the Macaque Monkey



Yufeng Zhang

DONDERS SERIES

RADBOUD UNIVERSITY PRESS

Radboud Dissertation Series

Towards Ecological Neuroscience

Studies of the Visual Perceptual System in the Macaque Monkey

Yufeng Zhang

The research presented in this thesis was conducted at the Ernst Strüngmann Institute for Neuroscience in Cooperation with the Max Planck Society (ESI) in Frankfurt am Main, Germany.

Towards Ecological Neuroscience Studies of the Visual Perceptual System in the Macaque Monkey
Yufeng Zhang

Radboud Dissertations Series
ISSN: 2950-2772 (Online); 2950-2780 (Print)

Published by RADBOUD UNIVERSITY PRESS
Postbus 9100, 6500 HA Nijmegen, The Netherlands
www.radbouduniversitypress.nl

Design: Yufeng Zhang
Cover image: Reproduction of *Impression, Sunrise* by Claude Monet, public domain
Printing: DPN Rikken/Pumbo

Chapter 1 of this thesis contains figures reproduced with permission from previously published works, for which the copyrights remain with their respective owners.

ISBN: 9789465150666
DOI: 10.54195/9789465150666
Free download at: https://doi.org/10.54195/9789465150666

RADBOUD
UNIVERSITY
PRESS

Towards Ecological Neuroscience

Studies of the Visual Perceptual System in the Macaque Monkey

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.M. Sanders,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

maandag 31 maart 2025
om 12.30 uur precies

door

Yufeng Zhang
geboren op 28 januari 1992
te Shanxi, China

Promotor:
prof. dr. P. Fries

Copromotor:
dr. C.M. Lewis (Universität Zürich, Zwitserland)

Manuscriptcommissie:
prof. dr. M.V. Peelen
prof. dr. P. König (Universität Osnabrück, Duitsland)
prof. dr. G. Gregoriou (University of Crete, Griekenland)

To Jiayue
and our soon-to-greet daughter, Siqi
with love

Summary

There is a discrepancy between how vision is studied in the lab and how vision is used in real life. In real life, an observer actively obtains information about their surroundings by shifting their gaze, turning their head, and changing their point of observation. By contrast, a participant in the lab usually passively receives stimulation that impinges on their retina. This is especially true when neuroscientists study the neural correlates of visual perception using macaque monkeys. More often than not, macaque monkeys are found in the lab head-posted, looking fiercely at an uninteresting fixation dot. This is not the type of vision we aim to understand. Following James J. Gibson, ecological psychologists have long recognized this discrepancy and suggested adopting an ecological approach to address it. Taking an ecological approach means considering the information available to the active observer and the process by which the active observer picks up this information. The work presented in this thesis advocates for a similar ecological approach in neuroscience.

Samenvatting

Er is een discrepantie tussen hoe visie in het laboratorium wordt bestudeerd en hoe visie in het echte leven wordt gebruikt. In het echte leven verkrijgt een waarnemer actief informatie over zijn omgeving door zijn blik te verplaatsen, zijn hoofd te draaien en zijn observatiepunt te veranderen. Daarentegen ontvangt een deelnemer in het laboratorium meestal passief stimulatie die op zijn netvlies terechtkomt. Dit geldt vooral wanneer neurowetenschappers de neurale correlaten van visuele perceptie bestuderen met behulp van makaakapen. Makaakapen worden in het laboratorium vaak head-posted gevonden, waarbij ze intens naar een oninteressante fixatiepunt staren. Dit is niet het soort visie dat we willen begrijpen. Ecologische psychologen, in navolging van James J. Gibson, hebben deze discrepantie al lang erkend en voorgesteld om een ecologische benadering aan te nemen om dit probleem aan te pakken. Een ecologische benadering houdt in dat men rekening houdt met de informatie die beschikbaar is voor de actieve waarnemer en het proces waarmee de actieve waarnemer deze informatie opneemt. Het werk dat in dit proefschrift wordt gepresenteerd, pleit voor een soortgelijke ecologische benadering in de neurowetenschappen.

Acknowledgements

In the first meeting with my supervisor, Pascal Fries, when I expressed that I did not want to have a short PhD, I could not have anticipated just how long this journey would be. Completing this long journey would not have been possible without the tremendous support and encouragement that I have received along the way.

Foremost, I want to thank Pascal for his mentorship and guidance. I owe every aspect of my academic training to him, and the work contained in this thesis is only possible with his valuable feedback. Members of the Fries lab have helped me as much. I learned so much from them in and out of the office. Particularly, I would like to thank Julien Vezoli for his guidance during my naivest early days as a doctoral student and Tim Näher, with whom the collaboration made the later stages of my PhD also exciting.

Much of what I did in my PhD, particularly the practical part, is not reflected in this thesis. I value them as much as the work presented in this thesis. This includes many cool projects, e.g., developing insertion tools for flexible electrodes, designing the various recording chambers, and, of course, the Kuka-robot to help with trepanations. For these projects, Rasmus Roese is almost always part of it as a source of inspiration and creativity.

I would also like to thank my co-supervisor, Christopher Lewis, and the members of my thesis advisory committee, including Prof. Dr. Manfred Kössl and Prof. Dr. Moritz Helmstaedter. They have helped me navigate my PhD to reach its end.

Completing a PhD abroad, especially during a three-year pandemic, comes with challenges. I am particularly thankful to my Chinese colleagues, Jianguang Ni, Yiling Yang, Jinke Liu, and Tianlu Wang, for their camaraderie.

Lastly, I am deeply thankful to my family for their unwavering support and encouragement. Above all, I owe the most to my wife, Jiayue Zheng. As Max Weber said, the academic life is "ein wilder Hazard"——a wild gamble. With her by my side, that gamble was never allowed to turn into harm.

Contents

List of Figures

CHAPTER 1

Introduction

The fundamental goal of visual science, as put by Davida Teller, "lies in the discovering and testing of linking propositions," which links the perceptual states (what people see) to physiological states (of the machinery that enables seeing) in our nervous system [1, 2]. This thesis presents some work using macaque monkeys to study the visual perceptual system. It advocates that the current search for linking propositions may be missing essential elements and requires an ecological approach [3]. Specifically, the perceptual and physiological states intended to be linked by the linking propositions shall be the natural states in the nervous system when the animal actively explores the environment that surrounds it (Figure 1.1).

This thesis suggests that an animal's active exploration is characterized by sequences of rhythmic sensorimotor decisions, which I call the action-perception loop (the AP-loop). The AP-loop provides the natural behavioral context in which the animal's visual perceptual system picks up information in the ambient light. Adopting an ecological approach involves providing information to the

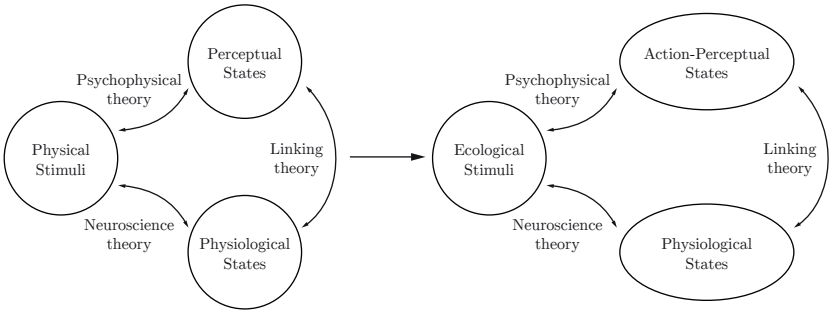


Figure 1.1: Linking theories. Davida Y. Teller divided the field of vision science into three types of entities and three corresponding mapping theories (Left). The fundamental goal of visual sciences is, thus, discovering and testing the linking theories that link an observer's perceptual states to their physiological states. The current thesis suggests that these states intended to be linked should be the natural states as the observer actively explores their environment (Right). Specifically, the physical stimuli should be the ecological stimuli representing the kind of stimulation the observer obtains in real life; the perceptual state should be the action-perceptual state as the observer engages in the AP-loop. The left panel is reprinted from Teller and Palmer [1] with permission.

animal, allowing the animal to actively explore this information with its eyes, head, and body, and identifying the persistence and changes in the nervous system as the animal engages in the AP-loop. The ensuing research chapters represent efforts in these directions.

The introduction begins with a presentation of relevant results from the neuroanatomy and neurophysiology literature. Due to my restricted reading, this presentation is admittedly limited and biased. Despite these shortcomings, I hope the readers can still appreciate the overall structure and, most importantly, the necessity of incorporating action into the study of perception. Well-informed readers may opt to skip this section.

The introduction then moves to the concept of the AP-loop and discusses its role in natural vision. This discussion sets the stage for the ensuing discussion on the subject matter of ecological visual neuroscience. At the end of this introduction, I overview the subsequent research chapters and explain how their respective results relate to the advocated ecological approach.

1.1 Context of Development

1.1.1 Neuroanatomy: The Structural Basis of Vision

The structural basis of vision is neuroanatomy. Modern investigations of neuroanatomy are often traced back to the pioneering work of Camillo Golgi and Santiago Cajal [4]. Golgi's discovery of the *Black Reaction* (reazione nera) made it possible to show the entire morphology of a neuron, including the detailed structure of its axons and dendrites. By applying and improving Golgi's technique, Cajal was able to describe the cell types and organizational principles of avian and mammalian retinas. Following Golgi and Cajal, researchers using essentially the same technique described cell types in the visual pathway beyond the retina. For example, Lund [5] described the neuronal organization of the macaque primary visual cortex and identified three basic cell types, each with distinct morphological features and laminar distributions (Figure 1.2 A). These three basic cell populations are spiny pyramidal cells, spiny stellate cells, and smooth non-spiny cells. Later immunocytochemical and physiological studies showed that these morphologically defined cell types also serve distinct functions in neural circuits: smooth, non-spiny cells are GABAergic inhibitory neurons, while spiny cells are excitatory neurons. Identifying such morphological and functional correspondences at the single-neuron level has been a lasting research topic that still interests scientists today [6].

Golgi and Cajal's work laid the foundation for the neuron doctrine [7]. The neuron doctrine states that our brain is composed of individual elements called neurons, and these neurons are the structural and functional units of the nervous system. Although our understanding of the functional units in the brain is shifting towards populations or networks of neurons, the neuron doctrine remains the most basic structural description of our nervous system [8]. As characterizing cell types continues to be a focus of modern neuroscience, the definition of a cell type has evolved with advancements in techniques to include not only morphological features but also physiological, immunochemical, and even transcriptomic data [6].

While the Golgi technique and alike have been instrumental in identifying cell types and hypothesizing about intrinsic connectivity on the basis of the observed distribution of axons and dendrites, it has limitations in revealing how distant neurons, not visible in a single section, are interconnected. Identifying such

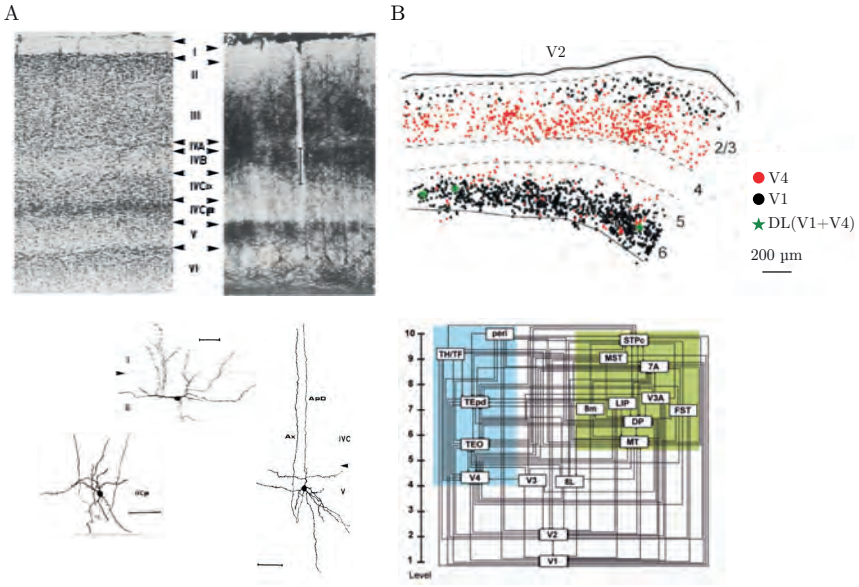


Figure 1.2: Neuroanatomy reveals the structural basis of vision. (A) The development of staining methods made it possible to visualize the entire morphology of a neuron, including the detailed structure of its axons and dendrites relative to the cortical lamination. The top panel shows the Nissl (left) and Golgi (right) staining of the macaque primary visual cortex. The bottom panels show example neurons from Golgi staining. Left: a spiny stellate neuron in layer IV. Middle: a smooth stellate neuron in layer II. Right: a spiny pyramidal neuron in layer V. Scale bar: 50 μm. (B) The application of tracers allows the mapping of long-tract connectivity. The resulting inter-area connectivity, in combination with the projections' laminar pattern, can be used to deduce a hierarchical model. The top panel shows the retrogradely labeled neurons in area V2. The feedforward projecting neurons (that project to V4) are predominantly located in the supragranular layers. In contrast, the feedback projecting neurons (that project to V1) are predominantly located in the infragranular layers. The bottom panel shows the deduced visual hierarchy. Panel A is reprinted from Lund [5] with permission from John Wiley and Sons. Panel B is reprinted from Markov et al. [9] under the terms of the Creative Commons license (`CC BY`).

long-tract connectivity in the visual system is essential for understanding how its different parts function together as a cohesive whole. The development of degeneration-based stains and axoplasmic transportation-based tracers provided this information. Notable achievements based on these long-tract connectivity data include the description of parallel pathways and visual hierarchy.

Three major concurrent pathways are identified with long-range connectivity data from the retina to the primary visual cortex: the M, P, and K pathways [10]. The names of M, P, and K pathways are derived from the cell populations of LGN. The primate LGN is organized into six layers: the two ventral magnocellular layers (Layers 1 and 2) and the four dorsal parvocellular layers (Layers 3 to 6). Anterograde and retrograde labeling have identified the input source of these layers in the retina and their projection targets in the primary visual cortex [11]. The magnocellular layers receive input from parasol ganglia cells in the retina and project to the primary visual cortex layer IVα. The parvocellular layers receive input from midget ganglia cells in the retina and project to layer IVβ of the primary visual cortex. More recently, a third major neuronal population immunochemically stained for the alpha subunit of type II calmodulin-dependent protein kinase was found below each of the six LGN major layers, forming the koniocelluar (K) [12]. Neurons in the K layers receive input from bistratified neurons in the retina and project to the superficial layers corresponding to the CO blobs. The separation of these three pathways starts as early as the first synapse after the photoreceptors: the parasol ganglia cells in the M pathway receive input from the diffuse bipolar cells, whereas the midget ganglia cells receive input from the midget bipolar cells. The segregation of visual streams continues in cortical processing. Specifically, the M stream is thought to dominate the dorsal "where" pathway, and the P and K streams are thought to dominate the "what" pathway [13]. Cross-talks between the two streams are abundant at each level of cortical processing.

In the neocortex, inter-area connectivity data laid the foundation for hierarchical models of visual processing. Rockland and Pandya [14] classified inter-area projections into feedforward (FF) and feedback (FB) connections based on their laminar origin and termination patterns. They found FF projections typically originate in the supra-granular layers and terminate in layer 4, while FB connections usually originate in the infra-granular layers and terminate outside layer 4. Using this classification, Felleman and Van Essen [15] summarized inter-area connectivity data and proposed the first hierarchical model of the visual system. They organized 32 visual areas into ten hierarchical levels of cortical processing. Because this model was based on qualitative data, the placement of visual areas on the hierarchy was indeterminate [16]. Despite this, Felleman and Van Essen's model marked a significant step towards understanding the neural basis of vision at the system level. Recently, Markov et al. [9] revised Felleman

and Van Essen's model through quantitative analysis of retrograde tracing data (Figure 1.2 B). Their model largely preserved the original hierarchical order but revealed essential features through quantitative retrograde tracing. They found that FF projections are not only predominant in the supragranular layer, but this supragranular bias increases with larger hierarchical distances, as measured by the percentage of supragranular labeled neurons (SLN). Similarly, FB projections, which preferentially originate in the infragranular layers, show a stronger infragranular bias with larger hierarchical distances.

As morphologically different cell types exhibit unique physiological properties and play different roles within the neural circuit, the visual brain's hierarchical structure has functional correlates [17]. Foremost, the serial, feedforward aspects of the hierarchy have been associated with the increasingly complex feature selectivity along the visual pathway. Pioneering works of Hubel and Wiesel applied the idea of feedforward processing to explain the emergence of orientation selectivity in the primary visual cortex. They suggested that the simple cells' orientation-selective receptive fields result from aligning input from multiple circular ON (or OFF) centered LGN neurons [18]. This hypothesis has later received direct physiological and anatomical evidence [19, 20]. Following these studies, various types of feature selectivity have been identified beyond the primary visual cortex, e.g., the selectivity for color and curvatures of V4 neurons [21], motion directions in MT [22], and the view selectivity in the inferior temporal cortex [23]. Such functional mapping along the visual hierarchy has been fruitful, and the idea that visual features are analyzed progressively in a distributive fashion is generally accepted as the starting point for explaining more complex visual phenomena, such as object recognition, decision-making, visual search, attention, and working memory. The power of converging feedforward connections is presumably best demonstrated not in biological visual systems but in the ability of similarly structured convolutional neural networks to solve complex visual tasks [24].

In addition to converging feedforward pathways, there are about twice as many feedback pathways in the visual hierarchy [25, 26]. Anatomically, compared to feedforward pathways, these feedback pathways: 1) have distinct laminar profiles, 2) are more divergent in their projections, and 3) can travel longer distances. Correspondingly, feedback pathways are also physiologically distinct from feedforward pathways. While feedback pathways can be driving, as shown

by phenomena like imaginary figures, they are often described as modulatory or context-dependent. The term "modulatory" does not imply that feedback connections are less important than feedforward connections; rather, it indicates their complexity [27]. There is growing interest in feedback projections, which are associated with functions that cannot be explained solely by feedforward processing. A prominent example is the study of covert attention, the selectively enhanced response to the attended stimulus. Feedback projections from frontal, parietal, and higher sensory areas have been found to enhance the processing of attended objects [28].

A particularly interesting development is the attempt to link hierarchical cortical processing to the constructive nature of visual perception within the predictive coding framework [29, 30]. Predictive coding assumes a hierarchical generative model from causes to sensory data, which the brain inverts to infer causes from sensory data by minimizing the free energy (or surprises). This computation requires updating predictions at each hierarchical level using the prediction error signal passed from the level below (bottom-up, feedforward), and updating prediction errors at each hierarchical level using predictions generated at the level above (top-down, feedback). Additionally, it involves separating units that represent predictions and errors, and subtracting prediction signals from higher levels. These computational requirements align well with the anatomical hierarchy of the cortex. Proposals have thus assigned feedback projections the role of broadcasting predictions generated in higher visual areas and feedforward projections the role of transmitting unexplained prediction errors. This mapping between computational and anatomical architecture was suggested to go even further with different cell populations and their intrinsic connectivity (the "canonical microcircuit") mapped to separate computational roles for implementing the predictive coding [31]. Such an elegant structure-functional correspondence suggests an asymmetry of frequency composition for feedforward and feedback signals. Because predictions are updated by accumulating prediction errors, high-frequency components in the feedforward prediction error signal are expected to be attenuated in the prediction signal that feeds back. A recent study shows that such asymmetry does exist [32]. Importantly, the suggested physiological asymmetry not only exists but also correlates with the anatomical hierarchy and serves the expected function.

1.1.2 Physiology: From Anatomy to Function

I now move gears to physiology, which brings vitality to the relatively static anatomical structure. The most crucial way neurons communicate with each other is through action potentials, colloquially known as spikes for their shapes when the fluctuating voltage is plotted against time. Just as staining and labelling techniques advanced neuroanatomy, progress in neurophysiology was enabled by the development of methods to detect spikes and their correlated signals.

Single Site Recordings: Cortical Columns and Noisy Neurons

Since the description of the biophysics underlying action potential generation, refinements in micro-electrodes have led to the proliferation of single-unit recordings in the visual system [33]. These recordings, predominantly obtained since the latter half of the 20th century, have been instrumental in unraveling the neural mechanisms of vision. The seminal work of Hubel and Wiesel on the functional architecture of the primary visual cortex, using deeply anesthetized cats and monkeys with tungsten microelectrodes, stands out as a landmark of this period (Figure 1.3 A).

The most influential idea put forward by Hubel and Wiesel is that the primary visual cortex is organized into functional columns [18, 34]. A cortical column, first termed by Mountcastle [35] based on his recordings from cat somatosensory cortex, refers to a periodically occurring, vertical cortical subdivision, wherein neurons, across the layers, show similar selectivity for a given receptive field attribute. Practically, this means if one records the neuronal activity with a single contact microelectrode, 1) neurons encountered in an electrode penetration vertical to the cortical surfaces display similar selectivity, for example, all respond strongest to a 1hr-7hr oriented bar compared to other orientations; 2) neurons encountered in an oblique electrode penetration, close to tangential to the cortical surface displays a periodically varying selectivity, for example, as the electrodes advance the neurons' preferred orientation moves gradually from 1hr-7hr to 5hr-11hr and eventually back to 1hr-7hr (Figure 1.3 B).

Several highly organized, periodic columnar systems have been identified in the primary visual cortex. These include orientation columns or pinwheels, ocular dominance columns, and cytochrome oxidase (CO) columns. These columnar

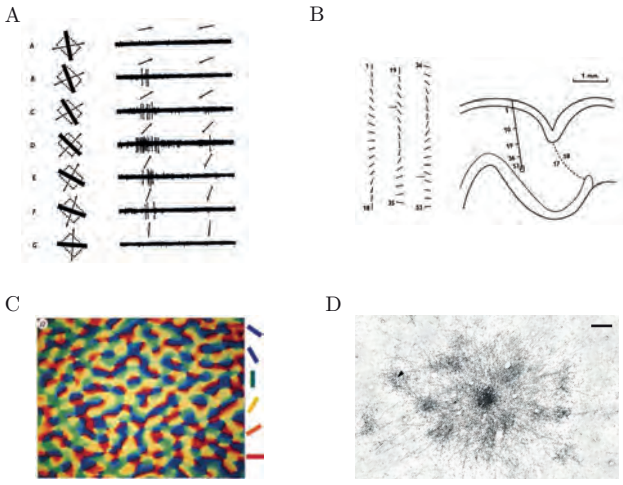


Figure 1.3: The orientation columns. (A) An example orientation-selective complex cell in macaque primary visual cortex recorded with a tungsten microelectrode. (B) The orientation selectivity changes gradually and periodically in an oblique electrode penetration. (C) The periodicity of orientation selectivity is visualized with voltage-sensitive dyes. (D) Anterograde tracing with biocytin reveals a patchy, start-like pattern of horizontal connections biased for similarly tuned orientation-selective neurons. Scale bar: 200 μm. Panel A and B are reprinted from Hubel and Wiesel [34] with permission from John Wiley and Sons. Panel C is reprinted from Blasdel and Salama [36] with permission from Springer Nature. Panel D is reprinted from Sincich and Blasdel [37]; copyright 2001 Society for Neuroscience. No permission required as per *JNeurosci* reuse policy.

systems were initially revealed through physiological or anatomical studies but are most clearly visualized using optical imaging methods (Figure 1.3 C) [36]. Within each system, columns are retinotopically organized and form a tiling of the visual space with their aggregated receptive fields. Connections among the columns are not random. Microinjection of tracers has shown that orientation columns are preferentially connected with other orientation columns with similar selectivity, resulting in a patchy, star-like pattern of horizontal connections (Figure 1.3 D) [37, 38]. The layouts of these columnar systems are related to each other. For example, cytochrome oxidase patches and orientation pinwheel singularities are preferentially located in the centers of ocular dominance bands. Iso-orientation contours are more likely orthogonal to the ocular dominance boundary, presumably to achieve a uniform coverage of stimulus feature combinations [39].

A key feature of these columnar systems is their uniform distribution across the retinotopically organized cortical surface. The periodicity of orientation pinwheels, ocular dominance columns, CO patches, and inter-patches all appear to be around 1∼2 mm across the primary visual cortex. This uniformity is intriguing, especially given the importance of the fovea in primate vision. Unlike the retina, where the fovea has a specialized central pit with densely packed photoreceptors and optimized connectivity, the cortex shows homogeneity in its columns. This suggests that the visual cortex employs a distinct strategy to emphasize the foveal processing. Instead of creating a specialized processing unit for the fovea, it allocates more of the same processing units to it. One advantage of this strategy is that, while the spatial extent of the fovea in the retina is dictated by optics, the cortex is free to allocate as many processing units (surface area) to the fovea as is ecologically optimal. Such preferential resource allocation is known as foveal magnification. Another advantage of this strategy is that it can be applied repeatedly in different visual areas, regardless of their specialized functions. Notably, the foveal magnification, in combination with the increasing receptive field sizes from the fovea to the periphery, results in a persistent cortical point image (the area of activated cortical surface), independent of the stimulus's retinotopic location (given that the stimulus is smaller than the corresponding receptive field size). As will be discussed later, understanding such persistence is particularly interesting for the envisaged ecological neuroscientists.

Columnar systems, like those observed in the primary visual cortex, have also been identified in the extrastriate visual cortex. For instance, V2 contains the periodically occurring thin-pale-thick cytochrome oxidase stripes, implicated in the preferential processing of form, color, and motion/disparity signals; V4 houses separate domains preferentially processing color, orientation, and curvature; and MT is known for its direction and disparity columns [40–45]. Adopting a looser definition for a cortical column, one that does not require discrete boundaries and periodic occurrence, reveals columnar-like organizations throughout the visual hierarchy from the base level, i.e., the primary visual cortex, up to the higher end, both in the ventral stream, e.g., the inferior temporal cortex (IT) where there are columns for object features including the well-known face patches, and the dorsal stream, e.g., the frontal eye field where saccade amplitudes are topographically represented and the saccade direction changes smoothly but shows occasional reversals similar to how the preferred

orientation changes in primary visual cortex [46–48].

In addition to the functional mapping across various levels of the visual hierarchy, single-unit recordings started to uncover the intrinsic dynamics of neural processing. Below, I introduce two basic dynamics relevant to the ensuing research chapters.

The first is the transient-decay dynamics following a stable stimulus presentation lasting hundreds of milliseconds. Rather than simply summing sensory input over time, as modeled by convolving a constant input signal with an impulse response, visual neurons display a non-linear time course. This includes a strong initial transient activity followed by a weaker sustained activity. The initial transient activity likely reflects the rapid feedforward processing [49, 50]. This is evidenced by lower visual areas closer to the retina exhibiting short latencies and *vice versa*. Functional mappings mentioned earlier often rely on the feature selectivity within these initial transients. Notably, within these transients, feature selectivity is already present in the very first few spikes. In contrast, the later, sustained response reflects recurrent processing involving feedback and horizontal connections. Signals in the sustained response are often associated with context-dependent cognitive factors such as visual attention and figure-ground segregation [51, 52].

Yan, Zhaoping, and Li [53] demonstrated the deferential role of the transient and the sustained responses. They showed that the early transient component reflects the relationship between receptive field stimulus orientation and the neuron's preferred orientation. Meanwhile, the sustained component reflects the task-relevant orientation contrast, independent of the receptive field stimulus orientation. Notably, this task-relevant signal was only found after the training.

Computationally, the transient-decay dynamics can be modeled by a delayed normalization model where the output of a linear-nonlinear process is divided by a low-passed, delayed copy of itself [54, 55]. This normalization model is closely linked to the predictive coding framework introduced earlier, where the delayed copy can be interpreted as a prediction that suppresses the feedforward signal [56]. Chapter 4 of this thesis uses the functional segregation of transient and decay components to study separately the contributions of stimulus-dependent initial responses and attention-related delayed responses to saccadic eye movements.

The second is repetition suppression (RS) [57]. RS describes the decreased firing rate upon repeated stimulus presentation. RS is observed in many brain areas and across different experimental conditions. Here I tentatively classify RS into three types based on the suppression's temporal dynamics. The fastest type, rapid RS, happens with an inter-stimulus interval (ISI) shorter than $\approx$ 500ms [58]. Within this range, the response magnitude to the second stimulus gradually recovers, and it fully recovers to the magnitude evoked by an isolated stimulus with ISI longer than 500ms. The above-mentioned delayed normalization mechanism describes the temporal dynamics of rapid RS reasonably well [59]. The second type, short-term RS, happens with an ISI of a few seconds. Typical short-term RS can be seen in a delayed matching to sample (DMS) task, wherein the test stimulus is presented a few seconds after the sample stimulus, and it could either be a match (repetition of the sample stimulus) or a non-match. Short-term RS is manifested in a reduced firing rate to the match stimulus compared to the sample stimulus. Notice that in a DMS task, ISI is often well beyond the range of ISI for rapid RS, and the observed short-term RS can even survive intervening stimuli being presented between the sample and match stimuli [60]. These two observations preclude the possibility that short-term RS shares identical neuronal mechanisms with rapid RS. The third type, long-term RS, manifests itself in a reduced firing rate after tens or hundreds of repetitions [61, 62]. Several features of long-term RS suggest a separate underlying neural mechanism compared to short-term RS: 1) The long-term RS's effect is additive to the short-term RS's effect. In a session of repeated DMS trials, the spiking activity to the sample stimuli is reduced over a session demonstrating the long-term RS; on top of this reduced firing rate over the session to the sample stimulus, in each trial, the matching test stimulus evokes even fewer spikes than the preceding identical sample stimulus demonstrating the short-term RS. 2) The magnitude of suppression increases with the number of repetitions but quickly saturates after dozens of trials, showing distinct dynamics over a few minutes. 3) In addition to the reduced firing rate, long-term RS is accompanied by enhanced gamma-band neuronal synchronization in lower visual areas. Gamma enhancement is particularly evident with many repetitions when the RS in firing rate is already saturated. Repetition-related gamma enhancement is observed in both passive and demanding tasks, with both artificial and parametric stimuli, in both monkeys and human participants. Chapter 5 of this thesis focuses on the characteristic dynamics of this robust repetition-related gamma change. It explores the possibility that gamma-band

neuronal synchronization tracks the change of stimulus occurrence frequency in the environment.

These two dynamics relate to the instantaneous firing rate. The instantaneous firing rate is the expected neural response, estimated by averaging neuronal responses over many trials. However, this definition overlooks an important aspect of neural response: the precise spike timing in each trial. This omission is considered acceptable if one assumes that precise timing carries little information and that the apparent irregularity in inter-spike intervals is due to intrinsic noise. There is evidence supporting this assumption. For example, the statistical properties of a spike train can often be well described by a Poisson model with a Fano factor close to one [63]. However, it is important to note that a stochastic spike train does not imply that: 1) the spike generation mechanism is noisy, nor 2) variability in spike counts or timing is simply noise and contains no information.

Indeed, Mainen and Sejnowski [64] showed that the output spike train is reproducible with sub-millisecond precision in response to a repeated current injection pattern (Figure 1.4 A). This means that, despite the apparently noisy spike train, the spike generation mechanism is very reliable and that spike timing itself can encode information. Note that spike timing, as a potential mechanism for neural coding, may be particularly effective in a noisy background [65]. Consider a simplified scenario: a downstream neuron (B) receiving excitatory input from two upstream neurons (A1 and A2). If A1 and A2 have independent Poisson spiking output with rate λ, neuron B will receive a Poission excitatory post-synaptic potential with the rate of 2λ. Next, we compare this independent firing scenario with a scenario in which A1 and A2 fire synchronously. In the synchronized firing scenario, neuron B would receive the same number of EPSP (same rate of 2λ), albeit in the form of coincident EPSP pairs. Because of the short integration time constant, EPSPs are summed supra-linearly at the post-synaptic neuron [66]. Consequently, coincident EPSPs will activate neuron B more effectively than the same number of EPSPs arriving randomly in time. As a result, neuron B will be driven more effectively in the synchronized firing scenario. The crux is that in this simplified scenario, a Poission-like spike train minimizes the likelihood of coincidence and thus maximizes the effect of coincidence when it indeed happens.

The question is, then, whether the brain indeed takes advantage of a temporal code. This question and alike involve enquiring about the interactions between neurons and can only be addressed directly by recording simultaneously from multiple neurons and preferentially neurons that communicate with each other.

Beyond Single Unit: Synchronization and Correlation

With simultaneous recordings from multiple electrodes or multiple contacts in one electrode, neural activities from multiple neurons or neuronal populations can be picked up simultaneously. Such data affords the investigation of interactions between neurons. Interactions between neurons are commonly measured with their correlations. In the previous section, I contrasted two different neural codes, the rate-code versus the temporal code. This section follows this splitting and introduces two example correlation structures, each emphasizing one of the two codes: the gamma-band synchronization for the temporal code and the spike-count correlation for the rate code.

When introducing the idea of temporal coding, I described a scenario in which two downstream neurons fire synchronously. This scenario is more than hypothetical. Simultaneous multi-site recordings in the visual cortex have shown that distant neuronal populations can synchronize their periodic spiking activity at the gamma frequency band (Figure 1.4 B) [67]. Such gamma-band neuronal synchronization is robust, ubiquitous, stimulus-specific, context-dependent, and an intrinsic network property of cortical connectivity. Presumably, because of its theoretical elegancy in implementing a form of temporal code, the discovery of gamma-band neuronal oscillations has spurred an abundant line of experimental and theoretical research [68]. Among these, Communication through Coherence (CTC) stands out as a prominent hypothesis linking rhythmic neuronal synchronization to efficient and selective neuronal communication [69].

The essence of CTC states that effective neuronal communication is achieved through rhythmic neuronal synchronization. In the case of visual cortical gamma, gamma-band neuronal oscillations represent fast sequences of high and low excitability network states. Inputs that are phase-locked to the high excitability phase are more effective in driving the post-synaptic neuron. Conversely, input that is either incoherent or phase-locked to the non-preferred phase is less effective. A scenario of particular interest arises when the post-synaptic neuron receives two rhythmic inputs, each from a separate neuronal group. In this

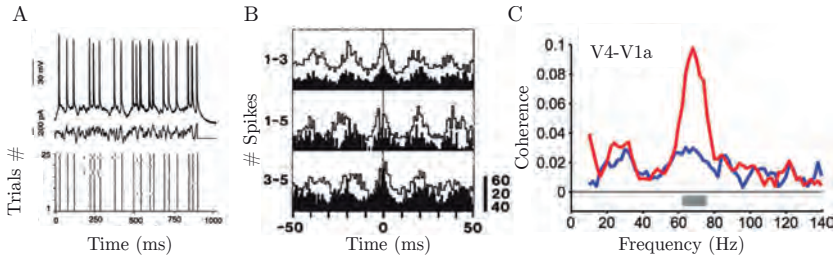


Figure 1.4: The temporal code. (A) With a current-clamp setup, Mainen and Sejnowski [64] showed that the spike timing is reproducible with sub-millisecond resolution with fluctuating current input. A precise spike generation mechanism allows the brain to use a temporal code. (B) Gray et al. [67] showed that distant neuronal pairs can synchronize their oscillatory response in the gamma band (≈40 Hz). Each row plots the spiking responses' cross-correlogram for a neuronal pair separated by ≈400 µm. (C) Bosman et al. [70] showed that the inter-area gamma-band neuronal synchronization is subject to attentional modulation. The red trace shows the coherence spectrum between a V4 site and a V1 site (V1a) when the animal's attention is cued to a location corresponding to V1a's receptive field. The blue trace shows the coherence spectrum when the attention is directed away from the V1a's receptive field. Panel A is reprinted from Mainen and Sejnowski [64] with permission from AAAS. Panel B is reprinted from Gray et al. [67] with permission from Springer Nature. Panel C is reprinted from Bosman et al. [70] with permission from Elsevier.

scenario, entrainment of the post-synaptic neuron to one of the two input rhythms would automatically render the other input less effective. This feature makes CTC an ideal mechanism for selective information routing. Indeed, simultaneous recordings from V1 and V4 in awake macaque monkeys have shown that gamma-band coherence between V1 and V4 sites is enhanced when the stimulus represented by the V1 site is selectively attended to (Figure 1.4 C) [70].

Gamma-band neuronal synchronization represents a type of correlation structure where the coordination of precise spike timing is crucial. Similarly, multi-site recordings have shown that the brain can also coordinate spike rates across different neuronal groups [71]. One form of correlation structure in spike rates is spike count correlation, which measures the co-variability of spike counts between two recording sites across experimental trials. This is also known as noise correlation, distinguishing it from signal correlation, which measures the co-variability of mean responses to different stimuli.

Noise correlation affects information encoding. Theoretical analysis, assuming stimulus-independent noise correlation, shows that high noise correlation between similarly tuned neurons can lead to saturation of encoded information. In a heterogeneous neuronal group, this effect can be summarized by examining the relationship between pair-wise signal correlation and noise correlation. A strong positive correlation between signal and noise correlation results in a greater limiting effect on encoded information. This suggests that the brain may adaptively modulate the correlation structure within a neuronal population according to task demands. Indeed, multi-site recordings from the V4 area of awake monkeys performing a selective visual attention task show that attention can reduce noise correlation among neurons representing the attended stimulus [72–74]. Additionally, noise correlation can be beneficial for encoding extra information when it is stimulus dependent. In such cases, interactions between neurons are synergistic. These synergistic interactions have been observed across layers of a cortical column [75].

Although I contrasted the coordination of spike rate and spike timing, there is an interesting link between the two concepts: The coordination of spike timing can be thought of as the coordination of firing probability in time, which is the limiting case of firing rate. In this sense, saying that two neurons fire synchronously at the gamma-frequency band is equivalent to saying that the two neurons have a strong spike count correlation over time instead of over trials. Will this gamma-rhythmic spike count correlation over time limit information coding similar to the correlation over trials, as we just discussed? The answer would be yes if the "signal" is homogenously distributed across a gamma cycle and the rhythmic fluctuation represents the shared "noise". What follows in this hypothetical scenario would be a reduced feature selectivity and elevated noise correlation (across trials) in gamma phase bins corresponding to the high firing rate. This is because the extra spikes in this scenario represent stimulus-independent noise. However, empirical results support the opposite view. Womelsdorf et al. showed that spikes that occurred in the neurons' gamma phase of the highest spiking probability are most orientation selective [76]. They also showed that the spike count noise correlation close to this mean gamma phase is minimal compared to other gamma phases. Gamma-rhythmic spike count correlation over time thus does not constrain information encoding, as we discussed for noise correlation over trials. More likely, it functions to facilitate information transmission. As the discussion of gamma-band neuronal synchrony

explains, this facilitation can be achieved via post-synaptic coincidence detection or/and phase locking of the spiking output to the high-excitability phase of post-synaptic rhythm (through CTC).

Studies on neuronal interactions emphasize the need to use awake-behaving animals rather than deeply anesthetized or passively fixating ones. In anesthetized animals, neurons in the primary visual cortex show substantial correlated variability, which diminishes significantly when the animal is awake [77]. Notably, as mentioned earlier, neuronal interactions during wakefulness, quantified by noise correlation and gamma-band synchronization, are adaptively modulated by moment-to-moment task demands. Furthermore, even without strong bottom-up stimulation, spontaneous activities in the visual cortex of an awake animal exhibit characteristic dynamics compared to those in anesthetized animals [78]. These dynamics are marked by an increase in high-frequency activity and a decrease in low-frequency slow fluctuations in both single-cell membrane potentials and local field potentials. These observations indicate that the brain operates in drastically different modes when an animal is awake versus anesthetized. Additionally, the brain displays diverse cortical states during wakefulness depending on the behavioral context, which can affect how stimuli are processed.

All these imply that to understand how stimuli are perceived in natural conditions, we need to control the cortical state of the animal in the lab when the brain processes the stimulus. This control is most commonly achieved by engaging the animal in a well-controlled behavior task. In the next section, I will focus on a particular type of behavior task, the selective spatial attention task, to illustrate how the details of a task can affect the cortical states it aims to control.

1.1.3 Behavior: Put Neural Processes into Context

As discussed in the section on physiology, early research on neural mechanisms of vision predominantly used anesthetized animals, viewing the brain as a complex but essentially passive machine processing sensory input from the retina. This "outside-in" tradition, focused on controlling the bottom-up input and analyzing the stimulus-driven response, penetrated the experimental design with recordings later in awake animals. A significant amount of electrophysiological data from awake monkeys is recorded with the animal passively looking at the screen center and flashed with behaviorally irrelevant stimuli in the visual periphery

to drive the recorded neuron as if the animal was virtually paralyzed. Such a passive approach contrasts with the generally accepted view that vision is constructive at its core, driven as much by the cognitive, top-down factors as by the bottom-up stimulation on the retina. To understand how vision works, we need to know how the top-down factors, from the "inside-out," interact with the stimulus from the "outside-in" [79]. The first step is to be able to control these top-down factors experimentally and the key to such control is behavior.

Visual attention is one of the most important and most studied top-down factors [80]. I will use selective visual attention tasks to illustrate the rationale of using behavioral tasks to create the necessary brain state contrast for studying this specific top-down factor in non-human primates. It unfolds as follows: (1) When a *human* participant is instructed to selectively pay attention to location A and ignore location B, their performance at location A would be better than at location B. (2) Assuming the condition "S attends to X if and only if S selects X to guide behavior", the link between attentional orienting and the improved performance can then be used to define attention operationally. Specifically, with other factors controlled, suppose performance is better at location A than at location B, then, we can say that the subject attended to location A because the stimulus at that location guided behavior. (3) This operational definition can now be applied to non-human primates and other animals that cannot communicate verbally with the experimenter. Namely, if an animal behaved according to the instruction and achieved a better performance at location A than location B, we say that its attention was successfully deployed at location A [81].

Importantly, this definition of attention is task dependent. Neural correlates of attention discovered in any task are thus inherently linked to that specific task. There is no guarantee that the exact same attention mechanisms are recruited in, for example, a detection task requiring the subject to saccade to the target compared to a match-to-sample task requiring the subject to release a bar, albeit both tasks require "attention" to guide the prospective action. Extrapolate it further; it is possible that the attention *everyone knows* in real life is very different from that *everyone studies* in the laboratory. This potential discrepancy between laboratory and real-life situations is one of the essential driving forces for this thesis to advocate for an ecological approach. The subtleties of how a behavioral task would affect the recruited neural processes would be more

specific after the classic Posner tasks are presented and subsequently analyzed below.

The original application of the Posner task in human participants was presented in 1980 by Michael I. Posner [82]. Its basic structure in an simplified monkey task is illustrated in Figure 1.5 A. It starts with the monkey looking at a central fixation dot. Two pedestal stimuli are then shown at separate locations. After a random delay, a cue is presented to indicate that one of them is the to-be-attended stimulus. Following another random delay, a probe stimulus is shown on one of the two pedestals. The animal needs to respond to this probe's appearance no matter on which pedestal it appears. In 80% of all trials, the cue is valid, and the probe stimulus would appear on the cued pedestal. In the remaining 20% invalid-cue trials, the probe would appear on the non-cued pedestal. Because attention is known to facilitate the detection of a weak signal and to shorten the response time, using stimuli that are hard to detect or rewarding a faster response encourages the animal to use the cue and attend to the cued location, which more likely contains the probe. Crucially, whether the animal used the cue and oriented attention accordingly is evaluated by comparing its cumulative performance when the cue is valid versus when the cue is invalid. In practice, validating that the animal learned to use the cue is done during the training phase before recording the electrophysiology data. Once the animal has learned the task, and its attention can be shown to be reliably manipulated by the cue, researchers can then use this animal and this specific task it learned to investigate the neural correlates of attention. During each recording session, the two pedestal stimuli are placed according to the neurons' receptive fields; for example, one stimulus can be placed within the RF, while another can be placed outside the RF. Varying the cued location constructs two conditions. The attend-in condition contains trials in which attention is cued to the stimulus within the receptive field, whereas the attend-out condition contains trials in which attention is cued to the stimulus outside the receptive field. While analyzing the data, trials belonging to these two conditions can be aggregated, and the corresponding neuronal response in these two conditions can be compared to deduce the neural correlates of attention.

The success of Posner-style tasks, particularly their application in human psychophysics, is evident in the original paper's over 13,000 citations as of writing this thesis in 2024 [82]. The key to its success lies in that a physically identical

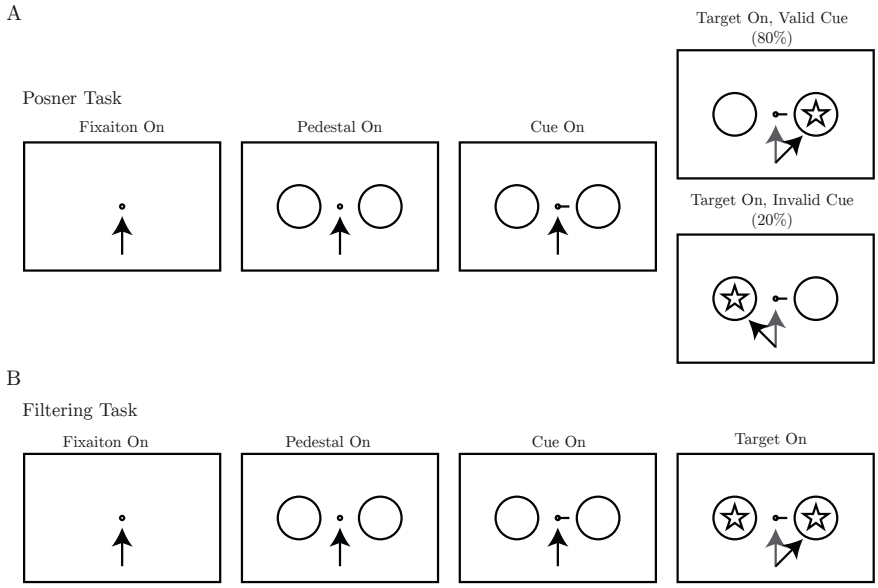


Figure 1.5: Two types of attention tasks. (A) The basic structure of a Posner task. In a Posner task, a trial starts with the animal fixating on the fixation dot. After a random delay, the pedestal stimuli are presented. After another random delay, the cue is presented. The cue illustrated is a small bar pointing to the cued location, which will contain the target probe in 80% of the trials. In the remaining 20% of the trials, the target probe is presented at the un-cued location. Importantly, independent of where the target appears, the subject needs to make a response to the target, e.g., saccade to it. (B) The basic structure of a Filtering task. In a Filtering task, the cue is always valid. Note that two probes appeared simultaneously during the Target On frame. The subjects need to ignore (filter out) the distractor probe and make a response to the target probe at the cued location.

stimulus can, in one trial, be the to-be-attended stimulus and, in another trial, become the to-be-ignored stimulus. Thus, any differential behavioral performance/neuronal activity observed in these two conditions is not due to differences in bottom-up stimulus drive but to the constructed attentional contrast. However, there are subtleties in the task design.

Foremost, there could be different strategies to solve the task equally well. These different strategies likely involve the activation of different neural processes. For example, in a detection task, according to the signal detection theory, in order to enhance the detection performance at the cued site, the brain can either increase the sensitivity for discriminating the probe signal from noise or lower

the criterion for asserting the response comes from a true probe signal instead of noise [83]. Both strategies can be adopted to enhance the detection rate. Indeed, when the task is designed to preferentially encourage the animal to adopt one of the two strategies, the brain gives a solution that implements that specific strategy adopted. Specifically, Luo and Maunsell [51] showed that the attention-related firing rate change in macaque V4 corresponds to the sensitivity shift but not the criterion shift. Separate studies showed that the criterion shifts may find its correlates in subcortical structures like the superior colliculus [84]. Another possibility is that the criterion shift is not coded in firing rate but in spike timing. Synchronized spikes can achieve higher gain through mechanisms like CTC and consequently implement the decision bias by effectively lowering the afferent firing-rate threshold for driving the downstream decision units [85].

Secondly, attentional allocation is not static. Because the probe stimuli appear at random times, at any moment, given that the probe has not appeared yet, there is a moment-to-moment varying likelihood of probe appearance at the next moment. This likelihood is known as the hazard rate [86]. Unless the distribution of probe onset time is exponential, which is never the case in a real task, the hazard rate is not flat. An uneven hazard rate implies that allocating attentional resources would be optimal if more is allocated to the moment associated with a higher hazard rate and less is allocated to the moment associated with a lower hazard rate. In line with this reasoning, the strength of attention-related firing rate increment has been found to correlate with hazard rate modulation within a trail [87]. Furthermore, the hazard rate may affect preferentially the processes related to the criterion shift as compared to processes related to the sensitivity shifts. With a very high hazard rate, for example, at the last moment of possible probe appearance, there is essentially no need to base the decision on the sensory signal (assuming that there are no catch trials). In support of this, Schoffelen, Oostenveld, and Fries [85] showed that, in a task where the hazard rate is modulated, gamma-band neuronal coherence between neuronal groups is modulated accordingly. Importantly, this modulation is found to correlate with a change in reaction time, a characteristic behavior maker of attentional orienting.

The observation that covert spatial attention is a dynamic, heterogeneous process informs future research directions. First, it requires a finer behavior control to isolate the individual sub-components, as Luo and Maunsell [51] did.

Second, we need to understand how these different sub-components interact, presumably at the system level, with multi-site recordings simultaneously from multiple relevant visual areas. Third, it is essential to investigate whether these sub-components and their interactions manifest themselves in tasks more akin to real-life scenarios.

The first two points emphasize the neural processes' interaction within a behavioral context. The third point, whether the neural processes identified in a Posner-style task manifest themselves in real life, is related to the task's ecological validity. The task's ecological validity can be reflected in how difficult it is to train a naïve animal on such a task. The importance of the task being ecologically valid and consequently requiring minimum training is well captured by György Buzsáki in his recent book, *The Brain from Inside Out* [79], where he wrote "...if training takes weeks or longer, the experimenter should seriously consider that the brain signals associated with such extended shaping of behavior may not reveal much about the intended question because each animal may use different tricks to solve the task, and these may remain hidden for the experimenter." So, how long does it take to train a macaque monkey on a Posner-style task? The answer is that it can easily take months! Such extremely long training time points us to yet another subtlety of the Poser task as it is used in non-human primates: the reward protocol.

In a Posner task, the reward is associated with successfully detecting the target probe, independent of where the probe appears. Notably, if the animal successfully detects the probe appeared at the un-cued location, in which case the cue was invalid, it will still be rewarded. During the learning phase, obtaining a reward in such an invalid-cue condition likely means the animal was attending to the un-cued location. However, attending the un-cued location is the opposite of what the experimenter wants the animal to learn. The to-be-learned behavior, correctly orienting attention according to the cue, does not benefit the animal in trials where the cue is invalid. It is only beneficial for the average performance. In other words, in a Posner task, there can be a mismatch between the trial-by-trial reward protocol and the to-be-learned behavior that benefits only in aggregated performance. This mismatch presumably explains why the animal learns a Posner-style task so slowly.

This mismatch would be greatly alleviated by the different operationalization of

attention in the filtering tasks, as illustrated in Figure 1.5 B. In a filtering task, the cue is always valid. The animal needs to ignore (filter out) an additional distractor probe presented at the un-cued location. Attention in the filtering task is operationally defined by the animal selectively responding to the target probe presented at the cued location and not responding to the distractor probe presented at the un-cued location. Unlike in the Posner-style task, the cue is always valid in a filtering task. Always valid cuing means correctly orienting attention according to the cue, and thus, responding selectively to the target probe would be rewarded at each correct trial. In other words, in a filtering task, the trial-by-trial reward protocol matches the to-be-learned behavior in such a way that the to-be-learned behavior benefits the animal at each trial. The advantage of such a match is significant. With a filtering task, training is much easier and thus more likely to reflect how covert attention is deployed in real life. For example, studies that employed the filtering task can often cue the animal to a random position that varies on a trial-by-trial basis. For comparison, studies that employed the Posner task most likely need to resort to block design, where the cued location is fixed for tens of consecutive trials. Attention maintained at a fixed location for repeated trials extending over a few minutes differs apparently from the attention we know in mundane life, which often changes its targets every few hundred milliseconds, if not faster [88].

The details of a reward protocol affect not only the animal training, as detailed above, but also the interpretation of the relevant empirical results. Assuming the intention of the test animal is to maximize the reward it can obtain in a given task, the reward protocol would consequently set the behavioral bias in a given task to facilitate reward harvesting. This facilitation is necessarily mediated by top-down factors modulating the processing of the relevant stimulus. Specifically, in the tasks illustrated in Figure 1.5, in order to get more reward, the animal would need to set its top-down factors to enhance the processing and, hence, detection of the change signal (associated with the probe appearance). Notice that even though this change signal is the signal that is directly relevant for behavior, it is often not the focus of data analysis. The focus of analysis is often the signal during the delay period prior to the change. During the delay period, the electrophysiology data is clean and not contaminated by the stimulus onset transient or movement noise. However, the neural activity in the delay period is driven by the pedestal stimulus, which is not directly related to reward. To be more specific, this pedestal-driven activity may be

related to preparing the brain for the change signal processing and thus being biased towards the to-be-attended location. However, this bias is not to be confused with attention because this biased neural activity in the delay period is not used directly to guide the correct, intended behavior, be it making a saccade to the target or pressing a level. Indeed, if this biased activity resulted in a response, this response would be treated as an unintended error, i.e., a false alarm. Consequently, data obtained during the delay period should be interpreted as reflecting the unintended biased processing incidentally caused by attending to the change signal to be detected at the cued location. One would need some additional assumptions to claim that this unintended bias is behaviorally relevant and reflects the same mechanisms underlying the selective behavior to the change signal. For instance, for the often-reported firing rate enhancement in the delay period to constitute attention, it is necessary to make an additional assumption that the specific mechanisms that result in the observed firing rate enhancement during the delay period would result in similar firing rate enhancement for the target signal to be detected. The above analysis is not just a linguistic play; it has real consequences. For example, in the case depicted in Figure 1-5, it is easy to conceive an attention mechanism that suppresses a neuron's response to the pedestal stimulus, say from firing 20 spikes per second to 10 spikes per second, such that it could better encode the change signal, say from firing 10 spikes per second to 30 spikes per second, a three-fold change due the suppression of the sustained response to the pedestal.

The assumption that neuronal modulation in the delay period can be transferred to neuronal modulation to the change signal is more likely to hold for tasks where the stimulus during the delay period is of the same kind as the change stimulus to be detected. Visual stimulation during the delay period is of the same kind as the to-be-detected stimulus in the modified match-to-sample task, where all test stimuli in a sequence, match or non-match, are transient flashes [72, 89]. So, it seems we are concluding that it is better to use flashes than to use pedestals. But wait! We do not see flashes everywhere in our everyday life! What we have reached is a clash between ecological and methodological validity. Another big clash in this regard is the use of peri-foveal stimuli. Methodologically, this is important for having an equi-eccentric control stimulus. But ecologically, it makes little sense, as primates typically foveate the attended stimulus.

Such a clash is revealing because it guides further research starting from either

side. Starting from the methodological side, e.g., using tasks with unnatural flashes, researchers would want to see if the same neural processes happen in more ecological conditions. Starting from the ecological side, e.g., using tasks with sustained pedestals, researchers would want to check how these processes affect the ensuing change signal processing, which constitutes the guidance for the action that leads to reward [90, 91]. Importantly, either case demonstrates the importance of careful analysis of the behavioral context established by a given task and subsequently putting the neural processes into this context.

Finally, I will touch on the elephant in the room: even though we emphasized the importance of behavioral context, in all behavioral tasks mentioned above, the animal's bodily movements are severely constrained to the extent that the subject's eye movements are limited to a tiny fixation window, often as small as one visual degree in radius. This type of control, being the defining feature of the covert attention task, is deemed helpful, because in the ideal scenario, fixed eye position ensures that the participant's retinal input is under the experimenter's control. However, this ideal scenario is probably neither true nor reflects how vision works naturally. It is not true because the eyes are never at rest. Even during intended fixation, there are high-frequency tremors, diffusion-like drifting, and micro-saccades. In particular, micro-saccades have been shown to be tightly linked to the attentional firing rate enhancement in V4 to the extent that when there are no micro-saccades, there is no attention [92]. Prolonged fixation also does not reflect how vision works naturally. In natural vision, we make saccades a few times per second without even noticing them. These saccadic eye movements fundamentally shape the input to our visual system [93]. Our visual system, with its intrinsic connectivity and dynamics, has evolved to deal with such pulsatile and directive input that bring objects of interest from the periphery to the fovea. When the input lacks this pulsatile and directive structure, the brain seeks it. Our brain's urge for such pulsatile and directive input is impressively demonstrated in a case study with patient AI [94]. Patient AI cannot make eye movements due to continental ophthalmoplegia. To compensate for the lack of saccadic eye movements, patient AI developed saccadic head movement that is in many ways similar to the regular saccadic eye movements, albeit the apparent uneasiness to do so. These observations, that it is hard to get rid of eye movements in the first place and that when the eye movements are indeed absent, the brain compensates them with similar head movements, strongly suggest that regular eye movements are integral parts of

natural vision. What follows is that to understand how vision works naturally, we need to allow regular eye movement in the laboratory. We need to study vision in the context of regular, naturalistic eye movements.

Allowing regular eye movements poses many challenges to the experimenter. Foremost, the majority of visual neurons, particularly those neurons in the early- and mid-visual areas, are sensitive to the stimuli's retinotopic position. The stimuli's retinotopic position necessarily changes with each eye movement. Ever changing retinotopic input makes it very hard to control the basic stimulus drive when the eyes are allowed to move freely. Failing to control the stimulus drive would severely constrain the detailed physiological investigation of neural processes underlying any visual perceptual phenomenon. However, it is not impossible to control the stimulus drive while allowing naturalistic eye movements. A promising direction is the use of carefully designed visual search tasks.

In a typical overt visual search task, a target is displayed to the participant among an array of distractors, and the animal's task is to locate this unique target [95]. Because stimuli constituting the search array are presented on an empty background, these stimuli afford "looking at", and the background does not. What follows is that the animal's gaze position is, to some extent, predictable. Having this prediction, to control stimulus drive for a given neuron, the experimenter just needs to place and space the search array such that when one of the stimuli is foveated, a planned stimulus will fall into the receptive field of that neuron. Also note that in a visual search task, the animal's intention can be controlled and thus studied by defining the search target. For example, in a trial where the search target is a red square, all red or square distractors would be selected by feature attention, for they share some features with the search target. The result is that a properly designed visual search task can achieve control of both the bottom-up stimulus drive and top-down intention, thus allowing the detailed physiological investigation in an ecologically valid behavior context. Chapter 3 of this thesis develops a visual search task in this direction.

A more general discussion of the role of a behavior task in visual neurosciences is covered in a recent review paper by Kay et al. [96].

1.2 The Action-Perception Loop

So far, I have argued that, to study natural vision, it is crucial to place the relevant neural processes within a natural behavioral context. In this section, I suggest that this behavioral context is characterized by the action-perception loop (the AP-loop). Engaging the animal in naturalistic behavior means engaging it in these AP-loops. In brief, the suggested AP-loop describes the agent's visual life being characterized by sequences of rhythmic sensorimotor decisions. Importantly, connecting each sensorimotor decision is the perception. On the one hand, perception in the AP-loop is the processed sensation that guides future action. On the other hand, action in the AP-loop brings about new information for perception. The suggested AP-loop differs from the sensorimotor loop in the enaction theories, where the perception is hypothesized to be "enacted" through the continuous sensorimotor loop. The perception in the proposed AP-loop is more of a Gibsonian flavor, being a process of active information pickup in each action-perception cycle as the agent actively explores the information-rich environment [3, 97].

1.2.1 Action Is Guided by Perception

The AP-loop is considered a fundamental concept for ecological visual neuroscience because it provides the behavioral context that organizes various neural processes for direct information pickup. To acknowledge the importance of the AP-loop, we first need to understand seeing as an active process biased by our intentions. This bias can be summed up by the slogan, "What we see is what we need." Most prominently, we see what we need by bringing what we want to see into our high-resolution fovea through rapid eye movements called saccades. Without noticing, we make several saccades every second. A sequence of saccadic eye movements and the intermittent fixations form the scanpath. The way intention biases our seeing is directly reflected in how it affects the scanpath. This is impressively illustrated in Yarbus's seminal work on eye movements in the 1960s [98].

In the last chapter of his book *Eye Movements and Vision*, Yarbus described an experiment in which he recorded a participant's eye movements while the participant viewed a picture of the Russian realist painting "An Unexpected Visitor" by I.E. Repin (Figure 1.6 A). In this painting, Repin depicted the family's response to the son's unexpected return from exile in Siberia. The

Figure 1.6: Eye movements are biased by intention. (A) The picture, *An Unexpected Visitor* by I.E. Repin. Yarbus recorded a participant's eye movements while the participant viewed this picture using different instructions. (B) The participant's eye movement under the instruction "to estimate the family's material circumstances". (C) Same as (B) but under the instruction "to estimate the ages of the people". The original results published in Yarbus [98] have the paint separately presented with the scan path. Archibald [99] recreated these images in high resolution and aligned the picture and the corresponding eye movements. Panel A-C are reprinted from these recreated images with permission from Immaterial Inc.

participant was asked to view the picture seven times. Each time before viewing the picture, the participant received a specific instruction that set the intention of the following picture viewing. What Yarbus found is that the participant's scanpath varied significantly with the received instruction. The participant's eyes dwelled more on the clothing and furniture when the task was to estimate the family's material circumstances (Figure 1.6 B). By contrast, the eyes stayed almost exclusively on the characters' faces when the task was to estimate the ages of the people (Figure 1.6 C). These task-dependent scanpaths showcase that the participant's seeing represented by the eye movements was biased by their intention to selectively sample information from the environment. Such intention-biased information sampling through sequences of sensory-motor decisions is the essence of the AP-loop. I elaborate on the AP-loop and intention-biased information sampling below.

First, intention biases action via the guidance of perception. This is best elucidated in the Perceptual Control Theory (PCT) [100]. PCT states that behavior is the control of perception. In short, within the framework of PCT, the behaving agent's intention sets the internal reference perceptual signal. This reference signal is then compared with the current perception (calculated based on the current and previous sensory input). The difference (or error) between

the current and reference perception signal is then used to guide behavior. The output behavior, bodily or not, will affect the future input to the agent's nervous system. Altered input will then affect the future perception. Feeding back this future perception to the perception comparator closes the loop. Many of these closed feedback loops form a hierarchical control system and eventually select a sequence of actions that achieve the agent's intention (via matching the reference perceptual signal). In the case of Yarbus's eye movement study mentioned above, part of a hypothetical control system can be described as follows: the task instruction, e.g., "to give the ages of the people", would, at the higher level of the control hierarchy, set the reference signal to be "feeling confident about the lady's age". This higher control loop will set the reference signal in the lower level of the control system to be "seeing clearly the lady's face". Given the current perception of seeing the lady's face not so clearly because the participant's current fixation is away from the lady's face, this error signal will cause an action, i.e., a saccadic eye movement that directs the gaze to the lady's face. This action will cause a new perception, seeing a clearer foveated face, satisfying the lower-level control loop. The lower-level control loop will also notify the upper-level control loop to update its own state, which would complete one cycle of the AP-loop. Suppose the final goal is not achieved after this cycle, i.e., being confident about the people's ages in the picture. In that case, the above-described AP-loop will continue and result in a scanpath with most of the fixations on the characters' faces.

PCT is not the only theoretical framework that puts perception in the central position to explain behavior. In Gibson's ecological approach, perception is said to be hungry for clarity, and this intrinsic hunger for clarity, a kind of reference perceptual state in PCT, drives the active information pickup. More recently, a similar account of action and perception can be found in the active inference framework [101, 102]. The active inference framework has gained popularity for offering a unified mathematical framework to understand (but not limited to) perception, action, and learning. In active inference, a sequence of action choices (a policy) results from inferencing the best policy over all possible policies. This inferencing process involves the evaluation of each potential action's consequence on the agent's future observation, i.e., to calculate the Variational Expected Free energy (VEF) over the potential actions. Crucially, calculating VEF requires defining a "prior preference distribution." This prior preference distribution describes the agent's preference over all possible observations. Essentially, this

prior preference can be thought of as the reference perception signal in PCT. As action and inference continue, similar AP-loops as described in PCT would be expected in active inference.

1.2.2 The Rhythmic AP-Loop in Rodents

Second, by using the words "sampling" and "loop", I want to implicitly convey the idea that the information sampling and the AP-loop implementing the sampling have characteristic timing. In other words, the sampling is rhythmic. In the context of the AP-loop, rhythmic sampling entails rhythmicity in both action and perception. To be more specific, rhythmic sampling entails that both 1) action, observable as the motor output and the associated pre-motor neural processes, and 2) perception, observable as either the underlying neural process or as the behavioral readout of perception, have a characteristic timing and show various degrees of rhythmicity. To illustrate these points, because 1) the AP-loop is a general concept not limited to vision and 2) the evidence for rhythmicity in rodents' whisking is unequivocal, I first illustrate the rhythmicity of action, perception, and their coupling in rodents whisking, and then from there, draw possible links to primate vision.

The motor output of rodents' rhythmic tactile information sampling, i.e., whisking, can be observed via a high-speed camera tracking the vibrissae movement in the awake animal. Two features of the observed whisking rhythm are worth noting. First, the whisking rhythm is under active control. The active control of whisking is demonstrated clearly in a conditioning study where the animal's whisking was selectively conditioned on the tone stimulus' presence or absence [103]. The control of whisking is also reflected in the rats' natural behavior. For example, depending on the behavioral context, a rat can spontaneously engage itself and readily switch between two distinct whisking modes: a high frequency, low amplitude foveal whisking mode around 15∼25 Hz when palpating an object ahead; and a low frequency, high amplitude explorative whisking mode around 7∼12 Hz when whisking in air without contact [104]. Furthermore, by using a tactile virtual reality setup, Sofroniew et al. [105] showed that the mouse's whisking is coupled with running: the faster the mouse rusn, the higher the frequency with which it whisks. These observations show that rodents can control their whisking rhythm to provide their desired sensory information. Second, the whisking rhythm is coordinated with other behavioral rhythms, forming a hierarchy of orofacial rhythms [106]. Most prominently, the rat's

whisking cycles are one-to-one locked to the sniffing cycle (≈8 Hz). Whisking is also coupled to the lower frequency basal breathing cycle (<3 Hz), which resets the faster whisking rhythm with each inhalation. In some sense, the coordination of rodents' orofacial rhythms is not surprising as the motor plant (particularly the extrinsic muscles for the mystacial pad motion) is shared in the whisking, sniffing, and breathing. As we will see shortly, the sharing of motor plants can be traced backward to the sharing/interacting of the corresponding upstream (pre-)motor neural processes in the brainstem.

The premotor process obligatory for the rodent's rhythmic whisking is an all-inhibitory premotor network in the brainstem [106, 107]. Specifically, neurons in the vibrissa intermediate reticular nucleus (vIRt) – a medulla nucleus located medial to the nucleus ambiguous – provide rhythmic inhibitory input to the vibrissa facial motor nucleus (vFMN). This rhythmic inhibition, which causes vibrissa retraction, together with the tonic excitation in the vFMN, which causes vibrissa protraction, constitute the whisking cycle. Central to this whisking cycle is the rhythmic activity in the vIRt. Furthermore, rhythmic activity in the vIRT is thought to be reset by the rhythmic input from the pre-Bötzinger complex (preBötC), which drives inhalation. Such resetting of the whisking oscillator by the breathing generator is likely the underlying neural mechanism for the observed coordination between the breathing/sniffing rhythm and the whisking rhythm.

In the context of the AP-loop, the rhythmicity in action, as described above, implies rhythmicity in the corresponding perception, i.e., somatosensation and olfaction, through shaping the temporal structure of the sensory input. Specifically, taking the example of whisking, we show below that 1) the whisking rhythm modulates the input tactile signal, and 2) the animal's perception is correspondingly rhythmically modulated.

The modulation role of whisking, *per se*, on tactile input is visible at the first neuron in the vibrissa afferent pathway, i.e., in the activity of trigeminal ganglion cells [108]. Such modulation can be seen as a kind of re-afferent signal as it is caused by the animal's own movement. Re-afferent signals may be essential for the animal to extract the genuine ex-afferent signal from the environment. However, presumably incidentally, because the re-afferent and ex-afferent signals are multiplexex on the same neuronal population, the rhythmically occurring

re-afferent signal likely means that the processing of the ex-afferent signal is also rhythmically modulated. For the rodent's somatosensation, this means an arriving tactile stimulation's effectiveness may depend on the whisking cycle's phase. A recent study by Isett and Feldman demonstrated such (whisking) phase-dependent sensory encoding [109]. In this study, the authors trained the mice on a rough-smooth surface discrimination task and recorded the neural activity in S1. They found that the surface whisking phase tuning of S1 neurons shapes the sensory encoding so that the encoding of surface roughness information is best at the protraction phase, which is also the preferred whisker phase for most recorded S1 units. It is worth noting that the (whisking) phase-dependent sensory encoding itself is context-dependent, as the phase information appears to affect sensory processing differently when the animal was whisking in the air compared to when it was whisking on the surface [110].

1.2.3 The Rhythmic AP-Loop in Primates

Rodents sample tactile information through whisking, while primates primarily sample visual information by making saccades. Previously, I examined the spatial aspects of a saccade sequence to illustrate how intention influences looking and seeing. Now, I focus on the temporal aspects of a saccade sequence to illustrate that visual information sampling, like tactile sampling in rodents, is rhythmic. Specifically, I examine the inter-saccadic interval (ISI) distribution, as it provides all necessary information if saccades can be modeled as a renewal point process.

The shape of the inter-saccade interval (ISI) distribution determines saccade rhythmicity. It is often stated that we make 2∼4 saccades per second, which specifies a rate parameter. However, a rate parameter alone does not imply rhythmicity. For example, a homogeneous Poisson point process can be described by a rate parameter (e.g., 4 events per second), but it is not rhythmic. Conversely, a perfectly rhythmic process, such as one event every 250 ms, can also be described by the same rate parameter (4 events per second). The degree of rhythmicity in a renewal point process is best captured by the shape of its inter-event-interval distribution. A sharp unimodal distribution indicates high rhythmicity, while a smooth, non-peaky distribution indicates low or no rhythmicity. Based on this, we can define a rhythmicity index (RI) for saccadic eye-movements as the mean of the ISI distribution divided by its standard deviation. Note that a homogeneous Poisson process has an exponential inter-

event-interval distribution where the mean equals the standard deviation (RI = 1), regardless of the rate. This supports the idea that the RI effectively captures the rhythmicity of a renewal process independent of its rate parameter. Mathematically, the RI is the inverse of the coefficient of variation (CV), a standardized measure of dispersion.

Now that we have a rhythmicity index (RI), we can investigate whether saccades are rhythmic and whether primates can control their saccade rhythm similarly to how rodents control their whisking rhythm. In a recent study, we calculated RI from saccade sequences in humans and macaque monkeys using a visual foraging task described in Chapter 3 [111]. The RI in both species was significantly higher than one, indicating some rhythmicity in saccadic visual sampling. Importantly, we found that task parameters can modulate saccade rhythmicity. A simple search array combined with a high saccade rate ($\approx$4 Hz) produced the strongest rhythmicity and vice versa. We extended this observation to a more naturalistic setup where subjects freely viewed a 10-minute first-person-view video. This setup allowed us to calculate RI over a very long saccade sequence. Like bouts of rhythmic whisking in rodents, there are bouts of rhythmic saccade sequences during the extended free viewing.

The oscillator circuit responsible for rhythmic whisking in rodents is in the brainstem. The brainstem also contains the premotor circuit for saccade initiation that may also control the saccade rhythmicity [112, 113]. Central in this circuit are the pontine raphe's omnipause neurons (OPNs). During fixation, OPNs receive tonic excitation from fixation neurons in various cortical and subcortical brain areas. Active OPNs prevent saccade execution by inhibiting the bursting neurons that innervate the oculomotor neurons. OPNs also receive inhibitory input from long-lead bursting neurons (LLBNs). Moments before the saccade execution, LLBNs are activated. LLBN activation will release the inhibition of oculomotor neurons from OPNs and cause the saccade execution.

As described above, the OPNs in the saccade initiation circuit are intriguingly similar to vFNMs in the whisking initiation circuit (Figure 1.7). Both of them receive tonic excitation, and both of them receive phasic inhibition that initiates the respective motor outputs. Because rhythmic inhibition of vFMNs by vIRt neurons forms the whisking oscillator, the network of reciprocally inhibited LLBNs and OPNs may similarly underlie the observed saccade rhythmicity. An

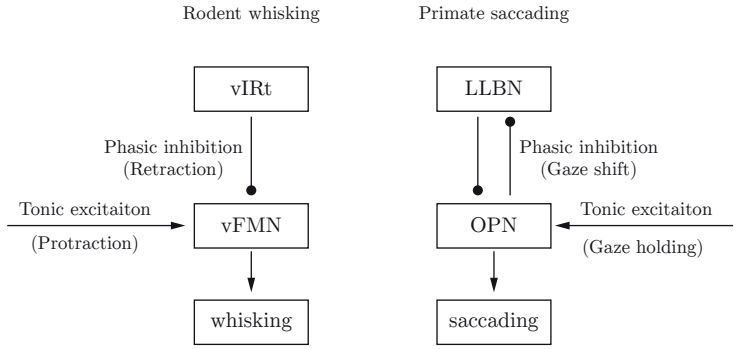


Figure 1.7: Rodent whisking and primate saccading have similar premotor circuits in the brainstem. The simplified brainstem circuit for rodent whisking is shown on the left. The core of this circuit is the vibrissa facial motor nucleus (vFMN). During the retraction phase, the vFMN receives tonic excitation. Rhythmic whisking occurs when this tonic excitation is interrupted by rhythmic inhibitory input from the vibrissa intermediate reticular nucleus (vIRt). The simplified brainstem circuit for primate saccading is shown on the right. The core of this circuit is the pontine raphe's omnipause neurons (OPNs). During fixation, the OPNs receive tonic excitation. Saccades are executed when this tonic excitation is interrupted by phasic inhibition from the long-lead bursting neurons (LLBNs). Both circuits are simplified to emphasize their similarities.

interesting observation supporting this hypothesis is the occurrence of staircase saccades resulting from persistent superior colliculus (SC) microsimulation [114]. These staircase saccades are very rhythmic, with stable ISIs despite the persistent non-rhythmic input from the SC, indicating that rhythmicity can be generated in the recurrent network below the SC. However, the ISIs of staircase saccades are very short, around 100 ms, much shorter than the ordinary ISIs observed in voluntary saccades (200∼300 ms). Presumably, the staircase ISIs represent the lower bound of a saccade period. The normal saccade periodicity of 200∼300 ms likely originates above the SC, in the cortical visuomotor areas projecting to the SC, reflecting a rhythmicity in saccade target selection [111, 115].

The cortical visuomotor processes deemed important for saccade timing are centered around the concept of a hypothetical stochastic accumulator: A saccade is initiated when this accumulator reaches a decision threshold [116]. For example, premotor saccade neurons in the frontal eye fields (FEF) and the SC exhibit accumulator-like firing patterns that predict saccade response times. While these models mainly explain saccadic reaction time for a single response

saccade, the time required for the accumulator to reach the decision threshold (plus the non-decision time) also fits the range needed to explain inter-saccadic intervals (ISIs) in a saccade sequence.

To reconcile the accumulator model for a single response with a continuous action-perception (AP) loop, we can use a simple notation trick: replace the excursion from baseline to the threshold firing rate with phase angles from zero to 2π and introduce a reset after the accumulator reaches the threshold. Using this notation, one can align neural processes in a sequence of sensorimotor decisions and understand how the processes in one action-perception cycle affect the processes in the next cycle. For instance, by aligning time-resolved neural activity to the moment of reset or its behavioral correlates (e.g., saccade onset), we can investigate how pre-saccadic activity in one neuronal population (e.g., peripheral neurons processing saccade targets) affects post-saccadic activity in another population (e.g., foveal neurons encoding the same stimuli after the saccade). Certainly, aligning analysis to behavioral events is not new, but linking processes before and after these events continuously as a rolling operation is relatively new. Correspondingly, this approach would be most effective if the process under investigation is recurring rhythmically, namely, if there is an ongoing AP-loop. Lastly, it is worth noting the connection between the suggested analysis aligned to a self-generated behavioral reset (e.g., saccade onset) and the classical analysis aligned to a stimulus onset. In the classical analysis aligned to the stimulus onset, it is often observed that with each stimulus onset, the visuomotor system is effectively reset, reducing the probability of initiating a saccade in the next few hundred milliseconds. This phenomenon, known as saccadic inhibition, can be compared to the reduced probability of making a new saccade immediately after the previous one.

There is ample literature on the effect of eye movement, particularly saccadic eye movements, on visual perception [117, 118]. The discussion has revolved around the problem of visual stability. That is, the world looks stable perisaccadically, as it should be, even though each regular saccade comes with a significant retinal image slip and blurring. To appreciate this true feat of our visual system, one can try shooting a short video clip with a handheld camera and jerking the holding hand a few times in between. Viewing the resulting video, I bet, will not be satisfying. However, this is the kind of visual input our eyes send to the brain. In fact, what the brain receives is even worse: because the visual

acuity drops quickly from the foveal to the visual periphery, what the brain receives is more like a discrete image sequence obtained through a tiny, blurred aperture connected by smeared motion streaks. Along the lines of this analogy, the efforts to solve the visual stability problem fall into two broad categories. First, on the suppression of the perisaccadic motion streaks. And second, on the construction of a stable percept from the discrete snapshots obtained during consecutive fixations.

The suppression of the perisaccadic retinal input can be readily revealed by one not being able to see their own eye movements [119]. Some have suggested that this suppression is largely a pure retinal phenomenon, caused by the clear retinal image obtained before and after the saccade masking the motion-induced retinal image gray-out during the saccade [120]. However, accumulating evidence has shown that the saccadic suppression is more than passive visual masking and is an active and selective process. For example, by measuring the contrast sensitivity to various horizontal gratings during a horizontal saccade, Burr, Morrone, and Ross [121] found that, in such a situation where the motion blur is minimum, the suppression was not only substantial but also selective for the magnocellular pathway: the suppression was strongest for low spatial frequency luminance contrast gratings and exempt for either gratings of chromatic contrast or high spatial frequency. A careful psychophysical analysis further indicates that the selective saccadic suppression finds its neural correlates very early in the visual processing, even before the start of the motion signal analysis, likely already in the LGN [122]. The early engagement of saccadic suppression has received direct evidence by using TMS stimulation inducing phosphenes at different sites along the hierarchy: a phosphene induced at the occipital cortex was shown immune to saccadic suppression, whereas a phosphene induced at the retina was not [123].

It is worth noting the similarity between this early saccadic suppression in the visual pathway and the early whisking-dependent modulation in the rodents' somatosensorial pathway described earlier in this section (Figure 1.7). This similarity suggests a common priority of integrating action into perception in the evolution.

That the brain actively constructs a stable percept from consecutive snapshots can be revealed by perisaccadically distorted spatiotemporal perception. A

popular hypothesis for perisaccadic spatial perception distortion is that visual space transiently compresses toward the saccade target just before the saccade. Psychophysical results show that this compression is so pronounced that a stimulus flashed up to 10 degrees to the left or right of the saccade target can be mislocalized to the saccade target's position. Notably, because the color and shape information is retained in the mislocalized object, the spatial distortion likely occurs later in visual processing, after basic feature extraction is completed [124–126].

The perisaccadic spatial distortion involves more than just compression toward the saccade target. For example, if a subsequent visual reference is presented, a bar flashed perisaccadically can be mislocalized to a broad range of spatiotemporal locations away from the saccade target [127]. Quantifying the effectiveness of the visual reference in attracting the flashed bar at saccade onset reveals a slanted spatiotemporal interaction field. Importantly, this slant parallels the spurious retinal motion during the saccade, implying that the saccade target's peri-saccadic retinal motion would fall within this interaction field throughout the saccade flight time. Tracking the saccade target's retinal motion as such may explain visual stability.

The slanted interaction field may find its neural correlates in the predictive receptive field (RF) remapping observed in many brain areas, including the SC, FEF, and LIP [128]. Many neurons in these areas shift their receptive fields predictively parallel to the saccade vector and then relax back to their original positions after the saccade. When plotted in space and time, this remapping forms a spatiotemporal RF slanted parallel to the retinal motion, just like the slanted interaction field.

Predictive remapping may also explain distorted temporal perception during saccades. At the time of a saccade, time perception is delayed, which can be attributed to the reported longer response latency for stimuli presented in future receptive fields [129]. If RF remapping is the physiological explanation for both spatial and temporal perception distortions, these two misperceptions should be linked and co-vary across trials. Indeed, correlated magnitudes of spatial and temporal misperception have been observed.

1.2.4 The AP-Loop and Perceptual Learning

Finally, I would like to propose that the AP-loop provides the natural context for (perceptual) learning. Perceptual learning is the very foundation of intelligence; it is "a change in what was perceived." For many neuroscientists, perceptual learning refers to the process by which practice or experience improves the ability to respond to sensory stimuli [130]. For ecological psychologists who do not like the world stimulus, perceptual learning reflects an education of attention to better pick up the information in the environment [131].

In the lab, such learning has been evaluated principally on a trial-by-trial basis. In a typical perceptual learning experiment, a subject is asked to repeatedly perform the perceptual task of interest over many trials or sessions. The subject's performance is then evaluated in a time-resolved manner, allowing the learning to be quantified by comparing the performance before and after training. Importantly, in such experiments, the subject's perception and response are not self-paced but temporally constrained by the specific trial structure employed. Depriving the natural temporal structure of the AP-loop may affect learning. As discussed earlier, rhythmic action entails rhythmic perception and vice versa; it is thus conceivable that the brain may be in different affective states through the different phases of the AP-loop, and consequently, the speed of learning may depend on when the stimuli are presented relative to the AP-loops. Furthermore, in a typical perceptual learning task, the subject's current actions and perceptions have no impact on the stimuli to be perceived in the next trial. In other words, the stimulus or information to be perceived is imposed on the observer instead of actively obtained by them. This differs significantly from perceptual learning in real life.

Perceptual learning in real life is an active process. We inspect an object by looking at its parts, viewing it from various angles, and manipulating it. In all these processes, we obtain information in the continuous AP-loops and detect invariants that are not detectable without the actions we make. Notably, the visual inputs before and after the observer's action are highly correlated and are often related by a simple geometric transformation. These correlated inputs are reminiscent of data augmentation techniques used in training convolutional neural networks (CNNs), where feeding the network modified versions of an existing dataset can improve performance and generalizability [132]. It would be interesting to explore whether such correlated inputs congruent with the agent's

voluntary actions invoke a different learning process compared to learning from stimuli imposed on the observer's retina.

1.3 Ecological Visual Neuroscience

Finally, it is time to discuss the subject matter of ecological visual neuroscience, to which this thesis hopes to contribute. The word ecological is taken from Gibson's series of works, culminating in his last book, *The ecological approach to visual perception* [3]. Gibson's ecological approach treated perception as a process of direct information pickup from the environment. For the visual perceptual system, he has provided a compelling argument that the information specifying the surface layout, object, and events in the environment is available in the structured ambient light array. An observer actively obtains this information by moving their eyes relative to the head, moving their head relative to their body, and moving their body in the environment where locomotion is afforded. Importantly, for Gibson, the registration of such information is direct, without the need of mediating processes, e.g., constructing a three-dimensional world representation from two-dimensional retinal images.

Regarding visual perception as direct seems to imply that there is no place for neuroscientists to adopt an ecological approach. Indeed, as the ecological psychologists followed the motto of Mace, namely to "Ask not what's inside your head, but what your head's inside of," they have been criticized for neglecting the brain in their studies of action and perception [133]. However, that taking the ecological approach implies denying the role of the brain, I believe, is a misunderstanding. Crucially, the starting point for the ecological psychologist to analyze what information is available in the environment is the postulation that the observer's perceptual system (including the eye and the brain) can register this information in the environment. That is to say, the observer needs to have "sense organs" to perceive. Gibson disregarded the question of how exactly the observer's sense organs are attuned to "directly" picking up information, probably because he considered it outside the scope of a psychological theory. Certainly, the classification of a process as either direct and requiring no further explanation or indirect and demanding detailed investigation depends on the level of analysis. The level of analysis Gibson chose allows him to focus on the available information for an active observer to perceive and regard the processes inside the sense organ registering the obtained information as automatic. This

automatic process shall be unpacked at the level of analysis that an ecological neuroscientist would take. To unpack or physiologize the process of information picking up is the Gibsonian neuroscientists' mission.

The above passage tried to find a place for neuroscientists within the framework of Gibson's theory of direct visual perception. Despite the philosophical debate of whether perception is direct or not, pursuing ecological neuroscience has pragmatic value in research practice. Specifically, incorporating the ecological approach as advocated by Gibson and his followers implies that when specifying the visual stimuli, designing the behavioral task, collecting the data, and interpreting the results, it is important to consider 1) that in real life, the animal perceives the environment that surrounds it and this environment has intrinsic behavioral meaning to it; 2) that the environment and its meaning are specified by the structure in the ambient light array; and 3) that the retinal image is simply a frozen section of this ambient light array. In the following text, I explain in more detail the ecological approach and suggest how its key concepts may guide neuroscience research. The relevance of these concepts to the works reported in this thesis are discussed.

1.3.1 The Environment, the Information in the Ambient Light, and the Ecological Stimulation

The ecological approach emphasizes that animals perceive the environment that surrounds them. Importantly, this environment does not consist of abstract lines, planes, or spaces. Instead, it consists of the medium, substances and surfaces that are intrinsically meaningful for the perceiving animal. For instance, air as a medium affords one to see through, flat ground affords one to move around, and a partially enclosed surface, depending on its exact layout, can serve as shelter. According to Gibson, animals perceive these affordances provided by their environment, instead of perceiving the abstract physical construct or retinal image.

If perception is not based on the retinal image but on the affordances, how do we understand viewing and perceiving images on a two-dimensional monitor? Certainly, an apple displayed on a monitor is not edible, but we still recognize it and perceive it. Moreover, most discoveries in vision science since the invention of electronic displays have used stimuli presented on these flat surfaces, and dismissing these findings as irrelevant to real-world vision would be unreasonable.

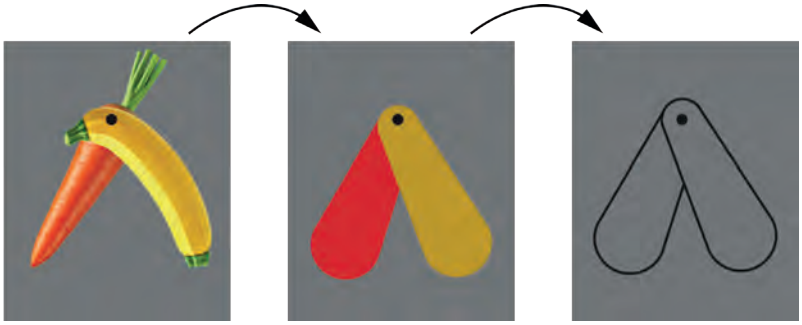


Figure 1.8: Stripping off the irrelevant information for identifying the relevant. Using overlapping natural stimuli as the spatial cue, Dowdall [134] trained the monkeys on a selective attention task in a surprisingly short time. The idea of using overlapping natural stimuli is illustrated on the left. In this illustration, the top zucchini is the cued stimulus to be selectively attended[1]. If one hypothesizes that the monkeys' fast learning is due to presenting information specifying the depth order, testing this hypothesis would then involve stripping off the irrelevant information from the natural stimuli and presenting the depth information in different formats. Chapter 2's results, obtained using the parametric filled-color (illustrated in the middle) and contour-only shapes (illustrated on the right), provides supportive evidence.

An appropriate ecological neuroscience theory should be able to explain the perception of flat images without resorting to the retinal image.

The ecological approach treats the perception of flat images on a display as indirect and mediated. Specifically, for Gibson, a two-dimensional display represents a frozen section of the ambient light array. Perceiving a flat image, e.g., an apple picture, is possible if this frozen section contains enough information to specify the corresponding distal object, i.e., a real, edible apple. This explanation fundamentally differs from the idea of a retinal image. What is perceived in this case is still the *real* apple; the relevant aspect for perception is the necessary information specifying the *real* apple. The retinal image of a miniature apple at the back of the eye is incidental to perception and is only relevant for understanding the process of perception in this specific setup and in animals with focusing lenses like ours.

Viewing flat stimuli as a frozen section of the ambient light links image-based perception in the lab to perception in the real world. Establishing this link is both theoretically interesting and practically useful, as it provides a criterion to

judge the ecological validity of different stimuli. For example, a line drawing has high ecological validity if it accurately follows perspective rules and represents perspective information as seen in the real world. Before delving into example studies reported in this thesis applying this idea (See outlined chapter 2 and chapter 5 at the end of this introduction chapter), it is helpful to specify the relevant concepts, namely ecological optics, the ambient light array, and information in the ambient light array.

The concept of ecological optics, proposed by Gibson in 1961, contrasts with physical optics. Physical optics describes the eye as a camera, where light enters either as emitted from a light source or reflected from a surface. In both cases, it is emphasized that light *radiates* from its sources. Particularly, physical optics treats an opaque surface as a collection of radiating points. Light rays radiating from each of these points are focused by the ocular lens to form a corresponding point image on the retina. The process of retina image formation, as just described, is physically correct. However, forming a retinal image itself does not solve the problem of perception. Indeed, to give meaning to the retinal image, a little man is required to see this retinal image. And if the physical optics are applied again to this little man's eye, another mini man is required to sit inside this little man to see this little man's retinal image.

Ecological optics, on the other hand, are optics with meaning to start with. It emphasizes the ecologically meaningful results of applying the physical optics in an environment that consists of scattering medium and reflecting surfaces. One of the key concepts in ecological optics is the ambient light array. The ambient light array describes the fact that after countless reflections and scatterings, the animal's surroundings are illuminated in such a way that at each potential observation point, light comes to this point from all directions. The light arrays

[1]The zucchini and carrot images used in this figure were generated using the Gemini 1.5 Flash text-to-image generation technique. The prompt for generating the zucchini was: "Generate a highly realistic image of a ripe yellow zucchini with a straight shape, pointing towards lower right. The zucchini should appear as if rendered in CAD software, with clean, defined edges and no visible shadows. The lighting should be even, with no reflections or shading. The background should be completely transparent, allowing for easy extraction of the zucchini. If possible, avoid shadows. If not possible, the zucchini should be floating to avoid in touch with the shadow." For the carrot, the prompt was: "Generate a highly realistic image of a ripe red carrot with a straight shape, pointing towards lower right. The carrot should appear as if rendered in CAD software, with clean, defined edges and no visible shadows. The lighting should be even, with no reflections or shading. The background should be completely transparent, allowing for easy extraction of the carrot. If possible, avoid shadows. If not possible, the carrot should be floating to avoid in touch with the shadow." The resulting images were processed offline to produce the left panel.

coming from all directions form the ambient light array at this point. The ambient light array at each potential observation point is structured, and it is this structure that contains information specifying the animal's surroundings, such as the surface layout. With ecological optics, seeing is to pick up information in the structured ambient light array. To pick up information, an active observer needs to sample sections of the ambient light array at a fixed observation point with gaze shifts and head movements and sample the ambient light arrays at different observation points by moving around. The active observer can explore in ways as they wish to reach the perceptual clarity or match the intended perceptual reference. Below, I focus on the information in the structured ambient light that is available for the perceiver to sample. Two important features of this information are worth noting.

First, information specifying the environment is available in various forms. For example, information specifying the depth order of two surfaces can be found in a section of the ambient array at a fixed observation point. At a proper section, the solid angles correspond to the farther surface is partially occluded by the closer surface. The same depth order information can also be specified by observing the occluding edge that progressively occludes or reveals parts of the farther surface as the observer locomotes. Importantly, it is the content of the information that is relevant for perception instead of the information's specific form. The perceptual system's invariant response to various forms of the same information in my opinion, is one of the core phenomena that needs to be explained by ecological neuroscientists.

Second, information in different formats is intrinsically linked to one another because of the common underlying surface layout it specifies. Starting from the information specifying the depth order in a section of the ambient light array, we can find the corresponding information specifying the same depth order in both the complete spheric ambient light array and the set of ambient arrays along the observer's path of locomotion. There is no ecologically meaningful information in a section of the ambient light array that cannot find its correlates in the full ambient light array or in the flow of the ambient light array during locomotion. What this implies is essentially a new definition of the naturalness of two-dimensional stimuli on a flat screen: we can now say 1) a stimulus is natural if and only if the information it specifies is natural. 2) the information is natural if the same information specified in the two-dimensional stimuli can

also be found in the ambient light array as the observer moves its eyes, head, and body.

A natural stimulus can be easily registered by the perceptual system because the perceptual system has evolved to pick up information specified by such stimulus. Conversely, an unnatural stimulus is hard to be picked up by the perceptual system because the perceptual system did not evolve to deal with such a stimulus. A direct consequence of stimulus naturalness is that it will affect the time required for training an animal to learn a specific stimulus-response association. Chapter 2 of this thesis provides an example of how stimulus naturalness, as defined here, would affect the time required for training an animal on a selective attention task. Similar ideas have appeared in literature addressing the classic stimulus-response compatibility problem [135].

Before closing this section, it is important to emphasize that an ecological stimulus does not need to be realistic. A realistic stimulus, such as those created with ray-tracing techniques, certainly has high ecological validity, but using such stimulus may not significantly advance our understanding of vision. Using realistic stimuli may introduce yet another type of retinal image that itself needs to be explained. To identify the relevant information in an information-rich environment, an ecological neuroscientist would need to present only the relevant information to the subject. They presumably need to present this information in various forms to demonstrate its relevance.

Chapter 2 of this thesis shows an example of gradually stripping off unnecessary information in the stimulus (Figure 1.8). Previous work using the depth order in overlapping natural stimuli as the spatial cue has shown that macaque monkeys can easily pick up the specified depth order and learn that this depth order specifies the to-be-rewarded behavior [134, 136, 137]. By reducing the spatial cue from natural stimuli to parametric filled-color shapes and further reducing the parametric filled-color shapes to contour-only shapes, the results obtained in Chapter 2 suggest that what is relevant for the observed fast-learning effect is probably only the information that specifies the depth order, largely independent of how this information is specified. Chapter 5 of this thesis moves even further by interpreting the repetition of a parametric stimulus as representing the information specifying stimulus occurrence frequency. The results in Chapter 5 suggest that the gamma-band neuronal oscillation power

may be the physiological correlate of registering such information in the brain.

Deciding what information is relevant and providing this information in a controlled way to the animal is practicing the ecological approach in neuroscience research.

1.3.2 The Active Observer and Their Brain

Above, I introduced the idea that for ecological psychologists, the (visual) information specifying the environment is contained in the ambient light. To pick up information in ambient light, an observer naturally moves around to sample the potential observation points. At each observation point, they also naturally move their head and eyes to sample the solid angles of the ambient light array. This is because, first, the ambient light array surrounds the observer from all directions, and second, some structures in the ambient light, e.g., the gradually occluded surface, are only revealed if the ambient light flows as the observer locomotes. In short, natural vision almost always involves an active observer.

Ecological psychologists have focused on the changes and persistence in the ambient light as the active observer interacts with the environment. By contrast, an ecological neuroscientist may be more interested in the changes and persistence in the observer's neuronal activity. Taking the ecological approach means considering these changes and persistence in the context of the observer's active sampling behavior. The observer's active sampling behavior is characterized by him engaging in the AP-loop.

An active observer exploring their natural habitat will exhibit a multifold orienting behavior, including striding, head turns, and eye movements. These movements coordinate with each other, and each has its own rhythm [138]. Here I focus on what Gibson calls the "aperture vision". In aperture vision, the subject can only move their eyes but not their head or body. However unnatural, the "aperture vision" is the first step towards understanding natural vision from the more constrained "snapshot" vision, where the eye movements are also restricted. In aperture vision, the AP-loop manifests most prominently in the rhythmic saccade sequence [111]. Understanding the changes and persistence in the neuronal activity across the saccadic eye movements is thus instrumental for understanding the active apertural vision and, eventually, the natural vision. In

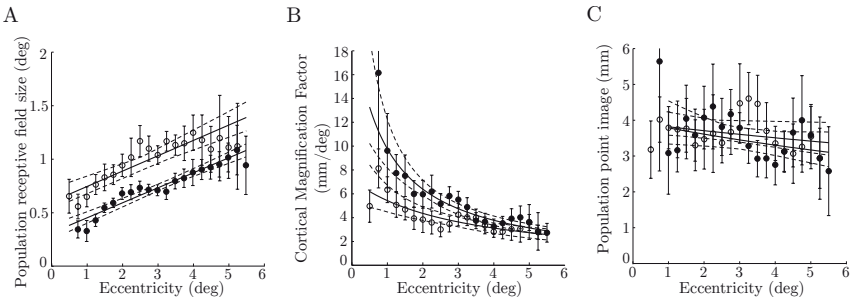


Figure 1.9: The persistent cortical point image. Using fMRI, Harvey and Dumoulin [141] examined the relationship between population receptive field (pRF) size and cortical magnification factor (CMF). As expected, in V1, pRF size increases with eccentricity (A), and CMF decreases with eccentricity (B). Importantly, the product of the two (the population point image) remains almost constant (C). Reprinted from Harvey and Dumoulin [141] under the terms of the Creative Commons license (CC BY-NC-SA 3.0)

what follows, I suggest a few mechanisms by which the environmental persistence may be registered in the observer's nervous system as they makes frequent gaze shifts (Figure 1.9, 1.10, and 1.11). Mechanisms related to receptive field remapping are left out, for it has been mentioned previously in the discussion of the AP-loop.

The Persistent Cortical Point Image

The persistence of an object that changes its retinotopic position due to eye movements may be registered in the brain by the persistent cortical point image independent of the stimulus's position (Figure 1.9). In early retinotopically organized areas such as the superior colliculus (SC) and V1, the speed of RF size increment with eccentricity is quantitatively matched to the magnification factor such that a point stimulus (smaller than the RF) in the environment will always activate a neuronal population of similar size in that area [139–141]. In other words, there is a persistence of the cortical point image independent of changes in the stimulus's retinotopic position caused by the observer's eye movement.

Results in Chapter 4 of this thesis are related to such point image persistency in the SC. Chapter 4 notices an incidental effect of reaching the constant point image via increasing RF size in the periphery. While increasing RF size compensates for the decreasing cortical magnification factor in the periphery,

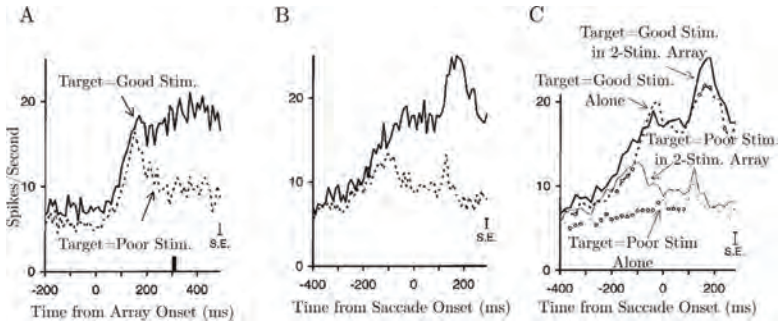


Figure 1.10: The persistent saccade target information. Recordings from [143] revealed a persistent saccade target information in IT neurons. The responses of IT neurons during a memory-guided visual search task are persistent, whether the response is aligned to the search array onset (A) or to the saccade onset (B). In both cases, the peri-saccadic selectivity is stable. Importantly, the responses obtained when there are two stimuli in the RF are similar to if only the saccade target is presented (C). Reprinted from Chelazzi et al. [143] with permission from the American Physiological Society.

it also means that physically identical stimuli in the peripheral visual field would occupy a smaller proportion of the RF of the corresponding neuron. Less receptive field coverage will likely cause the neuron in the periphery to be less activated, leading to a longer saccade onset latency. This is what we observed when the monkey was asked to make a response saccade to a peripheral stimulus onset. When we placed physically identical stimuli at various locations, the monkey's response time was longer with more peripheral saccade targets. Importantly, when we scaled the stimulus size at each eccentricity according to the corresponding SC neurons' RF size to equalize the receptive field coverage across target eccentricities, the response times to these scaled targets were equalized. The results in Chapter 4 thus not only support the existence of a constant point image but also show that the specific implementation of the persistent point image may have incidental consequences on the animal's action and perception.

The Persistent Saccade Target Information

Secondly, neurons in higher visual areas may register the persistence of the saccade target by mediating the communication between foveal and peripheral neurons in lower visual areas. This mediation is enabled by the size of RFs increasing progressively in the hierarchically organized visual pathways [142].

Consider, without loss of generality, neurons corresponding to the saccade target five degrees to the right of the current fixation [144, 145]. At the level of primate V1, these neurons' RF size is around half to one degree in diameter. At the same eccentricity, from V1 to V4, the RF size gradually increases to about 3 to 4 degrees in diameter. The relatively small RFs in early and intermediate visual areas imply that the saccade target will fall outside these neurons' RF after the eye lands on the saccade target. However, as the RF sizes continue to grow along the visual hierarchy and eventually cover the fovea, saccade targets will also eventually stay within the RF after the saccade. This is the case for the inferior temporal cortex (IT) neurons. In other words, if no additional stimuli exist, a persistent response can be achieved in IT neurons simply due to the RF expansion. When multiple stimuli coexist, attentional mechanisms can help reduce the multi-object scenario to the single-object scenario [143, 146]. Specifically, attentional selection can bias the IT neuron's response so that it responds to multiple coexisting stimuli in the RF as if only the attended one was present (Figure 1.10). Consider an observer making a saccadic eye movement in a crowded scene. Prior to the saccade, attention is lawfully directed towards the saccade target. After the saccade, attention dwells naturally on the same object that is now foveated. In both cases, it is the same object that is attended to and consequently dominates the IT neuron's response perisaccadically. Therefore, the IT neuron's response can still be persistent perisaccadically in a crowded scene.

Achieving persistent response in IT is only part of the story. The other part involves transmitting this persistent response to the foveal neuronal population in lower visual areas processing the foveated saccade target. Notably, because the IT neurons' RFs typically contain the fovea, it is conceivable that these IT neurons would also provide feedback signals to the foveal neurons in the lower visual areas with which they share receptive fields. Importantly, these feedback signals to the foveal neurons processing the saccade target post-saccadically result from the feedforward activation driven by the peripheral neurons processing the saccade target pre-saccadically. Accordingly, IT neurons can be described as mediating the communication between foveal and peripheral neuronal populations, each processing the same object at a discrete episode. In this suggested process, the information specifying the saccade target is persistent and registered in the feedforward and feedback connections of the visual hierarchy. Because of this persistent information, saccadic eye movements

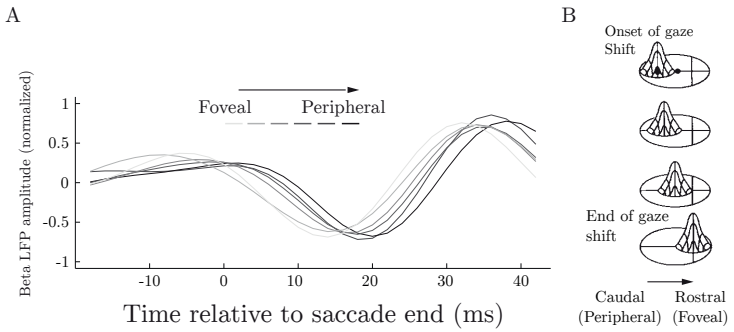


Figure 1.11: The persistent dynamics. (A)Zanos et al. [151] show that a traveling wave of oscillatory local field potentials in V4 is triggered by saccades. This wave moves from the fovea to the periphery and may be related to selecting the next saccade target. (B) Munoz, Pélisson, and Guitton [152] reported a moving hill during saccade execution from the caudal part of the SC representing the peripheral saccade target to the rostral end of the SC representing the fovea. This traveling wave is time-locked to saccade execution and may contribute to gaze control during saccade flight time. Panel A is reprinted from Zanos et al. [151] with permission from Elsevier. Panel B is reprinted from Munoz, Pélisson, and Guitton [152] with permission from AAAS.

can be viewed as part of the process in which the observer actively clarifies this information, from coarse (prior to saccade) to fine (after the target is foveated) [147].

The Persistent Dynamics: Two Traveling Waves

Finally, the persistence of trans-saccadic perception may be registered by the invariant spatiotemporal neural dynamics around the time of the saccadic eye movement. I suggest two dynamics. The first is a potential traveling wave from the fovea to the saccade target, corresponding to the obligatory pre-saccadic attentional shift [148, 149]. The second is also a potential traveling wave, but from the saccade target to the fovea, matching the main sequence of saccadic eye movements [150]. I discuss each of the two waves separately and in sequence below.

When there are no attention-capturing events, attention dwells naturally at the foveal, where we are looking at. As the observer explores the environment with saccadic eye movements, prior to each saccade, attention is obligatorily shifted to the saccade target. First, for linking this shift to a travelling wave from the

fovea to the saccade target, I note that the attentional shift is not a jump but takes time [153].

The time the shift takes can be estimated from visual search tasks. Most commonly, the attentional shift time can be estimated by examining the relationship between the set size of the search array and the time required to conclude that the target is absent. Using this method, the time required for each shift is estimated to be 25∼100 ms [88]. Such an estimation may be inaccurate, as it assumes a perfect one-by-one serial search. A different estimation can be obtained directly from physiological data. Recordings from macaque FEF, IT, and V4 with relatively simple search arrays have shown that the attention signal appears about 125∼175 ms after the search array onset [143, 154, 155]. After subtracting the visual delay, which may correspond to the arrival of pre-attentive information, the time required for attention to shift covertly from the fovea to the target item is about 75∼125 ms. Despite the discrepancy in the exact time it takes for an attentional shift, the fact that the shifting process *per se* requires time is generally accepted.

Chapter 4 of the current thesis provides evidence that the time required for attentional shifts increases with the shifts' amplitude. When the animal was signaled foveally to initiate a saccadic eye movement towards a peripheral saccade target, the time it took to initiate the saccade was found to increase with target eccentricity. One potential explanation for the observed results is that attention travels from the foveal to the periphery until it meets (or selects) the next saccade target. Physiologically, this may correspond to a traveling wave in retinotopically organized visual areas. Travelling waves from the fovea to the periphery have indeed been described. By recording with two dimensional multielectrode arrays chronically implanted in V4, Zanos et al. [151] found that saccades can trigger robust travelling waves of oscillatory local field potentials from the fovea to the periphery (Figure 1.11 A). Accompanying this wave is the propagation of high excitability from the fovea to the periphery, which may contribute to the selection of the next saccade target.

The second wave likely accompanying each saccadic eye movement is from the periphery to the fovea. Such a wave has been observed in the SC of cats and is described as a moving hill (Figure 1.11 B) [152]. At the saccade onset, the peak of the moving hill is initiated at the location corresponding to the saccade

target; during saccade execution, the moving hill moves towards the fovea; when the moving hill reaches the fixation zone at the rostral end, the saccade stops. Because the moving hill is observed at the motor layer in the SC and its dynamics is locked to the saccade execution, it is hypothesized to represent the gaze position error and, therefore, to provide continuous control of the eye movement during the gaze shift [156]. Similar traveling activity moving from the saccade target to the fovea has also been noted in the primate SC. However, there is some debate about whether this moving activity in the primate SC can contribute to gaze control during saccade execution [157].

So far, the discussion has been focused on the moving hill in the motor layer of the SC. Traveling waves from the periphery to the fovea may also occur in the sensory layer of the SC or other visual areas. In the SC, the visual layer, as the motor layer, is retinotopically organized. In addition, the retinotopic maps of these layers are registered. Neurons across the layers' corresponding locations are reciprocally connected via interlaminar connections. There is, thus, presumably, a corresponding traveling activity in the visual layer of the SC caused by the excitatory synaptic feedback from the motor layer. Other mechanisms may also give rise to sensory waves moving towards the fovea. The excitatory motor feedback causes the pre-saccadically enhanced visual response to the saccade target in various retinotopically organized visual areas, including the SC and V4 [158, 159]. This enhancement is akin to reactivating the saccade target whose onset transient has long subsided. Reactivating the saccade target at the peripheral site might trigger a traveling wave that propagates towards the foveal representation in retinotopically organized areas, similar to how a peripheral stimulus onset transient may trigger a traveling wave towards the fovea [160].

What could be their function if the traveling waves that move from the saccadic target to the fovea exist in the sensory areas? One possibility is that it may facilitate the communication between the peripheral neurons processing the saccade target pre-sacadically and the foveal neuron processing the same stimulus post-sacadically. Because the horizontal connectivity within a brain area is biased to similarly tuned cells (Figure 1.3 D), the traveling wave across such a surface will likely transmit feature-specific information and prime the foveal neurons for the upcoming stimulus [38]. Such facilitation may contribute to the preview-benefit in the reading literature and more recently reported

trans-saccadic memory in the post-saccadic foveal neuronal activity [161, 162].

1.3.3 The Call for New Behavioral Tasks

The previous two sections have discussed important aspects of the ecological approach. In the first section, by introducing the concepts of ambient light and the information it contains, I aim to convey that adopting an ecological approach involves presenting behaviorally meaningful information to the observer rather than merely stimuli to the retina. In the second section, by suggesting how invariances or persistence in the ambient light may find their correlates in the active observer's neuronal response, I aim to convey that adopting an ecological approach involves engaging the subject in the AP-loop interacting with the environment. The two requirements call for adopting new behavioral tasks to study visual perception.

These tasks should utilize ecological stimuli that specify natural information in the environment. These tasks should also allow the subject to respond accordingly to the specified information. For instance, if a stimulus appears in the periphery, the subject should be able to look at it. If multiple stimuli are presented, the subject should be allowed to fixate on them sequentially. Additionally, similar stimuli should have similar behavioral meanings. This list can continue, but importantly, just as presenting ecological stimulation does not necessarily mean using realistic pictures, meeting these requirements for ecological validity also does not necessarily mean abandoning the reductionist approach or losing control over retinal stimulation. Chapter 3 presents a controlled foraging task that represents work in this direction.

1.4 Overview of the Research Chapters

This section provides an overview of the ensuing research chapters, contextualizing their respective results within the ecological approach discussed earlier. Unlike the abstracts of the individual chapters, which are intended to be read independently, this outline highlights how each chapter's results contribute to the integrated message of the entire thesis: advancing towards ecological neuroscience. Some of the contents mentioned here have already been covered in the earlier sections of the introduction and are reiterated for completeness.

Chapter 2: Behavioral Benefits of Attention and Reaction-Time Dependent Learning in a Foveation-Based Selective Attention Task for Macaques

Taking the ecological approach emphasizes the importance of using behavioral tasks with high ecological validity. The ecological validity of a task is reflected in how fast an animal can learn this task. For simple tasks, the animal's speed of learning, i.e., to do the task according to the instruction, is not to be confused with the difficulty of the task itself, i.e., to achieve high performance in this simple task. The way task instructions are communicated to the animals can significantly affect their learning speed. In the context of taking the ecological approach, accelerating learning speed can be achieved by providing task instructions that the animal's perceptual system is attuned to.

Macaques, as often used to study the neural correlates of cognitive processes, are undoubtedly intelligent animals. However, they often take months to learn seemingly trivial selective attention tasks. This is likely because commonly used attentional cues, which serve to transmit task instructions, are not natural for them. Impressively, Dowdall et al. [136] developed a fixation-based selective attention task that macaques can learn in just a few sessions. The key to this success, in the context of the ecological approach proposed in this thesis, is probably the use of a natural surface layout, specifically, the overlapping surfaces, as the attentional cue. Overlapping surface layouts are ubiquitous in the environment. The animals' perceptual system has likely evolved to pick up the information specified in these layouts. Using the overlapping surfaces as the attentional cue can, therefore, facilitate learning.

Chapter 2 reports an adaptation of the fixation-based selective attention task developed by Dowdall et al. [136] (see also Dowdall [134]). This adaptation maintains the essential information specifying the overlapping surfaces but reduces the attentional cue from natural images to parametric shapes and eventually to shape outlines. Importantly, the animals' learning speed with the parametric shapes is similarly fast. This supports the idea that it is the naturalness of the information (in this case, the overlapping surfaces) rather than the specific forms of this information (realistic surfaces, filled shapes, or outlines) that determines the attentional cue's ecological validity. Additionally, Chapter 2 notes that when conflicting information is present, particularly if

it tends to be picked up by the animal and triggers a response of the same modality as the required response, the learning process is interfered with. Specifically, Chapter 2 shows that learning speed depends on the response times. When the response time is short, indicating that the response is likely driven by the cue-independent salient saccade target onset (which represents conflicting information), learning is hampered. Finally, Chapter 2 demonstrates the attentional benefits in behavior using the foveation-based tasks. This was achieved by converting the task into a subset of Posner's classic cueing conditions, specifically the neutral-cue and valid-cue conditions. The invalid-cue condition was intentionally omitted because it requires the animal to make an unnatural response.

Chapter 3: A Controlled Foraging Task for Studying Natural Vision

Adopting the ecological approach suggested in this thesis involves viewing perception as an active process in which the observer explores the environment to pick up information from the ambient light. This perspective necessitates new visual tasks that allow subjects to interact with the stimuli. For the visual perceptual system, the relevant actions include moving the eyes, head, and body during manipulation or through locomotion. A pragmatic approach would be to gradually incorporate these movements and their respective feedback into the behavioral tasks. The first step in this process is to allow eye movements in head-restrained animals.

Chapter 3 introduces a foraging task that allows animals to freely move their eyes in search of a hidden reward. The animal's search is guided by information within the search array, which can be flexibly modulated by the experimenter. By embedding symmetry in the search array, the experimenter can also control the retinal input during the search, akin to the control in traditional paradigms where the animal's eye movements are restricted. This proposed foraging task is thus not only interesting on its own but also enables the re-examination of classical results, such as feature tuning properties along the ventral pathway, within a behavioral context of high ecological validity. The foraging task's high ecological validity is manifested in the exceptional learning speed: in just one session, two macaques learned the expected behavior, a sequential search on physically identical search arrays where the target is probabilistically defined.

Chapter 4: Eccentricity Dependent Saccadic Reaction Time: The Roles of Foveal Magnification and Attentional Orienting.

Ecological psychologists focus on the changes and invariances in the ambient light as the observer actively explores the environment. The envisioned ecological neuroscientists, however, may be more interested in the changes and persistence in the observer's neural activity during this active exploration. For the simplest kind of visual exploration, which involves a sequence of rhythmic saccadic eye movements, I have identified earlier a few invariant neural activities across the composing saccades. Chapter 4 explores two of these invariant neural activities.

The first invariant neural activity explored is the persistence of the cortical image. In the primate visual system, foveal representation is disproportionately magnified in retinotopically organized visual areas. This foveal magnification can be quantified by the cortical magnification factor (CMF), which measures how much a brain area corresponds to the same visual field angle at different eccentricities. Notably, there is a reciprocal relationship between CMF and receptive field (RF) sizes in areas like the primary visual cortex and the superior colliculus. Multiplying CMF by RF size thus yields a constant value that corresponds to a persistent cortical image size for a point stimulus, regardless of its retinotopic location. This implies that despite changes in retinotopic location due to saccadic eye movements, the size of the cortical image remains persistent.

Chapter 4 explores the relationship between the CMF and RF sizes. It hypothesizes that the often-observed eccentricity-dependent reaction time increment is due to the increasing RF sizes as one moves from the fovea to the periphery. By scaling the saccade target size according to the CMF, we were able to equalize the response time across the tested eccentricities. These results not only support the existence of a persistent cortical image but also show that the specific implementation of registering environmental persistency—having the same cortical point image register the same point stimulus that affords looking at—may have incidental consequences on the animal's action and perception.

The second persistence relevant to the results presented in Chapter 4 is the invariant spatiotemporal dynamics of pre-saccadic attentional orienting. It is known that prior to each saccadic eye movement, attention obligatorily shifts

towards the saccade target. Earlier in the introduction, this shift was suggested to be linked to a traveling wave from the fovea to the saccade target. A traveling wave mechanism implies distance-dependent traveling time. Chapter 4 provides supportive evidence for this hypothesis by showing that the response time of a foveally triggered voluntary saccade depends on the saccade target's eccentricity. Notably, this relationship persists even when a foveal flash is applied to reset attention to the fovea (likely also reinitiate the traveling wave) during saccade planning.

Chapter 5: Gamma-band Neuronal Synchronization Tracks Prior Updates of Stimulus Occurrence Frequency

The core of the ecological approach is to view perception as a process of information pickup. This information is not the same as Shannon's information. The environment does not communicate to the observer, nor does the observer passively receive information through sensory channels. Instead, the information of interest to ecological psychologists and ecological neuroscientists is the "information-as-specification." Information defined as such specifies properties of the environment that are behaviorally relevant to the observer.

The information-as-specification can take various forms. It can be concrete, such as information specifying a falling-off edge, or more abstract, such as information indicating a resource-rich area where food is likely to be found. The richness of resources can be defined statistically, such as by the fruit density in forest canopies or the insect abundance on tree bark. For an active observer exploring the environment, the density or abundance of food can be translated into the frequency of stimulus occurrence. Chapter 5 investigates how the brain picks up information that specifies the frequency of stimulus occurrences.

Specifically, Chapter 5 proposes that the stimulus occurrence frequency is represented in the brain by the gamma-band neuronal synchronization, and picking up this information means integrating the prior knowledge about the stimulus occurrence frequency with newly obtained sensory evidence. If this hypothesis is correct, one should be able to see the corresponding changes in gamma-band neuronal synchronization as the new observations are obtained.

Chapter 5 tests this hypothesis by analyzing the dynamics of gamma-band neuronal synchronization in a repetition task. Specifically, the repetition task

was modeled as a series of Bernoulli trials with a beta-distributed prior over stimulus occurrence frequency. It was hypothesized that gamma-band local field potential (LFP) power indexes the mode of updating the prior distribution. Remarkably, with only three parameters, this model explained over 80% of the variability in the complex pattern of gamma-band LFP power variation over the session. This included the non-linear increase in power with repetition, the partial persistence of the repetition effect over an intervening break, and the decreasing power with interleaved repetition, but only after adaptation. Notably, if gamma-band neuronal synchronization indeed tracks prior updates, it offers a straightforward mechanism for updating prior information through Hebbian plasticity and computing with such prior information through coherence. Together, these mechanisms may underlie how the brain picks up stimulus occurrence frequency.

CHAPTER 2

Behavioral Benefits of Attention and Reaction-Time Dependent Learning in a Foveation-Based Selective Attention Task for Macaques[1]

2.1 Abstract

One of the most central cognitive functions is selective attention. It has been prominently studied in the visual system of macaque monkeys, for which they have been trained on a variety of tasks. However, this training often takes many months. Recently, Dowdall et al. [136] developed a foveation-based selective attention task and showed that macaque monkeys can learn the task surprisingly fast. Here, we describe an adaptation of the reported task to ease its usage and extend its application. The adaptation includes parameterization of the spatial cue, removal of the color cue, and placing additional target (distractor)

[1]Results reported in this chapter are based on data collected by Yufeng Zhang and Julien Vezoli from Monkey CH and HU. Learning trajectories for these two animals have been presented previously in a different format in Dowdall et al. [136, 137]

stimulus at the cued (un-cued) location. Interestingly, while analyzing the learning trajectories of two monkeys, we found that trials with longer reaction times showed more rapid learning, presumably reflecting the relatively late arrival of the attention-related signal in generating the response saccade. We also demonstrate attentional benefits in behavior: validly-cued compared to neutrally-cued probes resulted in faster reaction times in both monkeys and higher contrast sensitivity in one of them. The further development of the foveation-based selective attention task focuses on the essential aspects of selective attention, avoids unnecessary additional cognitive demands, and builds on natural, automatic processes. This approach might be extended to studying selective attention in other animal models.

2.2 Introduction

Selective visual attention is a central cognitive function [80, 163]. To understand its neuronal mechanisms, we need to study it in animals, particularly non-human primates. For this, we need to operationalize attention in a behavioral task. One class of these tasks (*Posner tasks*) has its root in Posner's seminal studies in the 1980s [51, 72, 91, 164–174]. It operationalizes attention as measurable behavioral benefits in a collection of trials, e.g., faster reaction time (RT) or higher sensitivity. In these studies, one out of several locations is cued, such that subjects are expected to pay attention to the cued location. Importantly, regardless of which location is cued, both locations can contain the probe stimulus, which the subject needs to respond to. In other words, the cue used in a Posner task is not always valid regarding where the behaviorally relevant probe will appear. Indeed, in a Posner task, successful manipulation of attention needs to be shown as enhanced performance when the cue is valid compared to when the cue is invalid (while the physical stimuli are kept identical).

A different operational definition of attention has also been used in monkey studies, in fact, more often [32, 70, 84, 87, 89, 90, 146, 175–210]; we refer to the respective tasks as *filtering tasks*. Compared to Posner tasks, the cue in a filtering task is always valid, i.e., the cued location will always contain the probe (unless a trial is a catch trial, in which the correct behavior is to not respond since no probe appears). Given the 100% cue validity, to demonstrate that the animal's attention has been successfully oriented to the cued location, a filtering task requires at least one task-irrelevant distractor to be presented at a different

site. To perform the task, the animal needs to ignore (filter out) the distractor. It is important to note that unlike in a Posner task, attentional benefits in a filtering task cannot be shown as better performance at the cued location than the un-cued location. This is because, by definition, the target probe never appears at the un-cued site. There are thus no trials like the invalid-cue trials in a Posner task to compare the behavior to. As a result, in the filtering tasks, attention is operationalized not as enhanced behavioral performance but as timely responding to the probe and ignoring (filtering out) the distractor when presented.

The above mentioned differences between Posner tasks and filtering tasks strongly influence the experimental design and animal training. Macaque training uses principles of positive reinforcement learning [211, 212]. To achieve the reinforcement, the animal needs to associate the reward with the desired behavior. With a Posner task, the benefit of correctly orienting attention to the cued location is only manifested when the average performance is considered in a collection of trials. This is inconsistent with the fact that a reward is provided after every successful trial, i.e., correct response to the probe at the cued or un-cued location. Such mismatch makes it difficult for the monkey to associate a reward value to the cue. For comparison, in filtering tasks, only the cued location contains the target. Correct orienting of attention is indicated by proper behavior in every single trial and will be rewarded immediately. Therefore, training is easier with filtering tasks, because it matches the temporal structure of the reward. To be more specific, in a filtering task, if the subject by chance reacts according to the cue, it will be rewarded right after the current trial. Thereby, this behavior, orienting attention according to the cue, will be reinforced. By contrast, in a Posner task, it will only get slightly better performance if it by chance orients attention according to the cue in a row of many trials. Trials with the invalid cue are particularly detrimental for training. Correct orienting of attention according to the invalid cue would actually be "punished" immediately for a delayed or absent response to the probe. Reinforcement learning of orienting attention to the cued location is thus difficult, especially when the overall cue validity is low. Accordingly, except for few cases [166, 169], previous studies using the Posner paradigm in macaques, even after extensive training, adopted high cue validity and block design [72, 168, 170, 172, 174, 213–216]. Note that Posner, in his original proposal [82], argued against block design: "...our basic method involves cuing on each trial.

If instead one spatial position is made likely for a whole block of trials, we found no benefits for the frequent position in comparison with conditions in which all positions are equally likely...". Block designs have also been used in filtering tasks. In these cases, the use of block design was most likely aimed at reducing the cognitive effort incurred by trial-by-trial reorienting of attention [217]. However, compared to a block design, trial-by-trial randomized cueing of attention has important advantages: (1) It minimizes any spurious correlation between experimentally controlled attention and slow changes in the state of the animal across blocks; (2) It minimizes interference by any slow changes that are unrelated to attention within a block (e.g. adaptation to repeated stimulation, learned location-reward association).

To achieve a trial-by-trial randomized design, it is helpful to make the task easier or more natural to the monkeys. Efforts have been made in this direction. For example, to make the cueing more natural, some studies used a colored fixation point as the spatial cue [191, 208, 210]. The to-be-attended stimulus was the one whose color matched the fixation dot color. Cues like this are natural, because the fixation point will always receive attention when fixated. If the fixation point is colored, and the fixation-related attention is thereby directed to that color, this feature-based attention is automatically also partly deployed to equally-colored stimuli in the periphery [95]. Importantly, in this case, neither the color of the fixation point nor the color of the target determines the location of the attentional target, rather it is the conjunction of both that defines the target location. By such design, the biggest confound is eliminated, i.e. any observed attention effect cannot be explained just by the difference in the color used for the fixation point or the target stimulus. This is important because when the cue itself is not a source of confound, it is possible to keep the cue on the monitor until the end of the trial, such that working memory demand is also minimized.

Studies that used the above-mentioned filtering task design often can achieve a fully random design where attention is assigned on a trial by trial basis [191, 208, 210]. But there are still other aspects that can be improved. In many above-mentioned filtering tasks, the distractor (foil) can appear before the target in half of the trials [70, 90, 146, 176, 190]. When this happens, requiring animals to restrain from responding to the early distractor onset is rather unnatural, because the distractor onset automatically captures attention and

tends to trigger a behavioral response (especially saccades). The cue can also be modified to be even more intuitive. An interesting development of the cue can be seen in the curve tracing paradigm. In this paradigm, a curve connecting the target to the fixation dot serves as the spatial cue [218]. By this arrangement, the target is actually part of the object that is fixated and thus naturally attended. Besides the intuitive cue, working memory demand and interference of the distractor onset are also minimized. All these, for reasons discussed above, make the task natural for the monkeys to perform. The application of the curve tracing task is, however, limited. The stimuli used in the curve tracing task were essentially segments of the curve. An ideal task should allow a flexible choice of stimuli. This is of particular importance if neurons in higher visual areas with their complex feature selectivity are of interest.

In a previous study, Dowdall et al. [136] developed a new foveation-based selective attention task. In this task, two or more stimuli (natural or artificial) overlap at the fixation point, and the depth ordering serves as the attentional cue: the foveated foreground stimulus cues attention to this stimulus and the probe that might later appear on it. They showed that macaque monkeys (including monkey CH and monkey HU from the present study) can rapidly learned this attention task (see also Dowdall [134]; Dowdall et al. [137]). This thesis chapter first, as background, presents and re-analyzes the learning trajectories of monkey C and monkey HU from Dowdall et al. [136, 137]. In addition to the already reported rapid learning, we report that: (1) Learning was facilitated by minimizing the possible stimulus positions. (2) Learning was faster when reaction times were longer. In addition to the learning trajectories, in this chapter, (1) we describe in detail a version of the task that uses parameterized color shapes, which allows to fully control the attentional cues (see “Experiment 1, multi-probe task” in the Methods section), (2) describe a Posner version of the task, which allows the investigation of attentional effects on RT and/or sensitivity (see “Experiment 2, single-probe task” in the Methods section), (3) describe a version of the multi-probe task in which the color is removed from the cue shapes, which further avoids potential influences of cue color and eases the use of arbitrary stimuli inside the cue shapes (see last paragraph of “Experiment 1, multi-probe task” in the Methods section).

These tasks are both intuitive and easy for the monkeys to learn and are also flexible in terms of stimuli selection and spatial arrangement. We trained

two macaques and recorded behavioral data during learning. We first trained the animals on a version of the task that is, by the above classification, a filtering task (multi-probe task). Once the macaques performed at a high level consistently for several sessions, we kept the filtering task for 20% of the trials, while 80% of the trials were adapted to constitute a Posner-like version of the task (single-probe task). In this version, trials with valid and with neutral cues were randomly interleaved. In this Posner-like task with trial-by-trial cueing, attentional benefits on reaction times were significant in both animals, and attentional benefits on detection performance were significant in one animal and trending in the same direction in the other.

2.3 Materials and Methods

2.3.1 Subjects

All procedures in this study complied with the German and European law for the protection of animals and were approved by the regional authority (Regierungspräsidium Darmstadt). Two adult male macaque monkeys (Macaca mulatta) participated in the current study. Monkey CH was 13 years old. Monkey HU was 11 years old. Both monkeys were implanted with a titanium head-post [219]. The two animals were housed in separate rooms that each contained other monkeys. Within the period of data collection, the animals' water intake was regulated so that they were motivated to perform the task. They received small amounts of juice as reward after each correct trial. Before participating in the tasks described in this study, both monkeys had been trained on a visual detection task. In the detection task, they needed to keep fixation on the central fixation dot and detect a local contrast change (target) in the periphery. The monkeys needed to report the change location by making a saccade to it. In all detection tasks that the two animals performed prior to the tasks described in the present paper, only one target was presented. Both animals therefore were familiar with other aspects of the main tasks described in this study, including the shape stimuli, but had no experience with tasks requiring the use of any kind of spatial cues.

2.3.2 Visual Stimuli and Behavior Control

All visual stimuli were presented on an LCD monitor (Samsung 2233RZ) controlled by custom software (https://github.com/esi-neuroscience/ARCADE).

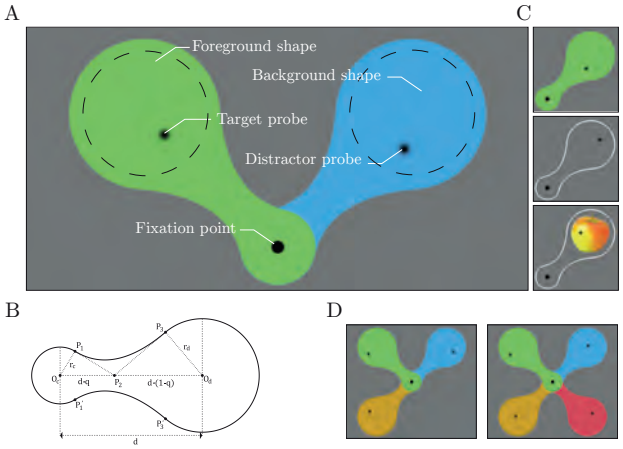


Figure 2.1: Stimuli arrangement. (A) An example display comprises a fixation point, a foreground shape, a background shape, and the later appearing target and distractor probes. The dashed circles indicate the area where the probes can appear and are not shown in the actual display. (B) The shape can be defined by color (top) or contour (middle). The target/distractor probe can appear on the distal part of the shape (top and middle) or on a separate stimulus placed on the distal part of the shape (bottom). (C) Parameterization of the shape. The shape used in the present study can be constructed by two open circles (O_c and O_d) and two 3-point Bezier curves ($\widehat{P_1P_3}$ and $\widehat{P_1'P_3'}$) tangent to the circles. Four parameters were used to define the illustrated shape: r_c, the radius of the central circle O_c; r_d, the radius of the distal circle O_d; d, the distance between two circles; and q, the ratio $\frac{O_cP_2}{O_cO_d}$. (D) In Experiment 2, different conditions contained different numbers of background shapes, as illustrated here.

Viewing distances for monkey HU and monkey CH were 75.5 cm and 77 cm, respectively. A video-based eye tracker (EyeLink-1000, SR Research Ltd.) was used to record gaze and pupil signals from both eyes. The gaze of one eye was used for experimental control. Before the start of each session, we performed a standard 9-point calibration. All stimuli presented in the present study were within the area covered by the calibration points.

A filled black circle (0.2 degrees) presented at the center of the screen was used as the fixation point in all tasks. Depending on the specific task, some of the following stimuli were presented.

1. A background shape and a foreground shape of the same form serving as the attentional cue (Figure 2.1). One end of the two shapes overlapped in the screen center. The other end of the two shapes pointed in different

directions. Two shapes were separable as two overlapping objects, with the foreground shape on top of the background shape. In the Neutral-cue condition in Experiment 2, the overlapping region is occluded such that the foreground-background relationship, or depth order, cannot be determined. An example parameterized shape (Figure 2.1 C) was constructed by two open circles connected by two 3-point Bezier curves. The control points of the Bezier curves were designed so that the two curves were (1) tangential to the central circle and the distal circle, (2) symmetric about the axis connecting the two circle centers. Two control points needed to be placed on the circles, and the third point was placed on the axis connecting the circles' centers to achieve these two goals. Four parameters were thus required to define the foreground or background shape: r, the radius of the center circle; R, the radius of the distal circle; d, the distance between the two circles; and q, the ratio of the segments connecting the circles' centers to the third control point.

2. A target stimulus and a distractor stimulus, which the monkey needed to pay attention to or ignore. These stimuli are intended to be used later, in neurophysiological experiments, to activate neurons in different visual areas; those neurons often have specific stimulus selectivities, e.g., for objects. The target stimulus could be an object of any kind presented on top of the distal part of the foreground shape. The distractor stimulus is of the same kind as the target stimulus but presented on top of the background shape. When target/distractor stimuli were not used, the distal part of the shapes functioned as the target/distractor stimulus on which the probe could appear. In the present study, for a given session, either the shapes themselves or some natural objects were used as the target/distractor stimuli. Natural object stimuli used in the present study were selected from Hermera Photo-Objects Vols. 1, 2, and 3 (Hemera Technologies) [61].

3. The probe(s). A probe stimulus was a black circular Gaussian blob. When the probe appeared on the target stimulus, it was a target probe; when the probe appeared on the distractor stimulus, it was a distractor probe. The monkey was required to make a saccade to the target probe and, at the same time, ignore the distractor probe if it appeared simultaneously. The target and distractor probes were of the same eccentricity (with

jittering of up to ±0.05 degrees) for every given trial and only differed in their spatial location. For a given session, the size of the probes was fixed. Sessions in the cue-learning phase used relatively large probes to make the detection easy. Smaller probes were used in Experiment 2 to assess the attentional benefit. In Experiment 2, the probes' contrast (Weber contrast relative to the shape's face color) was further adjusted by manipulating the opacity of the probe to control the task difficulty. To sample the psychometric curve, a set of contrast levels was used for each session. For each session, the range of possible contrast levels was manually selected based on the performance in the last session to achieve two goals: (1) an overall hit rate of ≈70%-80% to keep the animal motivated, (2) a sampling of the entire psychometric curve, including contrasts below detection threshold and those with close to perfect detection. In all tasks, the probe stimulus always appeared within the distal part of the shape, with the precise location within the distal part varying across trials. When natural objects were used, possible locations of probe appearance were further restricted to be on top of the object. The variation of the probes' precise location defines the spatial extent of the expected focal attention before probe presentation. We consider this useful, because we are interested in studying effects of attention on the colored distal part [220] or on extended stimuli placed on that distal part (see Results section on *Modification of the shapes*). We want to prevent the animal from paying attention to a single, fixed location.

In Experiment 2, more than one background shape could be presented (Figure 2.1 D). When this was the case, distractor stimuli and distractor probes were presented accordingly on these extra background shapes.

2.3.3 Behavioral Tasks

Experiment 1, Multi-probe Task

Dowdall et al. [136, 137] have previously outlined the multi-probe task used in the current study and referred to it as Task 2. These earlier works present two variants of Task 2: one version, used on four monkeys not covered in this report, uses natural objects as the cue against a grass-like background; the other version, used on monkeys CH and monkey HU, is identical to the one used here during the cue-learning phase and is detailed below. Note that the

four monkeys trained with natural object cues in Dowdall et al. [136, 137] were also trained on another task referred to there as Task 1. The two monkeys used in the current study were not trained on Task 1.

The multi-probe task is a filtering task by the above-described classification, with the foreground and background shapes serving as the spatial cue. It was used to introduce the spatial cue. Figure 2.2 A illustrates a typical trial. It started with the animal holding its gaze on the fixation dot for 1300 ms∼1800 ms. Following this period, the foreground and background shapes were presented for a random time chosen from a uniform distribution between 650 ms and 3800 ms. In the data collected, we used a uniform distribution between 650 ms and 2100 ms for monkey CH, between 650 ms and 5800 ms for monkey HU from session 1 to session 13, and between 650 ms and 3800 ms for monkey HU from session 14 to session 18. The relatively long duration was chosen, because it aids the investigation of neuronal responses during sustained attention, similar to several previous studies [70, 90, 176, 179, 185, 221]. The efficient use of this long duration is further aided by the uniform distribution, because it provides relatively many long trial durations. A noteworthy side effect of this is that the hazard rate increases with time, which can affect behavior and neuronal activity [85, 87, 222]. This can be mitigated by approximating a flat hazard function with an exponentially distributed delay duration, at the expense of less data from long-duration trials and including catch trials. After this period of presentation of foreground and background shapes (and eventually stimuli on top of the shapes), a target probe and a distractor probe were then presented simultaneously. Positions of these probes were randomly decided as long as they were located on the corresponding shape and were equally eccentric (with a jittering up to $\pm 0.05^\circ$). When the probes were presented on the screen, the animal needed to make a response saccade within 1000 ms to the target probe, while ignoring the distractor probe. If the animal held fixation on the target probe for another 250∼650 ms, a small amount of juice (as the reward for correct behavior) was delivered. If the animal broke fixation during a given trial (break-fixation trial), failed to make any response (miss trial), or made a saccade to the distractor probe (incorrect trial), that trial was aborted, and a new trial started. We decided the specific stimuli configuration for each trial as follows: First, we select the positions for the foreground and background shape(s), and subsequently, we randomly select the depth order of these shapes at the fixation point, which determined which shape would be fixated and thereby cued. The

depth ordering was, in all sessions, randomly selected on a trial-by-trial basis. For a given session, the positions of the two shapes, were either randomized or fixed (Note that even for sessions with fixed positions of the shapes, the depth ordering was still randomly selected on a trial-by-trial basis). Specifically, for monkey CH, we used fixed positions in sessions 1 to session 8 and introduced more positions starting from session 9. For monkey HU we used randomized positions in session 1 to session 6, session 10 and session 15 to session 18. This information is summarized in Figure 2.3 as the colored bars atop the plots (green, randomized positons; purple, fixed positions).

Experiment 2, Single-Probe Task

The single-probe task was used in Experiment 2 to test for an attentional benefit in behavior when oriented to the target shape. There were three conditions (Figure 2.2 B). The Valid- or Neutral-cue conditions were designed as in Posner paradigms. A third condition, the same as the multi-probe task, was included to ensure that the animal ignored the background shape. Trials in all conditions had the same temporal structure as described in the multi-probe task, except the probe-onset delay range was fixed to be 650∼3800 ms in monkey CH, and 150∼3800 ms in monkey HU. Each trial in the single-probe task was randomly assigned to be of either Valid-cue-condition, Neutral-cue-condition, or Multi-probe condition with the probability of 40%, 40%, and 20%, respectively. In all conditions, one foreground shape was presented together with either one, two, or three background shapes. The positions of the shapes were first randomly selected from four possible positions, corresponding to the four visual quadrants. For each shape, the color was randomly assigned. Finally, the depth ordering of shapes was determined randomly. In the Valid-cue condition, the target probe always appeared in the foreground shape. The exact location and time of probe appearance were randomly assigned for each trial, such that the monkey could not obtain reward above chance level without using the cue and waiting for the probe to appear. There were no distractor probes in the Valid or Neutral-cue condition. Trials in the Neutral-cue condition were the same as in the Valid-cue condition, except that the overlapping region of the shapes was masked out, rendering the shapes identical with regard to their attentional cueing state (Figure 2.2 B, middle). Probe contrast varied, as described above (*Visual stimuli and behavior control, the probe(s)*), for the Valid-cue and Neutral-cue condition, whereas probe contrast was always full for the Multi-probe condition.

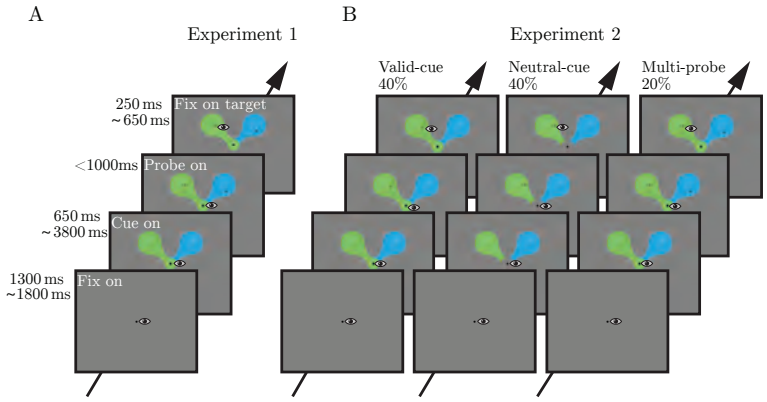


Figure 2.2: Behavioral tasks. (A) Experiment 1, multi-probe task. A typical trial starts with the monkey holding its gaze on a central fixation dot for 1300 to 1800 ms. Subsequently, the foreground and background shapes were presented. After a random delay (650∼3800 ms), probe stimuli of equal eccentricity appeared simultaneously on both the foreground and background shapes. The animal needed to make a saccade to the target probe on the foreground shape within 1000 ms, ignoring the distractor probe presented on the background shape. The animal would receive a juice reward after holding its gaze on the target probe for 250∼650 ms. Breaking fixation (break-fixation trial), absence of response (miss trial), or saccading to the distractor probe (incorrect trial) would abort the trial without reward. (B) Experiment 2, single-probe-task. There were three conditions. 40% of the trials were Valid-cue trials. In this condition, a single probe was presented on the foreground shape. 40% of the trials were Neutral-cue trials. Again, a single probe was presented on one shape, but the shapes were masked centrally such that there was no information available to tell which shape would contain the probe. The animal thus needed to monitor two shapes at the same time. The remaining 20% percent of the trials were the same as the trials in Experiment 1.

2.3.4 Data Analysis

Reaction Times

Reaction time was defined as the time between probe onset and saccade onset. Saccade onset was detected by a 2D velocity threshold as described in Engbert and Kliegl [223], using $\lambda = 10$.

Trial Exclusion

Trials during which the monkey diverted its gaze by > 1 dva from the fixation point were excluded from further analysis. In Experiment 1 (multi-probe task),

missed trials without a response either toward the target or distractor probe were also excluded. In both Experiment 1 and Experiment 2, trials with very fast reaction times (lower than 2.5th percentile) and with very long reaction times (longer than 97.5th percentile) were excluded. In Experiment 1 (multi-probe task), this procedure retained trials with reaction times in the range of 123∼329 ms for monkey CH. For monkey HU, the valid reaction-time range was 87∼363 ms. In Experiment 2 (single-probe task), this procedure yielded a reaction time range between 126 ms and 644 ms for monkey CH and between 115 ms and 706 ms for monkey HU.

Learning Trajectories

The learning trajectory was calculated as the moving-average performance, using a 100-trial boxcar. The boxcar was moved over the trials within a session, but did not move over session boundaries. If a session contained less than 100 valid trials, only one data point was calculated using all trials of that session. The confidence interval for each data point is the percentile confidence interval estimated based on 1000 bootstrap samples. The raw p-value for each session was calculated based on the one-tailed test for the following null hypothesis: the number of correct responses in the current session follows a binomial distribution with hit probability less or equal to the chance level. p-values were Bonferroni corrected for multiple comparison across sessions before the inference at $\alpha = 0.05$.

Psychometric Curves

We used *psignifit 4* [224] for psychometric function fitting in Experiment 2 (Single-probe task). In all fits, we used the Weibull function as the sigmoid. We estimated the threshold, width, lapse and eta from the data. We fixed the guess rate to be zero for the following reasons: (1) Chance level was low, because probe onset delays varied widely relative to the valid response window. (2) Trails with false-alarm saccades before probe presentation were terminated during the experiment; in the analysis, they were treated as break fixation and excluded. (3) Trials with unusually short reaction time (anticipatory saccades), which might correspond to false-alarm trials, were excluded. Together, these reasons reduced the expected false alarm rate to a value close to zero. Because the probes were presented on foreground shapes of different colors, we first fitted psychometric functions separately for the different foreground-shape colors.

For each individual fit, we used an automatic binning method for grouping the contrast levels. Specifically, for each fit, we started with a median split of stimulus contrasts. Then for each resulting half, we recursively split the contrast levels until the average performance in two halves were not significantly different (Fisher's exact test, at $\alpha = 0.05$). The psychometric function fit for individual colors (Figure 2.6 A and F) were then used to map the probe contrast (opacity) level x_i to a normalized stimulus level $\tilde{x}_i = \psi_{std}^{-1}(p_i)$; where ψ_{std} is the psychometric function with width and threshold equal to 1, and p_i is the hit rate at probe contrast x_i (Figure 2.6 B and G). These normalized stimulus levels were used to combine data with different foreground shape colors when comparing psychometric functions in Neutral-cue conditions and Valid-Cue conditions. We did the comparison separately for different numbers of background shapes. For each comparison, to test whether there was a difference in detection threshold between Neutral-cue trials and Valid-cue trials, we did a permutation test with 1000 randomizations (pooling Neutral-cue and Valid-cue trials) and reported the two-sided p-values.

2.4 Results

2.4.1 Cue-Learning with the Multi-Probe Task

Dowdall et al. [136] showed that macaque monkeys (including monkey CH and monkey HU) can rapidly learned the foveation-based selective attention task. This thesis chapter first, as background, presents and re-analyzes the learning trajectories of monkey CH and monkey HU. In addition to the already documented rapid learning in Dowdall et al. [136, 137], this thesis chapter describes that: (1) Learning can be facilitated by minimizing the possible stimulus positions (Figure 2.3); (2) Learning was faster when reaction times were longer (Figure 2.4).

We trained two macaques on the multi-probe spatial-attention task (Task 2 in Dowdall et al. [136, 137]) described above (Figure 2.2 A). Both animals were familiar with the non-cue-related aspects of the task (e.g., fixating on the fixation dot, presentation of the shape stimuli, making a saccade to a peripheral change, etc.), but they had no experience with spatial cues before the experiment. We recorded for both of them the behavioral data during the learning phase.

In the multi-probe task (Figure 2.2 A), the animal needed to pay attention to

the foreground shape, make a saccade to the target probe when it appeared, and ignore the simultaneously presented distractor probe on the background shape. The target and distractor probes differed only in their spatial locations (i.e., being positioned on the foreground and background shapes, respectively) and were identical in all other aspects. In all sessions recorded during the learning phase, both the target and distractor probes were salient so that there was no difficulty in detection *per se*. Thus, in those sessions, all errors were selection errors. The depth-ordering of the shapes varied randomly on a trial-by-trial basis, such that the animal needed to learn to use this cue to perform the task with a higher-than-chance-level performance.

In order to assess learning progress, we calculated the proportion of correct trials as a function of the number of overall completed trials (Figure 2.3). Monkey CH's performance was significantly higher than the chance level already in the fifth session ($H_0 : X \sim \text{Binomial}(n, 0.5)$, Bonferroni correction, $\alpha = 0.05$), approximately 300 trials after the spatial cue was first introduced. Monkey HU's performance was higher than the chance level starting from the seventh session, approximately 1500 trials after the cue was first introduced. Note that monkey HU's learning was slower, likely because in this monkey the shapes' positions were not fixed but selected randomly from four possible positions in the first six sessions. In session 7, we reduced the number of positions to two (as in monkey CH), and the performance quickly rose above chance level. Once the monkeys had learned the cue, their performance remained high, even when more positions were introduced later. For monkey HU, additional positions were introduced in session 15; for monkey CH, additional positions were introduced in session 9. This suggests that even though limited stimuli layouts were used initially, the monkeys were not trying to hard-code the relationship between the stimuli and target location. Instead, both monkeys learned the cue's abstract meaning, i.e., the fixated object is the target. At the later phase of learning, when the distractor/target probe was kept easy to detect, they rarely responded to the distractor, thereby satisfying the requirement of a filtering task.

Interestingly, we found that trials with longer reaction times (the time elapsed between probe onset and saccade onset) showed more rapid task learning. In Figure 2.4 A and B, we did a median split on reaction times for trials in each session and plotted the performance separately for the slow and fast halves—the curve rose above the chance level faster in the slow half for both monkeys.

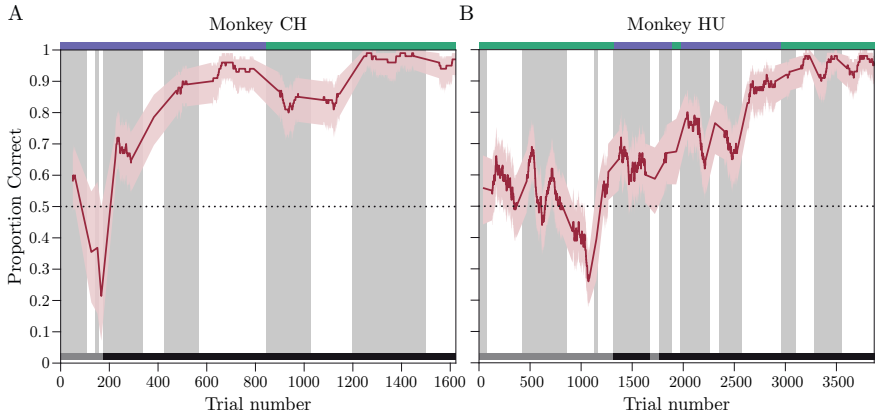


Figure 2.3: Learning trajectories. (A) Monkey CH's performance as a function of completed trials. The proportion correct and 95% bootstrap confidence interval were calculated using a 100-trial-wide boxcar window. The chance level is represented by the dotted line at 0.5. Alternating gray- and white-shaded regions in the background indicate session boundaries. Sessions with average performance significantly higher than the chance level are marked black in the bar at the bottom ($\alpha = 0.05$, Bonferroni corrected). The colored bar at the top shows for each session whether the shapes' positions were fixed (Purple: fixed; Green: random). (B) Same as (A) but for monkey HU. Learning trajectories for these two animals have been described previously in a different format in Dowdall et al. [136, 137]

Monkey CH's performance showed significant learning one session earlier in the slow as compared to the fast halves. Monkey HU's performance showed significant learning even four sessions earlier in the slow halves.

This effect might be explained by the relatively late arrival of the attention-related signal in the superior colliculus from the frontal/parietal cortex compared to the probe-onset related signal [225–227]. If this is the case, when all correct trials are combined, we should expect longer overall reaction times compared to incorrect trials. We did this analysis in Figure 2.4 C-H. We first separated the sessions into the pre-learning phase, peri-learning phase, and post-learning phase. Specifically, the pre-learning phase (Session 1 to session 4 for monkey CH; session 1 to session 6 for monkey HU) contained all sessions before the first session for which the detection rate was significantly above 0.5 (One-tailed test, $H_0 : X \sim \text{Binomial}(n, 0.5)$, Bonferroni corrected, $\alpha = 0.05$). Similarly, the post-learning phase (Session 6 to session 12 for monkey CH; Session 14 to session 18 for monkey HU) contained all sessions after the last session for

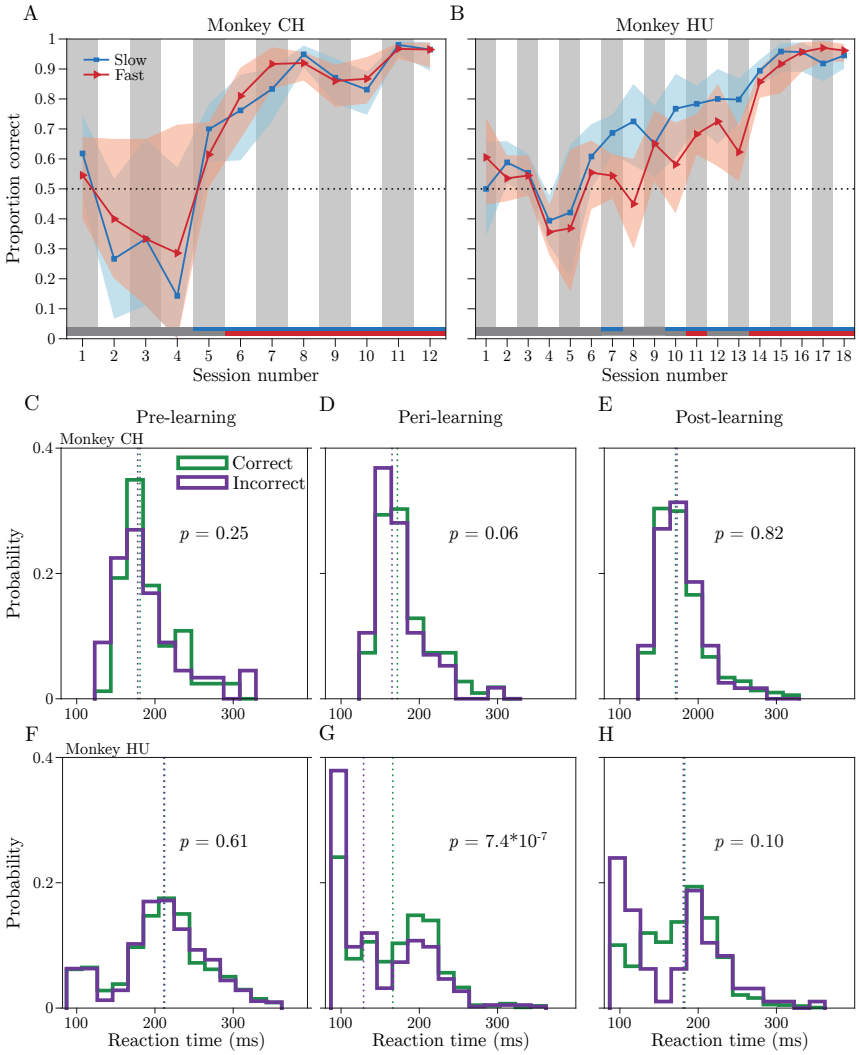


Figure 2.4: Trials with longer RTs showed more rapid task learning. (A) Monkey CH. Performance as a function of session number. Performance was calculated separately for the slow (blue) and fast trials (orange), based on the session-by-session median split on reaction times. Sessions with performance significantly higher than chance level were marked in the bar at the bottom, separately for the slow and fast half (same color code). (B) Same as (A), but for monkey HU. (C-E) Monkey CH. RT distributions for correct and incorrect trials. Wilcoxon rank sum test, H_0: equal median for correct and incorrect trials. (C) Pre-learning sessions. (D) Peri-learning sessions. (E) Post-learning sessions. (F-H) Same as (C-E), but for monkey HU. Results in the figure are based on further analysis of the data presented in Figure 2.3.

which the detection rate was significantly below 0.8 (One-tailed test, $H_0 : X \sim$ Binomial$(n, 0.8)$, Bonferroni corrected, $\alpha = 0.05$). The peri-learning phase (Session 5 for monkey CH; Session 7 to session 13 for monkey HU) contained all other sessions in between. As expected, in the peri-learning phase, reaction times for correct trials trended to be longer compared to reaction times for incorrect trials; this effect had a p-value of 0.06 for monkey CH and was highly significant in monkey HU. Note that monkey CH learned the task very quickly, which left only one session categorized as peri-learning session. Importantly, such a difference did not exist in the pre-learning phase, thus excluding the possibility that the difference is independent of the cue learning. The observed trend for a difference in RT distribution between correct and incorrect trials during the peri-learning phase suggests that for the attention signal to play a role (i.e., using the cue), the monkey needs to slow down, such that attention-related neuronal signals can contribute to the saccade initiation mechanism.

We considered an alternative explanation of the observed effects: Across trials, the delays between cue and probe onsets followed a uniform distribution, leading to an increasing conditional probability for probes to occur given that they had not yet occurred before, i.e., to an increasing hazard rate. An increasing hazard rate can lead to decreasing RTs [85]. This would associate long RTs with short delays between cue and probe and with the resulting reward, which might promote learning. To explore this alternative, we first investigated RTs as a function of cue-probe delays, and found them to show weak relationships with opposite directions in the two monkeys (Supplementary Figure 4-1). We then repeated the analyses of Figure 2.4 A and B, but rather than splitting trials according to RTs, we split them according to cue-probe delay (Supplementary Figure 4-2A, B). We also repeated the analyses of Figure 2.4 C-H, but rather than plotting RTs, we plotted cue-probe delays (Supplementary Figure 4-2C-H). These results differed markedly from the results in Figure 2.4 and did not support the alternative explanation.

2.4.2 Attentional Effect Observed in the Single-probe Task

To test whether the manipulation of attention using the shapes indeed increased the animals' behavioral performance, we further trained them on a single-probe task (Figure 2.2 B, Experiment 2). There were three types of trials in the single-probe task. 40% of trials implemented a Valid-cue condition. In these trials, the cue was 100% valid, but unlike the multi-probe task, only a single target probe

was presented during the Probe-on period. Another 40% of trials implemented a Neutral-cue condition. In these trials, the shapes were masked in the center, such that the depth order could not be resolved. The same single probe was presented during the Probe-on period in the Neutral-cue condition, such that the Neutral- and Valid-cue conditions were physically identical in the periphery. Because the masked shapes provided no information regarding where the target probe would be, the monkey was not able to benefit from selective attention, independent of its overall attentional strategy. As a result, we could assess the attention effect by comparing the monkey's performance between Valid-cue and Neutral-cue conditions. The remaining 20% of the trials were Multi-probe trials, as used above in Experiment 1. In these trials, the monkey needed to actively ignore the distractor probes that appeared on the background shape. Because stimulation was identical in the Valid-cue conditions and in the Multi-probe conditions until the probe was presented, a high hit rate in the multi-probe trials indicates that the monkey was attending or at least not responding to the background shape in the Valid cue condition. Both monkeys' performance in the Neutral-cue and Valid-cue conditions were around 70-80% for each session, and the range of probe contrast levels used were stable across sessions. In all sessions, for both monkeys, performance in the Multi-probe condition was above 90%. Note that the Multi-probe condition always used full-contrast probes, whereas the Neutral- and Valid-cue condition used variable-contrast probes as described in the Methods.

Both monkeys were faster in detecting probes at the attended site. Figure 2.5 A-D shows the RT distributions for monkey CH. When all stimulus configurations were combined for Neutral-and Valid-cue trials separately (Figure 2.5 D), the median RT for Valid-cue trials (221 ms) was significantly faster than the median reaction time for Neutral-cue conditions (235 ms) (Wilcoxon rank-sum test, $p = 7.3 \times 10^{-19}$). The same was true if we separately compared RTs between Valid-cue and Neutral-cue conditions for trials with 2, 3, or 4 shapes (Figure 2.5 A-C: 223 ms vs. 236 ms,$p = 9.5 \times 10^{-9}$; 214 ms vs. 238 ms, $p = 6.8 \times 10^{-12}$; and 221 ms vs. 231 ms, $p = 2.1 \times 10^{-4}$). The result was similar for monkey HU. The reaction time for Valid-cue trials was significantly faster compared to Neutral-cue trials in all comparisons (Figure 2.5 E-H; combined: 211 ms vs. 239 ms, $p = 3.5 \times 10^{-35}$; 2 shapes: 210 ms vs. 237 ms, $p = 2.7 \times 10^{-17}$; 3 shapes: 213 ms vs. 242 ms; $p = 2.0 \times 10^{-10}$; 4 shapes: 210 ms vs. 238 ms, $p = 3.4 \times 10^{-11}$).

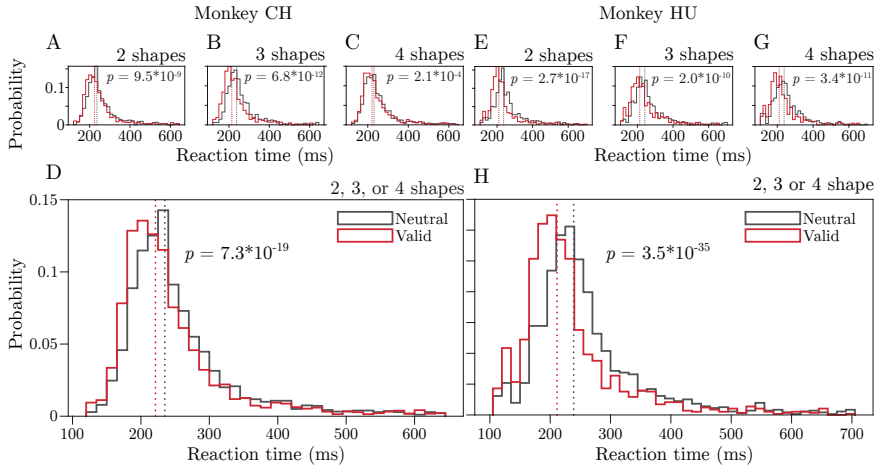


Figure 2.5: (A-C) Monkey CH. Reaction time distributions for trials with Neutral-cue (gray line) and Valid-cue (red line), separately for two (A), three (B) and four (C) shapes. Valid-cue trials had a faster reaction time (Wilcoxon rank-sum test, p-values listed as insets). (D) Same as (A-C), but pooled over conditions with different numbers of shapes. (E-H) Same as (A-D), but for monkey HU.

Detection of the subtle probes at the attended site was improved in one monkey and trended in the same direction in the other one, for the most attentionally demanding condition with four shapes. Because we presented probes of different contrast levels, we were able to fit psychometric functions. This was done separately for Valid-cue trials and Neutral-cue trials. By comparing the threshold of the two fits, we are able to assess the benefit of attention on the animals' contrast sensitivity. Specifically, for each condition, we fitted a psychometric function, separately for each color. We then used these fits to obtain the normalized stimulus level for each trial, independent of the foreground-shape color. We then fitted a single psychometric function using these normalized stimulus levels for all shape colors combined, yet separately for Valid- and Neutral-cue conditions. This procedure is demonstrated in Figure 2.6 A-B. We did this analysis separately for trials with 2, 3, or 4 shapes (Figure 2.6). With four shapes, monkey CH showed a significant decrease in the detection threshold (taken as the normalized stimulus level at 50% detection rate) (Permutation test, $p < 0.001$). The benefit of using the cue was not significant when there were two or three shapes (2-shapes $p = 0.13$; 3-shapes $p = 0.85$). The same trend was present for monkey HU; the significance test on the change in the detection threshold when there were four shapes resulted in a p-values of 0.15

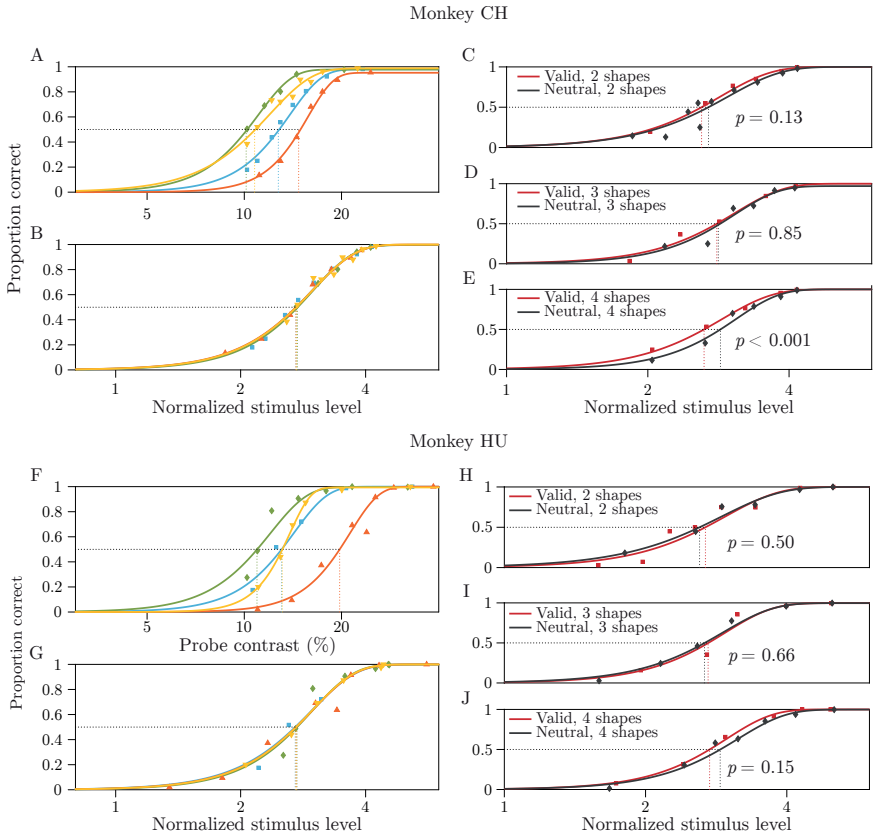


Figure 2.6: Attentional effect on contrast sensitivity. (A-E) Monkey CH. (A) Psychometric functions separately fitted for the different foreground shape colors; the data points and fitted functions are illustrated with the corresponding foreground shape colors. The fits obtained for the different colors were used to map probe contrast to a normalized stimulus level. (B) Psychometric functions, using the normalized stimulus level. (C) Psychometric functions for trials with 2 shapes using the normalized stimulus level as in (B), separately for the Neutral-cue condition (gray) and the Valid-cue condition (red). The threshold of the fits at 0.5 detection rate were compared between Valid-cue and Neutral-cue conditions through a Permutation test with 1000 randomizations, and the resulting p-value is listed as inset. (D) Same as (C), but for trials with 3 shapes. (E) Same as (C), but for trials with 4 shapes. (F-J) Same as (A-E), but for monkey HU.

(permutation test). Such a benefit was insignificant when there were only two or three shapes (2 shapes, $p = 0.50$; 3 shapes, p=0.66).

2.4.3 Modification of the Shapes

In the experiments described so far, we used solid-color shapes as the cue. However, using colors might interact with the probe stimuli and thereby introduce unnecessary variance. Such variance can be seen in Figure 2.6, where separate fits needed to be obtained for the different shape colors. In one monkey, monkey HU, we therefore aimed at removing the colors completely and using only contour-defined shapes (Figure 2.1 B, Figure 2.7). The left part of Figure 2.7 shows the monkey's performance as we gradually faded out the color. The removal of filling color, in particular the final step from weak to no color, led to a temporary decrease in performance. However, performance increased back to normal levels (above 90% correct) within 100 trials, and performance (calculated on a 100-trial running average) never dropped below chance level.

For many applications of the attention paradigm to neurophysiology, it is crucial to be able to present stimuli tailored to the recorded neurons' selectivity. For neurons in higher visual areas, this often requires the presentation of natural objects. Therefore, in the same monkey, on top of the contour-defined shapes, we subsequently introduced natural objects as target and distractor stimuli, and the right part of Figure 2.7 shows the resulting performance. This change of stimuli was less of a disturbance. We were able to fade-in the object completely over the course of a few dozens of trials.

2.5 Discussion

This thesis chapter describes the training of two macaque monkeys on a new type of attention task recently presented by Dowdall et al. [136], in which the cue was the foveated shape, and the cued location could be changed by changing the depth order of the shapes at the fixation point. The task was intuitive and easy to learn for the monkeys. Both monkeys learned the trial-by-trial cueing within few sessions. In a previous project in our lab, four other monkeys were trained on the same type of task, yet using natural stimuli instead of colored shapes as cues, and three of them learned the task within few sessions [134]. With the colored shapes employed here as cues, and maintaining trial-by-trial cueing, in a separate Posner-like task, both animals were able to use the cue to achieve a faster reaction time at the attended site, and a better detection rate was achieved by one monkey, with the same trend in the other monkey.

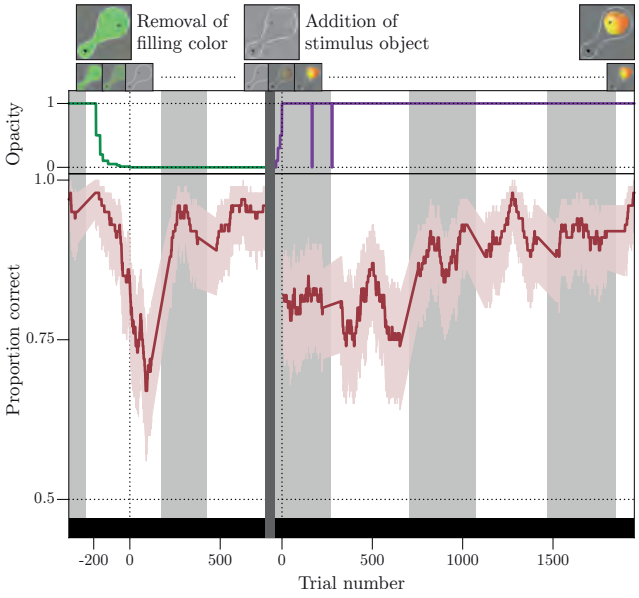


Figure 2.7: Removal of the shapes' filling color and addition of stimulus objects. Left bottom (red): Proportion correct around the time of color removal (trial number zero indicates the first trial with color completely removed). Left upper (green): opacity of shapes' filling color. Example shapes with different filling color opacity are illustrated above the opacity trace. Right bottom (red): Proportion correct around the time of addition of stimulus objects (trial number zero indicates the first trial with object at full contrast). Right upper (purple): opacity of stimulus objects. Example overlaying stimulus objects with different opacity are illustrated above the opacity trace.

Moreover, we have shown that the proposed cueing was very flexible: Once the animals learned the cueing through the shapes' depth ordering, changing the cue's exact form, spatial arrangement, or adding other stimuli to the display hardly impaired the performance. Even when, in one monkey (monkey HU), we removed the shapes' color completely, constituting a quite dramatic change in the display and a dramatic reduction in cueing information, this monkey was able to adapt quickly within a single session (Figure 2.7).

Macaques are intelligent animals. However, training them on a trial-by-trial cued selective attention task can easily take months [72, 165, 171, 184, 213, 216]. In our experience, this holds even when they already mastered general task aspects like fixation and the detection of peripheral changes, and the additional training pertains specifically to learning the cue and reaching high hit rates. In the introduction section, we proposed that long training times are likely due to the fact that classical tasks place additional demands beyond selectively attending to the target stimulus. Briefly, these demands include (1) understanding unintuitive cues; (2) learning the cue validity (in a Posner task); (3) withholding responses to the salient distractor onset (in a filtering task); and (4) holding the cue information in working memory. To eliminate these additional demands, the task employed in Dowdall et al. [136, 137] and Dowdall [134] and further developed here used a composite strategy: (1) the cue was natural: the fixated shape is automatically, i.e. naturally, attended; (2) the general task design was a filtering task, which kept cue validity at 100%; (3) target and distractor probes appeared simultaneously, avoiding an isolated salient distractor onset; (4) the cue was visible throughout the trial, avoiding any working memory demand; (5) the initial training used two shapes, whose depth ordering changed, but whose positions remained fixed, thereby eliminating non-essential task variability (in one monkey, early sessions with more positions gave poor learning, and later sessions with two positions gave better learning – see Results and Figure 2.3 B) [217]. We suggest that this combination was essential for the observation that both monkeys learned trial-by-trial cueing and responded almost exclusively to target probes after merely few sessions. Note that we do not claim (1) that the two studied macaques might not have learned an alternative selective-attention task equally quickly or even more quickly; (2) that the training time observed in these two macaques is representative for the macaque population; (3) that in the population, the training time for this task is shorter than for other tasks. Rather, we aim to share observations

obtained with this task design, which might motivate further labs to try the task with other animals. If learning is similarly quick, this will likely be helpful; if learning takes longer, it can still be terminated after a few weeks, and merely a negligible amount of time will have been lost, compared to the typical time for training macaques on attentional cueing.

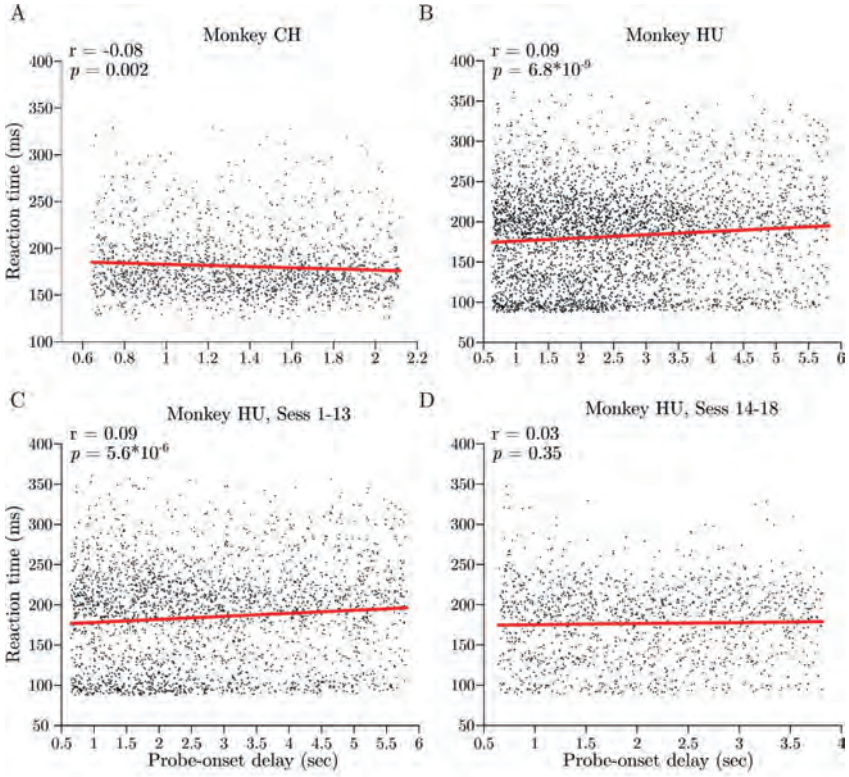
The ease of task learning is one important aim of the task design. Yet, another important aim is to minimize any factors that are confounded with the attention conditions, and therefore to keep the two attention conditions as physically identical as possible. Complete physical identity in both space and time is impossible, if the attentional conditions are to be cued on a trial-by-trial basis. Each trial needs to convey that trial's attentional condition through some cue, this cue can only be another physical stimulus, and thereby this cue stimulus is necessarily confounded with the attentional condition. Thus, the confound cannot be eliminated but merely be minimized. Different strategies have been used, sometimes in combination. Here is an incomplete list: (1) The cue can be placed outside the receptive field(s) of the investigated neuron(s) and thereby at different positions than the stimuli inducing the neuronal response under investigation; this separates the cue stimulus and the activity-inducing stimulus in space [190, 191, 203]; (2) The cue stimulus can precede the activity-inducing stimulus transiently, thereby separating them in time [176, 194, 206]; (3) The cue stimulus can differ between conditions merely in aspects that are not obviously related to the spatial or featural preferences of the recorded neurons and/or the activity-inducing stimulus; this is often used in studies with human participants and referred to as symbolic cueing, e.g. with letters or words. Approach (3) is by definition not building on natural, automatic attentional mechanisms, which makes learning more difficult and lengthy. Approach (2) requires the subject to hold the cue, after its disappearance, in working memory, which is demanding, particularly for animals [228]. Approach (1) has often been used with lines pointing from the fixation point to the target. The lines can get close to the RF of the recorded neurons and thereby lead to relevant confounds. Also, the lines might in fact hinder attention to reach from its natural location at the fovea to the cued location, i.e., it is not clear whether this cueing strategy capitalizes on natural, automatic attentional mechanisms. By contrast, the strategy employed here restricts physical differences to the immediate vicinity of the fixation center, leaving physical stimulation identical for the entire rest of the visual field. Furthermore, the strategy builds strongly on the natural,

automatic allocation of attention to the fovea, and the natural, automatic spread of attention from one part of an object to the entire object [218].

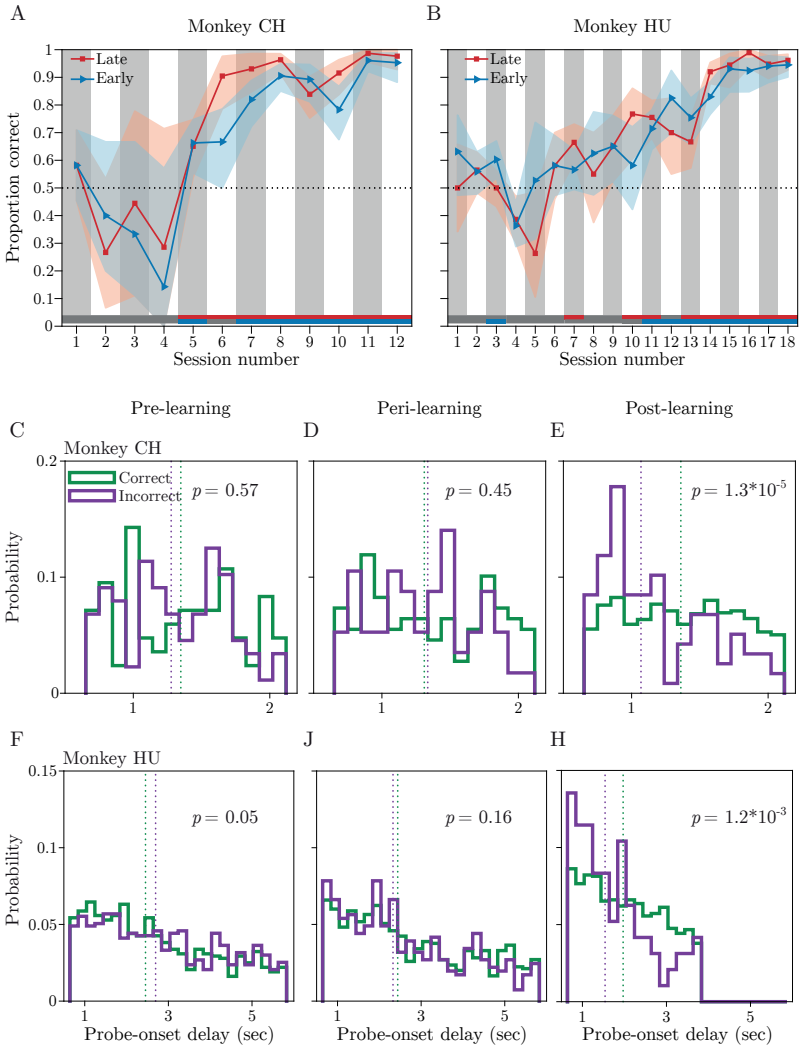
A large body of literature has investigated the effects of attention on perception and behavior. We replicated the two most commonly reported attentional effects on behavior: faster reaction times and, in one animal, lower detection thresholds. Reaction times shortened with attention by 14 ms for monkey CH and 28 ms for monkey HU and thereby showed an effect size comparable to previous studies [229, 230]. The detection threshold showed a small decrement with attention that reached significance in only one animal and trended in the same direction in the other; also the effect reached significance only in the most difficult condition with four shapes. The relatively weak evidence for an effect of attention on detection threshold might be due to the following reasons: (1) The detection of luminance difference might not be a capacity-limited process. The size of an attentional effect on sensitivity is thus expected to be small and may require a larger sample size to show. (2) Distractor shapes in Neutral-cue trials can be considered as independent external noise sources. More external noise sources can lead to higher decision criteria at all locations. Such higher criteria would also affect the target and thereby result in a lower hit rate. This might explain why we see detection-rate changes only in the 4-shapes condition [231]. (3) We used a reaction-time task in which a speeded, directed saccade was used to report probe detection. The outcome is a very natural stimulus-response pair for training the monkey. However, it is not optimal for the behavioral readout of attentional effects on sensitivity, compared to, for example, a 2-AFC task with limited stimulus presentation time, backward masking and post-stimulus probing. (4) Our experiment adopted a relatively long SOA, i.e., a long gap between cue onset and probe onset, which might reduce the attentional benefit [232, 233]. (5) In our task, shapes with different face colors were used such that the same probe had different local contrasts when placed on different foreground shapes. This created different signal strengths, and the animals might have applied different decision criteria for each combination of location and shape color. These two effects in isolation or in combination, might have decreased the power of our analysis. This latter problem can potentially be resolved with contour-only shapes without color. We trained one of the monkeys with such contour-only shapes, and he learned it within one session (Figure 2.1 B, Figure 2.7). (6) The employed contrast levels were manually selected and likely did not sample the psychometric function optimally with the limited number of

trials. The sampling can probably be improved by (a) including pilot sessions that estimate the overall psychometric function for a given animal, face color, etc., and (b) determining the optimal sampling schemes for this psychometric function by taking into account insights from previous systematic comparisons [234].

We found that the subset of trials with longer RTs showed more rapid cue learning (Figure 2.4). Because responses were given as saccades, any RT difference implied a difference in the saccade generation process. In this process, a very important structure is the superior colliculus (SC), because it sends direct output to the brain-stem circuitry responsible for motor output [235]. If SC output had a longer delay in long-RT trials, this was likely due to the respective SC neurons reaching their threshold later. Interestingly, in the context of the current task, the SC received two major inputs with different latencies: an early probe-onset related signal and a late attention-related endogenous input from frontal and parietal cortex [225–227]. The fact that we observed more rapid task learning in trials with longer RTs suggests that the learning is associated with reducing the relative weight of the probe-onset related signal and increasing the relative weight of the endogenous input in the decision process underlying saccade generation in the SC. The observation that trials with short RTs presumably contributed less to the learning gives hints about how to improve the task further. One possibility is to reduce the strength of exogenous input, e.g., reduce the contrast of the target/distractor probe during learning, or gradually fade in the probes to reduce onset transients. Another possibility is to give the monkey plenty of time to make the choice or even allow eye errors (e.g., break fixations and secondary saccades) during the learning phase.



Supplementary Figure 4-1: Reaction times as a function of probe-onset delay. (A) Reaction times as a function of probe-onset delay, for monkey CH. (B-D) Same as (A), but for monkey HU and with the following specifications: (B) All sessions combined. (C) Session 1 to session 13, with probe-onset delays in the range of 650∼5800 ms. (D) Session 14 to session 18, with probe-onset delays in the range of 650∼3800 ms. (E-L) Same analyses as Figure 4A-H, but replacing behavioral reaction times by probe-onset delays, as explained in the results section.



Supplementary Figure 4-2: Effect of probe-onset delay on learning. Same analyses as Figure 2.4 A-H, but replacing behavioral reaction times by probe-onset delays, as explained in Resutls.

CHAPTER 3

A Controlled Foraging Task for Studying Natural Vision[1]

3.1 Abstract

Vision research in the laboratory, particularly neurophysiological studies involving non-human primates, often adopts a passive approach using head-posted, eye-fixating animals to control the retinal input. This chapter presents a novel strategy to achieve controlled retinal stimulation through a carefully designed behavioral task. Central to this approach is a saccadic foraging task, in which the monkey is trained to fixate sequentially on potential targets to search for a hidden reward. By embedding symmetry in the search array, the task ensures that the pre-defined stimuli automatically fall within the receptive fields of recorded neurons at each fixation. This foraging task is easy for macaque monkeys to learn, adaptable to the desired visual stimulation and behavior, and thus well-suited for electrophysiological studies interested in natural vision in primates.

[1]Results reported in this chapter are based on data collected by Yufeng Zhang and Martina Pandinelli. Tim Näher contributed to the task design.

3.2 Introduction

Vision in our everyday lives is an active process. We exert control over what we see through our actions. As we navigate our environment, approach objects of interest, manipulate them, and shift our gaze to sample different parts of the visual scene, our actions determine what we see in the next moment, and conversely, what we see influences our actions in the subsequent moment. This dynamic interplay between looking and seeing is a fundamental and intrinsic aspect of natural vision [3, 93, 117, 236–239].

However, how vision is studied in the laboratory, especially in neurophysiological experiments involving non-human primates, often takes a passive approach [240, 241]. Animals are typically head-posted and instructed to fixate on a central fixation point while a peripheral stimulus is presented repeatedly. Crucially, unlike natural vision, there is no interaction between what the animal does and what the animal sees. This difference makes it difficult to transfer what we have learned about how vision works in the lab to the natural vision we experience every day in the real world [117, 238, 242].

Beyond the technical considerations for handling the animal and managing the recording apparatus, the choice of using head-posted, eye-fixating animals is rooted in the need to control the retinal input on a trial-by-trial basis. This control is particularly important for electrophysiological studies due to the presence of retinotopically organized neurons throughout the visual hierarchy. What follows is a common belief that only by controlling the input to these retinotopically organized neurons can we begin to understand their function [243–247].

It is important to note that achieving controlled retinal input does not necessarily require eye-fixating animals. In theory, as long as the experimenter knows the animal's current gaze direction, the stimuli can be positioned accordingly. Instructing the animal to maintain stable fixation only serves to simplify this process because the gaze direction is supposedly constant during fixation. Following this idea, various gaze-contingent techniques have been used to achieve retinotopically controlled visual stimulation while allowing the participant's eyes to move [248]. However, gaze-contingent techniques have their limitations. Firstly, there are noticeable delays of $\approx$20 ms for updating the stimulus position according to the participants' real-time gaze direction [249,

250]. For comparison, change detection is possible only 6 ms after the end of a saccade [251]. More importantly, gaze-contingent stimulation is unnatural. In our everyday experience, objects do not move with our eyes. In fact, if an object moves perfectly contingently to our gaze (e.g., the retinal blood vessels), we do not even see them [252].

In this chapter, I introduce an alternative approach to achieve controlled retinal stimulation while allowing animals to explore the visual scene naturally. This is done by embedding symmetry in the visual display, ensuring that the planned stimulus automatically falls within the receptive field (RF) of the recorded neuron even when the eyes jump between predefined locations. At its core, this method involves training the monkey on a saccadic foraging task wherein it sequentially fixates on potential targets to uncover a hidden reward [253]. I demonstrate that such a task is exceptionally easy for animals to learn, adaptable to the desired visual stimulation and behavior, and thus well-suited for electrophysiological studies interested in natural vision in primates.

3.3 Materials and Methods

3.3.1 Subjects

Two adult rhesus macaques (Macaque mulatta, denoted C and G in the main text, both 19 years old) were trained on the basic saccadic foraging task described below (Figure 3.1). All experimental procedures adhered to German and European animal protection laws. The experiment was approved by the responsible local authority (Regierungspräsidium Darmstadt). Both animals were implanted with a titanium head-post [219] and were trained on basic lab routines such as eye tracker calibration and passively fixating on a central fixation point. Monkey G had a recording chamber implanted to address other scientific questions. Water intake was regulated for Monkey C. Both animals received a juice reward after each correctly completed trial.

3.3.2 Visual Stimuli and Behavior Tasks

All visual stimuli were presented on an LCD monitor (LG 32GK850G-B) through custom software (https://github.com/esi-neuroscience/ARCADE). For both Monkey C and Monkey G, the viewing distance was 78 cm. A video-based eye tracker (EyeLink-1000, SR Research Ltd.) is used to track the real-time gaze

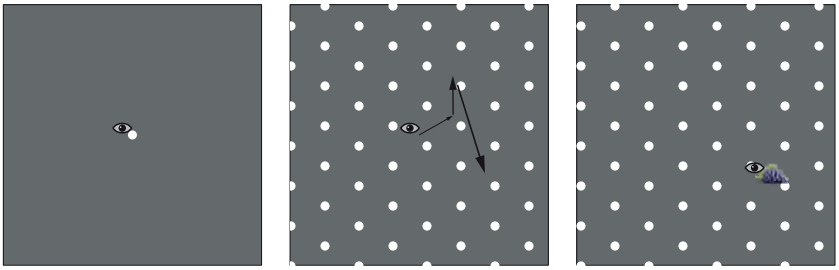


Figure 3.1: A basic foraging task. The task begins with a fixation period, where the animal focuses on a white circle at the screen's center for 800-1000 ms (Left). This is followed by the display of a search array consisting of uniformly spaced, identical white circles (Middle). Only 20% of the circles are true-targets associated with a reward, while the rest are pseudo-targets. The animal's task is to find and fixate on one of the true-targets by sequentially inspecting the circles ("foraging"). When a true-targets is located, it is replaced by an image of grapes, and a juice reward is immediately triggered (Right). By continuing to fixate on the the grapes, the animal can earn additional rewards, up to a limit of 2000 ms("harvesting").

directions. The signal from one of the two eyes was used for online experimental control. Before each session, a standard nine-point calibration/validation was applied. Typical gaze estimation errors were below 0.5 degrees of visual angle (dva).

In the subsequent sections, I first outline the structure of the basic foraging task. This task uses a search array that contains only one stimulus type and is thus called the fixed-target search task (Figure 3.1). Next, I describe in detail how to construct a fixed-target search array and how to build, using the same principle, more general search arrays that include distractors and neutral stimuli (Figure 3.2 and Figure 3.3). Example search arrays highlighting different features of the proposed foraging task are provided in Figure 3.4. Finally, I describe the training steps that led to two animals learning the basic foraging task in one session (Figure 3.5).

A Basic Foraging Task with Fixed Target Type and No Distractors

Figure 3.1 illustrates the structure of a basic visual foraging task with a fixed target type and no distractors. The task starts with the presentation of a filled white circle at the center of the screen, serving as the fixation point (diameter, 0.1 dva). After having acquired the animal's fixation for 800 to 1000 ms, the

fixation dot disappears. Following a ≈100 ms gap, the search array is displayed. This search array contains identical, white-filled dots (diameter, 0.3 dva) spaced regularly. All the dots are search targets, but only 20% of them are true-targets associated with a reward once visited. The rest are pseudo-targets. The animal's task is to sequentially look at these dots to locate any one of the true-targets (foraging). If, during foraging, the animal's gaze falls within a virtual target window surrounding a true-target (radius, 1.0 dva, not visible to the animal), an image of grapes will replace that target dot, and a small juice will be provided. Additionally, if the animal maintains stable fixation on the grapes (harvesting), a juice reward will be provided every 200 ms until a maximum harvesting window of 2000 ms elapses. If the maximum search time has elapsed (5000 ms) and none of the true-targets have been foveated, the trial will be terminated without a reward.

Search Array Construction

Constructing a search array starts with specifying a hexagonal grid layout (Figure 3.2). This layout partitions the 2D visual space into hexagonal tiles. A hexagonal grid layout can be completely specified by a seed hexagon (the shaded, central hexagon in Figure 3.2). The seed hexagon, in turn, can be specified by three parameters: its position, size, and angle. Figure 3.2 B and Figure 3.2 C illustrate the role of these parameters in controlling the grid layout by rotating and scaling the seed hexagon in Figure 3.2 A.

Every hexagonal tile is assigned a unique two-vector coordinate (q, r) in the hexagonal grid. The seed hexagon is assigned with a coordinate $(0, 0)$. In a flat-top hexagonal grid (as in Figure 3.2), the positive q direction $(+q)$ corresponds to moving to the right on the grid, and the positive r direction $(+r)$ corresponds to moving to the lower left. Depending on the parity of the q and r coordinates, each hexagonal tile is categorized into one of the four color groups. Tiles with even values for both q and r belong to the Red tiles, those with even q and odd r are Green tiles, those with odd q and even r are Blue tiles, and those with odd values for both q and r are Yellow tiles. In Figure 3.2, the tiles's color groups are indicated by the dot's color at the center of each tile. Note that the grid layout stays invariant if the entire grid is translated in a way that maps one tile to another tile of the same color group.

The above-described hexagonal grid layout, including its color group and coor-

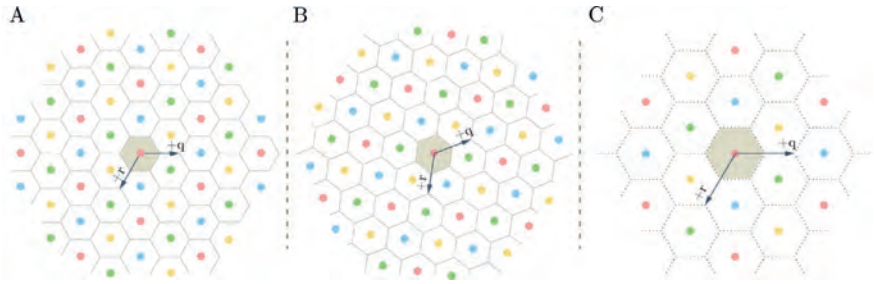


Figure 3.2: The hexagonal grid layout. (A) The hexagonal grid layout partitions the 2D visual space into hexagonal tiles. Parameters of the seed hexagon (the shaded hexagon at the center) set the full layout. Each tile is assigned a unique two-vector coordinate (q, r) and categorized into color groups based on these coordinates' parity, as shown by the dot color in each tile's center. (B) The layout in (A) rotated. (C) The layout in A scaled. The hexagonal grid layout including its color group and coordinates, remains invisible to the animals. They serve solely as a design guide for placing the actual visual stimuli. In practice, parameters of the hexagonal grid layout should be adjusted according to the RF of the recorded neuron(s).

dinates, remains invisible to the animals. They serve solely as a design guide for placing the actual visual stimulation. Placing these actual visual stimuli onto the hexagonal grid layout is hereafter referred to as filling.

Filling the grid layout begins with assigning the search target to at least one tile color (Figure 3.3 A). When mapping the search target to a color, all tiles belonging to the same color group will be assigned a copy of the search target to maintain the required symmetry in the search array. A search target is a stimulus that could potentially turn into a reward when fixated. The search target serves as a lure to guide the animal's attention. A true-target is a search target that is indeed rewarded when fixated, while a pseudo-target is a search target that does not lead to a reward. Since the animal lacks prior knowledge about which search targets are true-targets and which are pseudo-targets, it needs to check them sequentially much like it would do in the real-world during foraging. Ideally, with a well-behaved animal, these search targets, including both true- and pseudo-targets, would be the only positions to be visited (fixated) during foraging.

If the same search target is assigned to all four tile color groups, the resulting search array is a fixed-target search array used in the basic foraging task

(Figure 3.1). Alternatively, if not all four tile color groups are used, we have the option to assign distractors to the remaining tile color(s) (Figure 3.3 B). Distractors are characterized by 1) their lack of association with a reward, and 2) in separate trials, they can switch with the former search target to be the new search target. A foraging task with a search display containing distractors is hereafter called a "varying target foraging task" because the search target can vary across trials. The distractor stimulus may share some features with the search target but differ in other aspects. In an optimally behaving animal, the distractors would not be visited because they do not afford a reward.

Finally, we have the option to assign neutral stimuli to one or multiple remaining tile colors (Figure 3.3 C). These neutral stimuli are behaviorally irrelevant elements in the search display in all trials. Completing the assignment of these neutral stimuli marks the completion of the filling process. The filling process results in a trans-sacadically invariant display that contains objects of various types (Figure 3.3 D). Importantly, the association of object identity and its behavioral relevance can be altered on a trial-by-trial basis, which is essential for controlling the stimulus-specific biases.

Example Search Arrays

Figure 3.4 illustrates some example search arrays constructed following the procedures described above. An example scanpath is shown on top of each search array. Alone with the scanpath, the RF of a hypothetical neuron is plotted at each fixation to illustrate how this neuron might sample the search display during the animal's active visual foraging.

Figure 3.4 A-D illustrates four types of fixed-target search arrays. In Figure 3.4 A, the search target (green dot) is assigned to all four tile color groups and is the sole type of stimulus visible on the screen. Note that the grid layout is scaled and rotated in such a way that at each fixation, one and only one of the search targets falls within the hypothetical neuron's RF. Because the RF stimulus remains identical at each fixation, each fixation in the scanpath functions as a repetition unit, similar to a "trial" in a passive viewing task. A fixation epoch in a scanpath offers more than a passive "trial" because it is part of the animal's active exploration. For example, at each fixation, analysis of the neuronal responses can be aligned to either the start or the end of the fixation to explore the neural dynamics within a natural behavioral context [128, 143, 239,

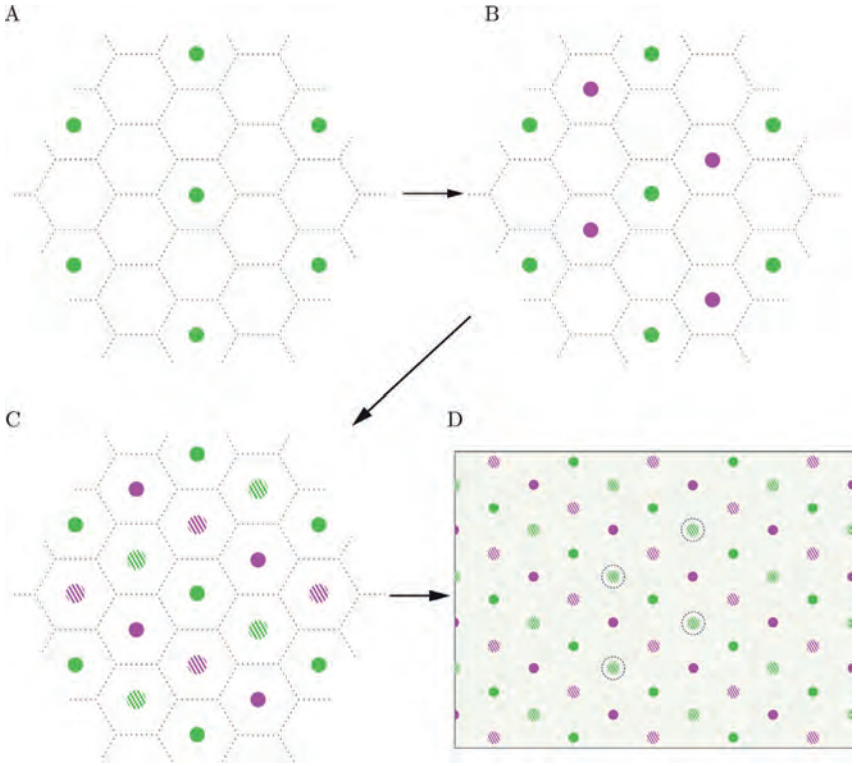


Figure 3.3: Filling a grid layout. (A) Placing the search targets. Search targets (the green circles) are assigned to the red tiles (see Figure 3.2 for tile color groups). (B) Placing the distractors. Distractors (the purple circles) are assigned to the yellow tiles. (C) Placing the neutral stimuli. The neutral purple and green gratings are mapped to the remaining blue and green tiles, respectively. (D) The resulting search display and one example scanpath during saccadic foraging. In this example display, the search targets are the green circles. The dotted line indicates how a hypothetical neuron's RF moves alone with the gaze shifts. Note that the layout filling procedure ensures the same stimulus (the green grating) falls in the RF at each fixation.

254, 255]. We can also ask how the planned saccade direction might influence the neuronal activity by comparing the fixations followed by a saccade toward the RF (as illustrated in the second saccade in Figure 3.4 A) with the fixations followed by a saccade away from the RF [256]. One apparent disadvantage of analyzing these fixation epochs is that, compared to a classic trial, the duration of each fixation is relatively short (200∼300 ms). The short analysis window may pose challenges for some data analysis methods. Potential strategies to encourage longer fixation durations are suggested in the Discussion.

With the search target assigned to all four tile color groups, the neuronal response is limited to that driven by the search target. This limitation can be easily addressed by adding neutral stimuli to the search array. Figure 3.4 B shows an example search array where the search target (green dot) is assigned to only one tile color, and the remaining tiles are assigned with an oriented grating as the neutral stimulus. Now, across fixations in a scanpath, this grating stimulus, instead of the saccade target, repeatedly falls within the hypothetical neuron's RF. As a result, unlike the search array in Figure 3.4 A, the RF stimulus is no longer behaviorally relevant and can be freely chosen independently of the primary foraging task.

Interestingly, performing a visual search on the search display depicted in Figure 3.4 B would result in a retinal stimulation sequence across fixations similar to the retinal stimulation sequence across trials when the same stimulus is presented repeatedly in the RF across trials while the participant is asked to fixate on a central target in each trial [61, 62, 257]. However, a key difference is that in our foraging task, the stimulus repetition is internally paced by the animal, while in the repetition task, it is externally controlled by the experimenter.

The search array shown in Figure 3.4 C adds further complexity to the fixed-target search display. In this case, the search target is assigned to two tile colors, and a neutral grating stimulus is assigned to one of the remaining colors. As a result, the search display entails two fixation types: one with the grating stimulus in the RF and another with no stimulation in the RF. The latter type occurs after a saccade that brings the empty tile into the RF (e.g., the first saccade in Figure 3.4 C). These fixations can act as a control condition with no visual stimulation but the same behavioral context.

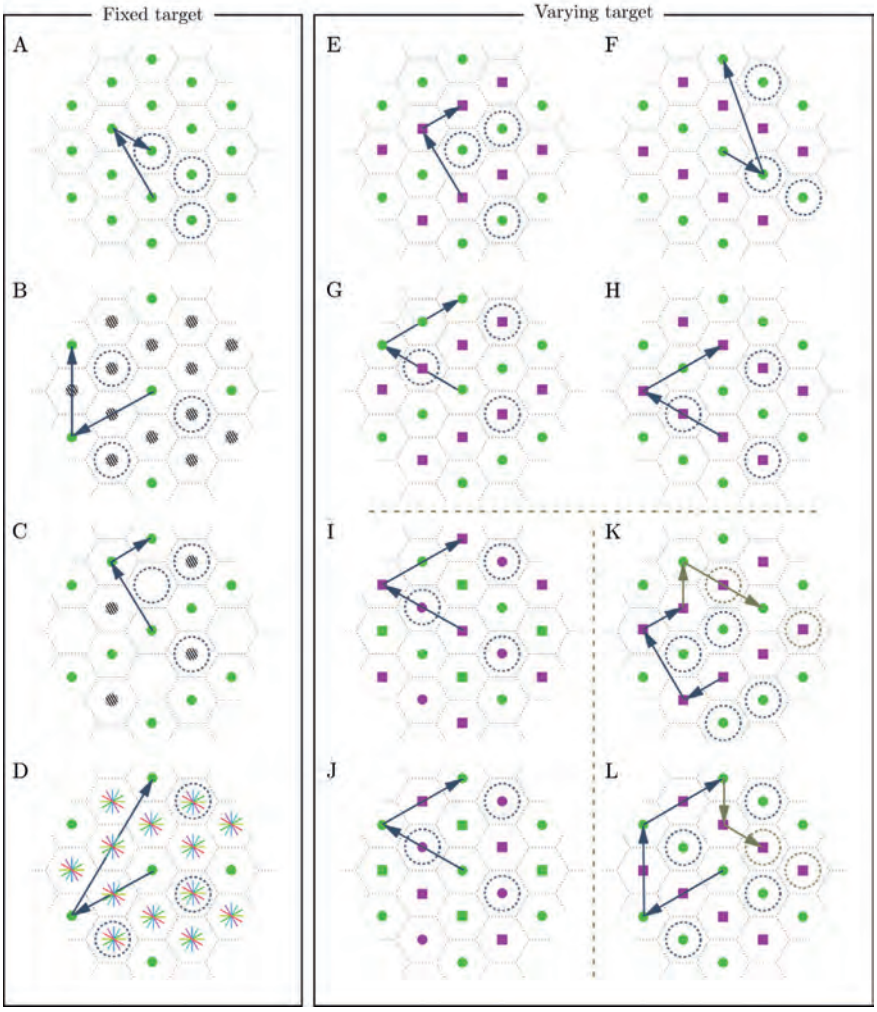


Figure 3.4: Example search arrays. In all panels, arrows illustrate the example scanpath, and the dotted circles illustrate a hypothetical neuron's RF along with each fixation.

The search array shown in Figure 3.4D contains neutral stimuli that vary randomly in orientation and color in different hexagonal tiles. This type of search display can be used to map the tuning curves of the neuron being recorded. Note that in this approach, the variable stimuli are pre-positioned onto different hexagonal tiles and brought into the neuron's RF through the animal's own action during foraging. This approach contrasts with the more typical mapping procedures, which rely on the stimuli's onset transients [258, 259]. As a result, the tuning curves obtained using this method would more accurately reflect the neuron's natural responses as the animal explores its environment.

The search arrays shown in Figure 3.4 E-L include distractors, which can alternate with the current search target to be the search target in a separate trial. Figure 3.4 E-H illustrates four conditions of a search task with one search target and one distractor. In each of these search arrays, the search target (the purple square in Figure 3.4 E and H; the green circle in Figure 3.4 F and G) is assigned to two of the four tile colors while the distractor (the green circle in Figure 3.4 E and H; the purple square in Figure 3.4 F and G) is assigned to the remaining tiles. Across these four conditions, the search target type is orthogonal to the stimulus type that enters the RF at each fixation. Specifically, when the search target is a purple square, the RF stimulus could be either a purple square (Figure 3.4 H) or a green circle (Figure 3.4 F). This design allows us to ask questions such as whether the firing rate would increase when the RF stimulus is the search target compared to when the same RF stimulus is the distractor. Note that when the RF stimulus is the search target, it is not necessarily to be selected for the next saccade target. In other words, the stimulus type in the RF is also (largely) independent of the direction of subsequent saccades. This decoupling enables us to investigate further questions, such as whether there is differential processing between the actually selected target and the physically identical but ignored target (e.g., in Figure 3.4 F, the RF stimulus is selected for the first saccade but ignored by the second saccade).

The search arrays shown in Figure 3.4 I and J are classical conjunction search arrays containing one search target and three distinct distractor types [88, 260]. In both search arrays, the stimulus in RF is a purple dot, but the search targets in the two search arrays are different. In Figure 3.4 I, the search target is a purple square sharing the same color as the RF stimulus. In Figure 3.4 J, the search target is a green circle sharing the same shape as the RF stimulus.

Figure 3.4 K and L illustrates a particularly interesting task design. This design includes two types of trials identical to those depicted in Figure 3.4 G and H, but with one key difference: in this case, the animal does not know in advance which of the two stimulus types is the search target and which is the distractor. The animal must learn this information through its actions. This task design can be implemented by randomly alternating between the purple target blocks (Figure 3.4 L) and the green target blocks (Figure 3.4 K) within a continuous session, with a variable trial length per block. When switching to a new block, the animal is expected to initially check the previous target type and quickly learn that the target has changed. For example, during the first trial(s) of a green target block (as shown in Figure 3.4 K), the animal would initially check the old target, the purple squares. After failing to receive rewards for a few fixations, it would start to check the green circles. In this process, as the animal explores, the animal's belief about the value of each stimulus is updated. Green dots gain value as the animal loses confidence in the purple squares. Figure 3.4 L illustrates the complementary scenario, where the green dots lose value as the animal explores, and vice versa for the purple squares.

Animal Training

In this section, I outline the key training steps that led to two animals achieving proficiency in the basic foraging task (Figure 3.1). Since the task parameters were adjusted on a trial-by-trial basis according to the animals' real-time performance, and because the animals learned the task remarkably quickly, I describe only the main steps. These steps may have varied slightly with each animal. Additionally, I focus on the steps after the animals have been trained on the basic laboratory routines, such as head-posting, eye calibration, and passive fixation.

The basic task in Stage 1 of the training is illustrated in Figure 3.5 A. The task starts with the animal fixating on a central dot. After the animal maintains fixation for 800∼1000 ms, the fixation dot is removed and replaced by a search target randomly placed on one tile of the invisible grid layout. This target is guaranteed to be the true-target and, when foveated by the animal, is replaced by the reward image. A fluid reward is provided as reinforcement immediately after the reward image appears. Since the search array contains only one stimulus, the animal would almost always look at it and receive the reward. At this stage, the sole search target's location shall be randomly selected across

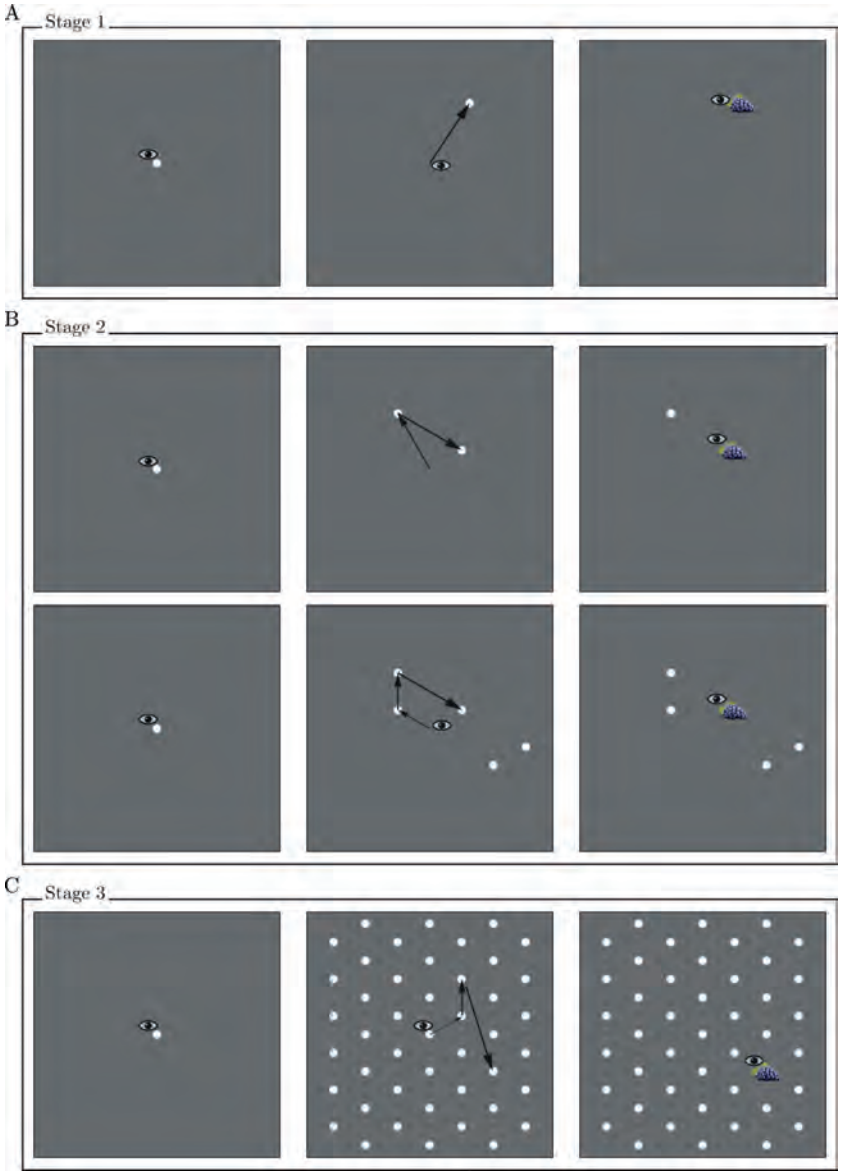


Figure 3.5: Training protocol. (A) Stage 1 of the training features a single true-target presented at a randomly selected tile within the grid layout. (B) Stage 2 of the training features the gradual addition of more search targets, including both the true-targets and pseudo-targets. (C) Stage 3 of the training uses the full search array, and the goal is to familiarize the animal with various task parameters such as different grid layouts, different true-target ratios, and stimulus timing.

trials to help the animal learn the association between the reward and looking at the search target rather than the association of the reward to a fixed spatial location.

In the next stage (Stage 2, Figure 3.5 B), additional target stimuli are gradually introduced to the search display, including both the pseudo-targets and true-targets. At this stage, the animal learns to explore rather than fixate on a single target. This is where the animal may experience some confusion. From our experience, two strategies have been helpful in maintaining the animal's motivation as more stimuli are added to the search array. First, the number of true-targets should be increased before introducing more pseudo-targets, such that minimal search is required early in the training. Secondly, it is helpful to keep some trials where a true target is positioned at the screen center. Since the animal starts the trial by fixating at the central fixation dot, these trials guarantee a reward. Even though these trials do not particularly encourage exploration, they are useful for keeping the animal motivated.

Through Stage 2, we will eventually have a full search array with all tiles of the grid layout filled with search targets, whether true-targets or pseudo-targets. The training then progresses to Stage 3 (Figure 3.5 C). At this stage, we only need to reduce the reward probability (the ratio between true-targets and pseudo-targets) to the target value and familiarize the animal with different grid layout and search target parameters, such as the layout's orientation and spacing and the search target's size and color.

3.4 Results

We trained two rhesus macaques on the basic foraging tasks (Figure 3.1). Before this training, both monkeys were familiar with standard laboratory procedures such as head-posting, passive fixation, and eye calibration but had no prior experience with visual search tasks or similar visual stimulation to what is depicted in Figure 3.1. We recorded their behavioral data during the learning phase.

The training followed the steps outlined in Figure 3.5. Throughout all training stages, the search display included at least one true-target and potentially multiple additional true-targets and pseudo-targets. Importantly, these true-

targets and pseudo-targets were physically identical and placed randomly on each trial. The monkey's task was to locate one of the true-targets by sequentially inspecting these target stimuli ("foraging"). We consider the monkeys to have learned the task if they are actively engaged in checking individual search targets.

I computed the observed hit rate to evaluate the learning progress. The observed hit rate is the proportion of trials in which one of the true-targets was successfully located. The observed hit rate was then compared with the no-search hit rate. The no-search hit rate is the expected hit rate if the animal were to employ a passive fixation strategy. The no-search hit rate is not zero because one of the true-targets could, by chance, be placed at the screen center, formerly occupied by the fixation dot. If this happened, the monkey would get an immediate reward at the search array onset. If the animal achieved a hit rate higher than the no-search hit rate, it is considered to have actively engaged in the visual foraging task. It is worth noting that the absolute hit rate could be influenced by task difficulty, which might vary across/within sessions according to the experimenter's specifications.

Both monkeys successfully learned the task within one session (Figure 3.6). For Monkey G, we started with a search display containing only a single true-target (Figure 3.6 A, Stage 1). This target varied its position at each trial. Given the large number of possible locations, the only stimulus was unlikely to be located at the screen center. As a result, we got close to zero no-search hit rate at Stage 1. The training quickly entered Stage 2 at the 28th trial. Monkey G's performance was maintained until the end of this session. We continued Stage 2 in Session 2 and the first half of Session 3. Throughout Stage 2, the number of true-targets and pseudo-targets were increased gradually. Monkey G entered Stage 3 at trial number 1022, with all tiles in the grid layout either filled with a true-target (20%) or a pseudo-target (80%). For monkey C, we skipped Stage 1 and started directly with a display full of true-targets. We then decreased the number of true-targets and at the same time increased the number of pseudo-targets. The number of true-targets was later increased again along with the pseudo-targets to the planned final value. It took 292 trials for Monkey C to enter Stage 3 (20% true-targets and 80% pseudo-targets) with the search array as shown in Figure 3.1.

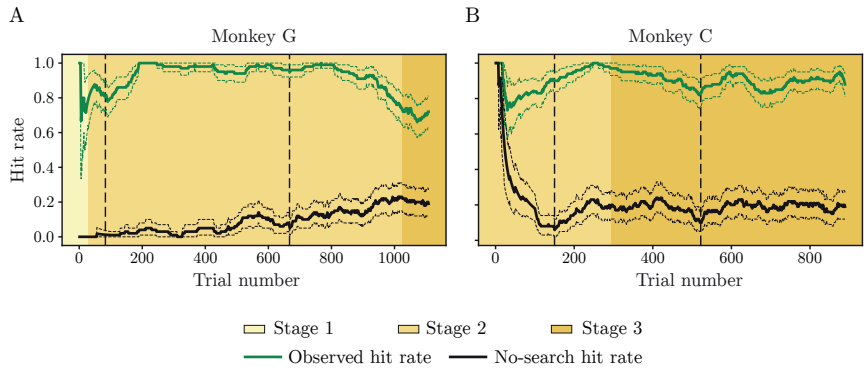


Figure 3.6: Learning trajectory. The observed hit rate (green) represents the percentage of trials in which one of the true-targets is located, calculated over a moving 100-trial window. The no-search hit rate (black) represents the percentage of trials where a true-target is at the screen center, ensuring a guaranteed correct trial without the need to search. The dashed line marks the 95% confidence interval derived from the binomial distribution of correct trials within the moving window. The background color reflects different training stages, as per Figure 3.5. The dashed vertical lines indicate session boundaries. (A) Monkey G. (B) Monkey C.

3.5 Discussion

Looking is a fundamental aspect of seeing, with direct effects on our visual perception [93]. These effects can be intentional, like bringing object from the periphery into the fovea; or intrinsic, like overt attentional orienting prior to saccade onset or incidental like the inter-saccadic motion streaks and visual transients accompanying each fixation onset [117]. However, in laboratory settings, especially in electrophysiological studies using non-human primates, the motor aspects of vision are often minimized. Much of the training for these animals focuses on fixation, teaching them to fixate steadily on a central point on the screen. This method, while simplifying the delivery of controlled retinal stimulation, overlooks the dynamic nature of natural vision. In this chapter, I introduced an alternative approach. This approach, by embedding symmetry in the search display, utilizes the animal's natural behavior, instead of limiting it, to provide controlled retinal stimulation. Because of its high ecological validity, I showed that the two macaque monkeys learned the basic foraging task within one single session. I also demonstrated how different search displays can be constructed to study various interesting questions in active vision (Figure 3.4).

I have posited that our search array construction method allows invariant retinal input across saccades. However, this holds true only under ideal conditions, such as using an infinitely large, homogenous screen and having each saccade landed precisely on the target center. These ideal conditions apparently cannot be met in practice. In the current study, the screen size was approximately $40 * 30$ dva. The symmetry of the search array breaks when the animal gazes toward the edges of the screen. Nonetheless, this is not a major concern as most saccades were made towards centrally placed search targets (because the animals are head-posted and naturally oriented forward) and the foveal magnified visual processing can mitigate the asymmetry at the periphery. Using a larger screen could further mitigate the edge effects. In human studies, using a virtual reality setup can enable a truly unbounded search array.

The issue of saccade landing errors is indeed noteworthy. Both monkeys displayed sub-optimal search performance, primarily due to many saccades missing the intended 1dva radius fixation window. This lower precision may stem from the animals' strategy to prioritize faster sampling over accuracy, a behavior we encouraged because we want to avoid procrastination for the purpose of using these animals in a separate study. Previous research indicates that macaques can achieve high precision in search tasks and visually guided saccades [261]. Increase the reward for precise saccades could potentially reduce the landing position variance. Furthermore, since landing errors vary across saccades and trials, their impact on the main research question can be systematically evaluated.

While introducing the task, I suggested that each fixation during the visual foraging can be treated as a single trial in a traditional passive fixation task. This analogy makes sense because each fixation is a repetition unit of consistent visual input just as a trial being a repetition unit in the passive fixation task However, there is a notable difference: whereas a typical trial lasts a few seconds, each fixation usually lasts just a few hundreds of milliseconds. This discrepancy poses challenges for data analysis pipelines that rely on the system maintaining a steady-state. While we could design a search task that requires longer fixation times, e.g. by using search targets that require further foveal processing, the short fixation durations are characteristic of natural vision [111, 115, 262]. Therefore, I consider these brief fixations a feature of our study, necessitating adaptations in the analysis methods rather than alterations to the task design.

On the surface, our foraging task closely resembles a conventional visual search task, where the subject identifies a specific target's presence or absence [88, 95]. However, two key differences distinguish our approach from the visual search task. Firstly, in the proposed task, the true-target is probabilistically defined and physically indistinguishable from the pseudo-targets. Having the true-targets to be probabilistically defined not only allows straightforward comparisons between conditions but also opened the door to study the probabilistic computation in the brain (e.g. Figure 3.4 K and L). Secondly, we employ a hexagonal grid layout for the search display. The hexagonal grid layout optimizes the likelihood that pre-defined stimuli fall in the recorded neurons' receptive fields. While other types of regular tessellations could serve our purpose, the hexagonal grid is chosen for its simplicity, and some preferred features such as equidistant neighboring tiles, that can simplify the subsequent data analysis.

The diverse visual-motor behaviors observed while the animals were performing the proposed saccadic foraging task invite us to ask numerous intriguing questions. For instance, are the revisited targets processed differently from novel ones, and if so, how are they different from each other [263]? What are the underlying search strategies, and how might they change if cues indicating the local probability of true-targets are introduced [264]? Can animals integrate these cues, and if so, how this integration is reflected in the neural activity and behavior [265]? These questions, rooted in ecological validity, mark our goal for developing the saccadic foraging task: understanding how vision works under naturalistic conditions as in our everyday lives.

CHAPTER 4

Eccentricity-Dependent Saccadic Reaction Time: The Roles of Foveal Magnification and Attentional Orienting[1]

4.1 Abstract

A hallmark of primate vision is the emphasis on foveal processing, accompanied by frequent saccades that bring the fovea to salient parts of the scene, or to newly appearing stimuli. A saccade to a new stimulus is one of the most fundamental sensory-motor transformations. In macaque monkeys, we show that foveal magnification is not only the reason for saccades, but it also affects the dynamics of saccade initiation. In a task where the monkeys made saccades to peripheral target onsets, saccadic reaction time (SRT) increased with target eccentricity. Notably, we effectively eliminated this increase by scaling the target size according to the foveal magnification factor in the superior colliculus. We

[1]An earlier version of this chapter is available on bioRxiv: Yufeng Zhang and Pascal Fries. *Eccentricity-Dependent Saccadic Reaction Time: The Roles of Foveal Magnification and Attentional Orienting*. en. Aug. 2023. DOI: 10.1101/2023.08.08.552339. URL: http://biorxiv.org/lookup/doi/10.1101/2023.08.08.552339 (visited on 07/03/2024)

repeated the comparison between non-scaled and scaled targets, while changing the task to a delayed saccade task. In this task, the target was presented long before the saccade, and the saccade was triggered by foveal fixation offset rather than target onset, such that target onset long before fixation offset was essentially irrelevant for SRT. In this task, we found that SRT increased with target eccentricity, with similar rate for both non-scaled and scaled targets. Furthermore, this increase survived the addition of a salient distracting flash resetting attention to the foveal. The results obtained with the delayed saccades task are consistent with an attentional scan from the fovea to the target, a recently hypothesized general mechanism of attention.

4.2 Introduction

We make several saccades per second, redirecting our gaze to bring objects of interest into our high-resolution fovea. However, the time it takes to initiate a saccade to a novel target, i.e., the saccadic reaction time (SRT) can vary significantly, even when the task is simple and highly repetitive. Part of the variability comes from the size of the impending saccade. The relationship between SRT and the saccade size has a characteristic bowl shape with long SRTs for the shortest (<1 degree of visual angel, dva) and longest (>10 dva) saccades, and shorter SRTs for medium-sized saccades [267–271]. Within these medium-sized saccades the relationship between SRT and target eccentricity for medium-sized saccades remains unclear and seems task dependent [268, 269, 272–275]. Yet, these medium-sized saccades between 2 and 10 dva constitute the most commonly executed saccades in daily life [144, 150] and in laboratory settings, where SRTs are widely used as a tool to characterize the cognitive process of interest [230, 276, 277].

Previous studies describing the relationships between SRT and saccade size often utilized a Step task. In a Step task, the fixation dot steps into the periphery and becomes the saccade target. This design usually comes with a significant confounding factor: saccade targets that are physically identical at various positions vary in visibility with corresponding saccade sizes [278]. This is partially due to the fact that a visual stimulus of a given physical size undergoes foveal magnification early in the visual processing pathways [140, 245, 279] and, therefore, has a decreasing drive for increasing eccentricity or saccade size, respectively [280]. At the single-neuron level, this decreasing drive

is correlated with the stimulus occupying a smaller proportion of the receptive field (RF), the size of which increases as the RF moves away from the fovea [140, 245, 281]. Further, accompanied by foveal magnification and the corresponding change in RF sizes, neurons representing the visual periphery prefer lower spatial frequency components as compared to neurons representing the fovea [282]. Given that a dot is a broad-spectrum stimulus, its spectral components will be amplified differentially at different eccentricities. In addition, low background lighting commonly employed in previous studies might have resulted in the target onset signal being too strong, such that the SRT showed a floor effect, which might have masked the effect of target eccentricity [283]. It is, therefore, important to control the strength of the afferent signal at different eccentricities to draw conclusions about the dependence of SRT on target eccentricity *per se*.

Besides physical stimulation, the nature of the impending saccade, as determined by the specific task structure, can also influence the relationship between SRT and target eccentricity. Recently, Hafed and Goffart [268] reported a clear effect of increasing SRT with target eccentricity in a visually guided saccade task. Despite the confounding factors mentioned earlier, when the same stimulus set was used for a delayed visually guided saccade task, the observed increase of SRT with eccentricity became much less prominent. This is interesting because the main difference between these two tasks is that in the visually guided saccade task, the response saccade was exogenously driven, dominated by the transient onset of the target stimulus; in contrast, in the delayed visually guided saccade, the response saccade was endogenously driven, triggered by fixation dot offset, far away from the target stimulus. The observation that the SRT in these two tasks exhibited different patterns of eccentricity dependence motivated us to investigate these two types of saccades (endogenous saccades and exogenous saccades) separately.

Given the same physical conditions and task structure, SRTs across trials can still exhibit high variability [284, 285]. The participant can very commonly make a response saccade with SRT as fast as 150 ms in one trial and as slow as 300 ms in another trial. Such high variability requires many trials to reveal a genuine but modest shift of the SRT distribution. For example, more than 200 trials per condition are needed to reveal a shift of 10 ms for a typical SRT distribution (calculated with an ex-Gaussian peaked at 200 ms with a standard deviation of 38 ms, two sided Mann-Whitney U test). Such trial counts per

condition were generally not obtained in previous studies.

The current study aims at addressing the aforementioned issues in macaque monkeys, because they offer a rich history of studies of oculomotor control and attention, and also the potential to use the current results to motivate future circuit investigations. We selected two sets of stimuli as saccade targets located between 2 and 10 dva (Figure 4.1 A). One set is Equal, having the same physical size at different eccentricities. The second set is Scaled based on the foveal magnification in the superior colliculus (SC), aiming to equalize the targets' corresponding afferent input strength to the SC across eccentricities [245]. To avoid a flooring effect due to high input strength, in all tasks, we presented the target on a gray background. We used the same stimuli sets in both, a Step (visually guided saccade) task where the response saccade is exogenously driven by the target onset, and a Delayed (delayed visually guided saccade) task where the response saccade is endogenously triggered by the fixation dot offset (Figure 4.1 B-C). Because the scaling is supposed to mainly affect target onset transient, we expect to see the scaling reduce SRT significantly in the Step task but affect SRT minimally in the Delayed task. Additionally, in the Delayed task, we incorporated an extra condition including a transient attention-capturing distractor flash to investigate whether the eccentricity-dependent SRT changes for the endogenously driven saccades are consistent with an attentional scan from the fovea. Last but not least, for each monkey and each condition, we collected more than 200 trials per condition to gain the necessary statistical power.

4.3 Results

4.3.1 Foveal Magnification Explains Eccentricity-Dependent SRT Increase for Exogenously Driven Saccades

We start with exogenously driven saccades in the Step task (Figure 4.1 B). The hypothesis is that, for exogenously driven saccades, the reported SRT increase with eccentricity [268] is due to foveal magnification. Specifically, accompanying foveal magnification is the increasing RF size with higher eccentricity in the SC to equalize the size of the active population for the same targets located at different eccentricities [140, 245]. Increasing RF size, however, decreases the visibility of the stimulus because it occupies a smaller portion of the (excitatory)

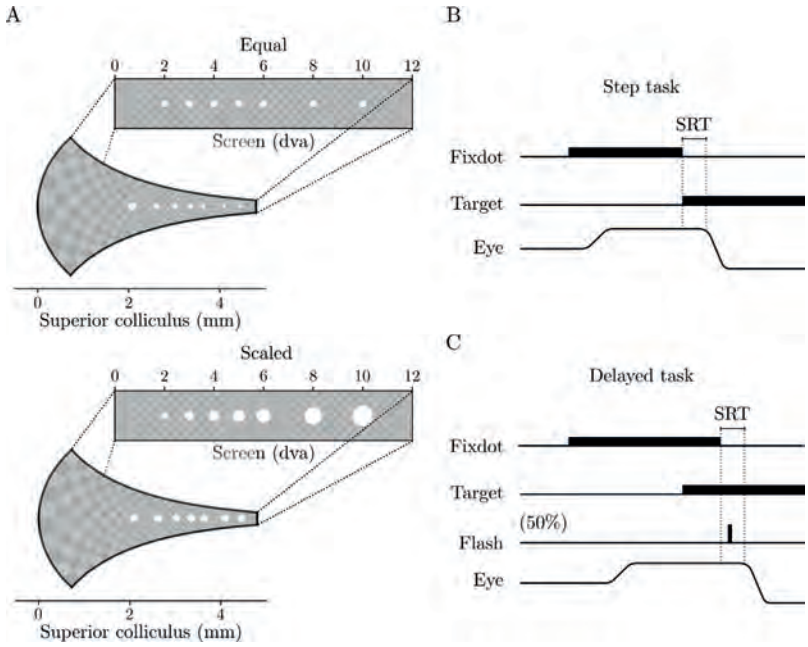


Figure 4.1: Stimuli and task. (A) Representation of target stimuli on the screen and their corresponding afferent input in the SC map, based on the foveal magnification reported in Chen et al. [245]. Top: Equal targets at varying eccentricities on the screen exhibit a diminishing size of corresponding afferent input in the SC. Bottom: Scaled targets on the screen have the same size of corresponding afferent input in the SC map. Note that the checkerboard background is shown for illustration purposes and was not presented during the experiment. (B) In the Step task, the fixation dot stepped into the periphery and became the target stimulus. (C) In the Delayed task, the target presentation preceded the fixation dot removal. The animal was required to wait until the removal of the fixation dot before shifting its gaze. In 50% of trials, a transient foveal flash was presented at ≈100 ms after fixation dot removal. The flash had no task relevance and served to capture attention back to the fovea while the response saccade was planned but not yet executed.

RF. This hypothesis also aligns with the recently reported observation that SC neurons at higher eccentricities prefer lower spatial frequencies (corresponds to a larger dot) [282].

To test this hypothesis, we designed two stimulus sets as saccade targets. The first set, Equal, was similar to what has been used in previous studies (Figure 4.1 A, top) [268, 272]. In this set, saccade targets at all positions were white filled circles of diameter 0.1 dva. The second set, Scaled, consisted of filled white circles whose sizes were scaled according to the reported magnification factor in the SC, such that they provided the same afferent input measured in the SC map (Figure 4.1 A, bottom). Both sets consisted of stimuli positioned at 2, 3, 4, 5, 6, 8, and 10 dva to the right of the fixation point along the horizontal median. To facilitate comparison, the target stimulus at 2 dva was of the same size in both sets.

Figure 4.1 B illustrates the task structure of the Step task. In this task, each trial started with the monkey maintaining fixation on a fixation dot (0.1 dva filled white circle) for 800 to 1500 ms. Subsequently, the fixation dot was turned off. At the same video frame, a single saccade target randomly chosen from the combined Equal and Scaled stimuli set was presented. The monkey needed to make a saccade within 500 ms towards the target and hold its gaze on the target for another 800 to 1500 ms. If it completed the trial correctly, a juice reward was provided. Only correct trials were included in the analysis.

The resulting SRT distributions in each condition from one example monkey (HO) are plotted in Figure 4.2 A. The raw data shows that for this monkey: 1) Both location and shape of the SRT distribution varied across conditions. 2) Similar to what has been reported before, the peak of the SRT distributions for Equal targets (blue) gradually increased with increasing target eccentricity. 3) This shift was significantly reduced for Scaled targets (orange). 4) SRTs also became more variable as they became longer. We next quantified these observations.

To quantify these observations, we first fitted an exponentially modified Gaussian (ex-Gaussian, solid lines in Figure 4.2 A) for each SRT distribution. The ex-Gaussian is the most commonly used parametric function to describe a reaction time distribution [115, 286]. We used the mode of each resulting ex-Gaussian fit to represent the typical SRT, as it denotes the most likely occurring value in the

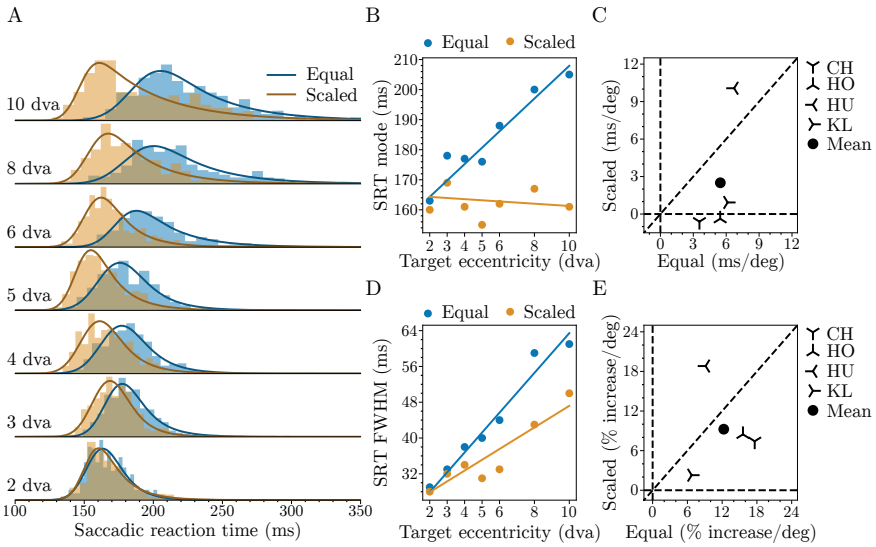


Figure 4.2: Eccentricity-dependent SRT in the Step task. (A) SRT distributions from monkey HO for Equal (Blue) and Scaled (Orange) targets. Solid lines represent the fitted ex-Gaussian distributions. (B) Multiple linear regression of SRT mode on target eccentricity and scaling for monkey HO. (C) Scatter plot of SRT mode-eccentricity slopes for Scaled vs. Equal targets in all tested monkeys. For each monkey, the slopes are derived from the regression coefficient, as demonstrated in (B). The filled circle shows the mean slopes. (D) Same as (B) but for FWHM. (E) Same as (C) but for FWHM. The raw slopes were converted to percent increase per degree of visual angle.

distribution. Additionally, we used the full width at half maximum (FWHM) to represent the SRT variability. Having these values, we first examined how the typical SRT varied with saccade size, separately for Equal and Scaled targets. To this end, we fitted a simple linear regression model ($Mode \sim \beta_0 + \beta_1 * (targetEccentricity - 2)$). For monkey HO, SRT mode increased with target eccentricity for the Equal targets but not for the Scaled targets (Equal: $CI_{\beta_1,95\%} = [4.33, 5.73]$, $p < .001$; Scaled: $CI_{\beta_1,95\%} = [-0.55, 0.71]$, $p = .82$). Next, to confirm that scaling reduced the slope of SRT as a function of target eccentricity, we fitted a multiple linear regression model ($Mode \sim \beta_0 + \beta_1 * (targetEccentricity - 2) + \beta_2 * (targetEccentricity - 2) * isScaled$) on the combined data from Equal and Scaled targets (Figure 4.2 B). Scaling indeed reduced the slope (Combined: $CI_{\beta_2,95\%} = [-6.45, -5.25]$, $p < .001$). We repeated the same procedure for the other three monkeys. The resulting slopes for each monkey in both conditions are plotted in Figure 4.2 C. Notice that data from three out of four monkeys (CH, HU, and KL) lie very close to the horizontal zero line meaning that, for these monkeys scaling the target stimulus according to the SC foveal magnification factor effectively eliminated the eccentricity-dependent SRT increase. Having only four monkeys, we are essentially limited to the fixed effect analysis and draw inferences on these four monkeys instead of on the population [287]. Fixed effect analysis shows that for these four monkeys, scaling significantly reduced the average slope($CI_{\overline{\beta_2},95\%} = [-3.42, -2.59]$, $p < .001$). We repeated the same analysis for the variability of SRT (ex-Gaussian FWHM). Similarly, for monkey HO, FWHM increased with target eccentricity for Equal targets ($CI_{\beta_1,95\%} = [3.43, 5.11]$, $p < .001$). Scaling reduced the speed of increase for FWHM but this reduction was not as effective as for the mode (Figure 4.2 D, Scaled: $CI_{\beta_1,95\%} = [1.84, 3.24]$, $p < .001$; Combined: $CI_{\beta_2,95\%} = [-2.80, -1.37]$, $p < .001$). Fixed effect analysis using averaged percent change per degree on FWHM agreed with the pattern observed in monkey HO (Figure 4.2 E, $CI_{\overline{\beta_2},95\%} = [-4.41, -1.95]$, $p = .001$).

4.3.2 Scaling Reduces SRT Most Effectively for Low-Contrast Targets

The results in the previous section suggest that the observed SRT increase with increasing target eccentricity could be explained by foveal magnification. This is prethe RF sumably because, accompanied by foveal magnification, RF size increases as the RF moves away from the fovea [140, 245]. Consequently, a

physically identical stimulus closer to the fovea would drive the corresponding neuron stronger in the SC because it occupies a larger proportion of the RF. This stronger stimulus drive would then result in a faster saccade initiation [282]. Also, a stimulus with higher contrast evokes a stronger and faster response throughout the visual hierarchy, leading to a faster reaction time [288–293]. Furthermore, these two bottom-up factors, stimulus size and stimulus contrast, have been shown to interact with each other in determining the magnitude of response transient in the SC [282]. We next investigated whether such interaction between stimulus size and contrast would be reflected in the eccentricity dependent SRT increase.

To this end, we slightly modified the Step task that we had used so far. Firstly, we reduced the screen background luminance to allow for a wider range of target contrast manipulation. Secondly, at each target location and for each target size, we selected three target luminance levels such that the resulting contrasts were approximately evenly spaced on a logarithmic scale (Weber contrast equals 1.4, 3.1 and 6.9 for Low, Medium and High contrast respectively). Lastly, we limited target locations to 2, 4, and 6 dva to ensure an adequate number of trials for each condition. We collected data from one monkey (HO) for this modified Step task.

We plotted the resulting SRT distributions in Figure 4.3 A-C. At each contrast level, SRT tended to be longer for more eccentric Equal targets. Scaling reduced this increase, and the effect of scaling, in terms of reducing the eccentricity-related SRT increase, appeared to decrease with increasing contrast. As before, we used linear regression to quantify these observations (Figure 4.3 D-K).

To evaluate how SRT mode depends on target eccentricity, we first fitted a simple linear regression model (Mode $\sim \beta_0 + \beta_1 * (targetEccentricity - 2)$) at each contrast level, separately for the Scaled and Equal targets. For the Equal targets, the SRT mode increased with target eccentricity at each contrast level. For the Scaled targets, the SRT mode decreased slightly with eccentricity at low contrast and increased with eccentricity at higher contrast. At each contrast level, the slopes are lower for the Scaled compared to the Equal targets. This difference in SRT-Eccentricity slopes between Scaled and Equal targets indexes the effect of scaling. To quantify the effect of scaling reducing the slope, we then fitted a multiple linear regression model (Mode $\sim \beta_0 + \beta_1 *$

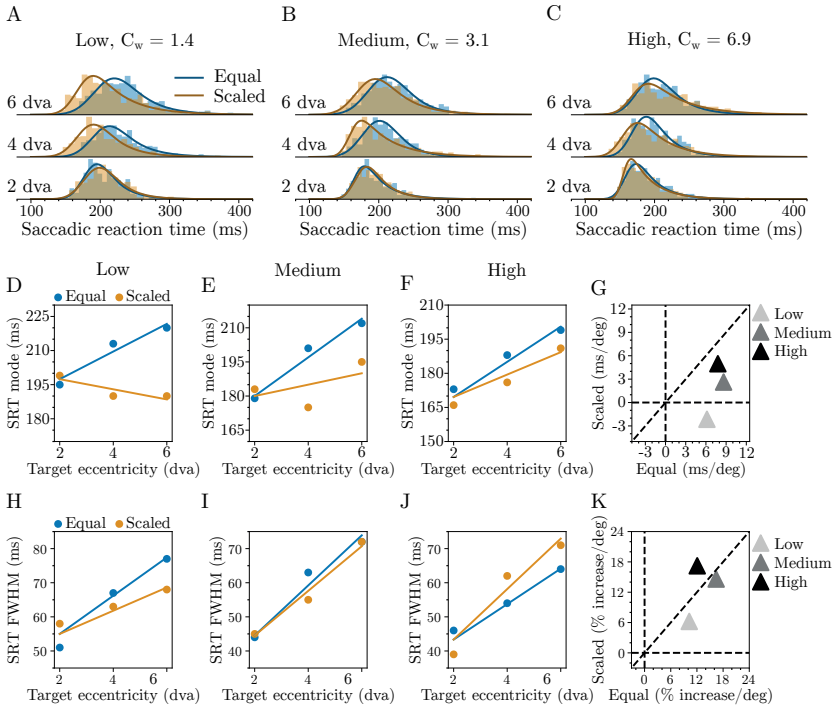


Figure 4.3: Eccentricity-dependent SRT for targets of different contrast in Monkey HO. (A-C) SRT distributions for low-, medium-, and high-contrast targets shown separately for Equal (blue) and Scaled (orange) targets. Solid lines represent the fitted ex-Gaussian distributions. Weber contrasts, C_W for low-, medium-, and high-contrast targets are 1.4, 3.1 and 6.9, as also listed on top of each column. (D-F) Multiple linear regression of SRT mode on target eccentricity and scaling for the low-, medium-, and high-contrast, respectively. (G) Scatter plot of SRT mode-eccentricity slopes for Scaled vs. Equal targets at all contrast levels. (H-K) Same as (D-G), but for FWHM.

$(targetEccentricity - 2) + \beta_2 * (targetEccentricity - 2) * isScaled)$, combining data from Equal and Scaled targets (Figure 4.3 D-F) at each contrast level. At low and medium contrast, scaling reduced the slope significantly, but the reduction of slope due to scaling was marginal for high-contrast targets (Low contrast: $CI_{\beta_2,95\%} = [-10.30, -6.00]$, $p < .001$; Medium contrast: $CI_{\beta_2,95\%} = [-8.40, -3.80]$, $p < .001$; High contrast: $CI_{\beta_2,95\%} = [-5.80, 1.20]$, $p = .063$) The above results from the multiple linear regression analysis suggest that the effect of scaling decreased at higher target contrast. This trend is illustrated graphically in Figure 4.3 G: with increasing contrast, data points moved gradually towards the equal-slope diagonal. As before, we applied the same analysis to FWHM (Figure 4.3 H-K). FWHM increased with more peripheral targets for both Equal and Scaled targets. Stimulus scaling had a smaller effect on FWHM. It reduced the slope only marginally for the low-contrast targets (Low contrast: $CI_{\beta_2,95\%} = [-4.30, .50]$, $p = .099$; Medium contrast: $CI_{\beta_2,95\%} = [-3.10, 2.30]$, $p = .537$; High contrast: $CI_{beta_2,95\%} = [-.60, 5.90]$, $p = .171$).

4.3.3 Eccentricity-Dependent SRT Increase for Endogenously Driven Saccade is Consistent with an Attentional Scan from the Fovea

So far, we focused on saccades in the Step task driven by the appearance of a peripheral saccade target, which occurs simultaneously with the disappearance of the fixation point; we refer to these as exogenously driven saccades. In the this section, we switch to the Delayed task, where the response saccades were driven by the disappearance of the fixation point, after the peripheral saccade target had already been present for 800-1500 ms, and we refer to these as endogenously driven saccades (Figure 4.1 C). The same two stimulus sets, Equal and Scaled with filled white circles were used as target stimuli. Like the Step task, each trial in the Delayed task started with the monkey holding fixation on a fixation dot for 800 to 1000 ms. The target stimulus was then presented for 800 to 1500 ms. Importantly, during this period, the monkey was required to keep fixation on the fixation dot and restrain from making a saccade. Next, the fixation dot was turned off, which signaled the monkey to make a response saccade to the target stimulus. The monkey needed to shift its gaze to the target stimulus within 500 ms and hold its gaze on the target for another 800 to 1500 ms to complete the trial and receive a reward. Only correct trials were included in our analysis. Additionally, in 50% of the trials we included a

short ≈30 ms flash at the fovea (1.0 dva diameter filled white circle) starting ≈100 ms after the fixation dot offset. The flash had no task relevance and was intended to capture attention back to the fovea when the response saccade was planned but, in most trials, not yet executed [294]. We collected data from three monkeys (KL, CH and HO) for this task.

We first present results from trials where no flash was presented. Figure 4A shows the SRT distributions from monkey KL. These SRT distributions are clearly different from the SRT distributions obtained in the Step task, wherein saccades are more reflexive. First of all, SRT in the Delayed task is generally longer. Second, SRT in the Delayed task is generally more variable. Last but not least, the effect of scaling was much smaller, with the SRT distribution for Scaled targets and Equal targets almost entirely overlapping for targets up to 8 dva. There is actually a small effect of scaling increasing SRT. All three observations are consistent with the fact that the sustained stimulus drive after the initial transient response is small and insensitive to the target size. It is worth noting that a larger stimulus may invoke lower sustained activity in the SC neurons even though its evoked transient response is larger, which agrees with the observed effect of scaling increasing SRT in the Delayed task [282].

As before, we used linear regression to quantify how the mode and FWHM of SRT distributions change with target eccentricity and stimulus scaling. Simple linear regression shows that with monkey KL, SRT mode increased with target eccentricity for both Equal and Scaled targets (Equal: $CI_{\beta_1,95\%} = [2.00, 3.22]$, $p < .001$; Scaled: $CI_{\beta_1,95\%} = [3.63, 4.75]$, $p < .001$). Interestingly, scaling slightly increased the slope, showing the opposite effect as in the Step task (Figure 4.4 B, $CI_{\beta_2,95\%} = [.85, 1.91]$, $p < .001$). Similar results are obtained with fixed-effect analysis pooling the three monkeys tested in this task (Figure 4.4 C, Equal: $CI_{\overline{\beta_1},95\%} = [1.50, 2.66]$, $p < .001$; Scaled: $CI_{\overline{\beta_1},95\%} = [1.08, 2.35]$, $p < .001$; $CI_{\overline{\beta}_2,95\%} = [.32, 1.41]$, $p < .001$). We repeated the same analysis for FWHM and the results is qualitatively the same as for the mode (Figure 4.4 D, Equal: $CI_{\beta_1,95\%} = [.51, 2.20]$, $p = .009$; Scaled: $CI_{\beta_1,95\%} = [.90, 2.52], p < .001$; Figure 4.4 E, Equal: $CI_{\overline{\beta}_1,95\%} = [1.06, 2.84]$, $p < .001$; Scaled: $CI_{\overline{\beta}_1,95\%} = [1.99, 3.92]$, $p < .001$).

The observed increase of SRT with target eccentricity in the Delayed task is intriguing. Since the delay between target onset and fixation dot offset was

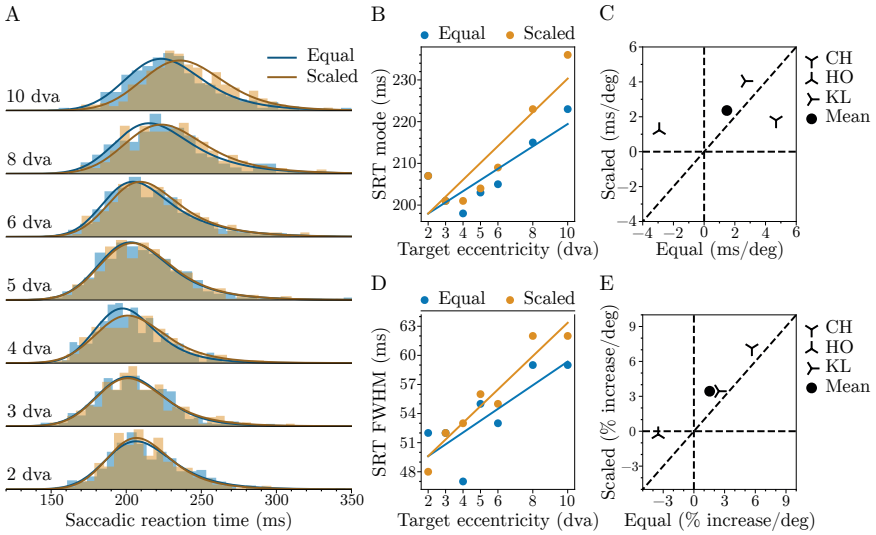


Figure 4.4: Eccentricity-dependent SRT in the Delayed task without foveal flash. (A) SRT distributions from monkey KL for Equal (Blue) and Scaled (Orange) targets. (B) Multiple linear regression of SRT mode on target eccentricity and scaling for monkey KL. (C) Scatter plot of SRT mode-eccentricity slopes for Scaled vs. Equal targets in all tested monkeys. For each monkey, the slopes are derived from the regression coefficient, as demonstrated in (B). The filled circle shows the mean slopes (D) Same as (B) but for FWHM (E) Same as (C) but for FWHM. The raw slopes were converted to percent increase per degree of visual angle.

at least 800 ms, the monkey had enough time to prepare for the impending saccade. The increase is therefore probably not related to motor planning before the "Go" signal. Rather, it is likely related to processes that happened after the "Go" signal. Additionally, because the effect is eccentricity dependent, it is likely caused by processes that are spatially relevant. One potential candidate is attentional orienting. Note that 1) because the "Go" signal in the Delayed task is the fixation dot offset, to make a timely response, the animal needed to pay attention to the fovea, even though there were two stimuli (fixation dot and peripheral target) presented at the same time, 2) before actual saccade execution, attention is mandatorily shifted to the location of the saccade target [148]. Thus, the observed increase of SRT with target eccentricity can be explained if attentional orienting from the fovea takes longer for more eccentric targets [295].

To provide further evidence that the eccentricity-dependent SRT increase in the Delayed task is related to attentional orienting from the fovea to the saccade target after the "Go" signal, in 50% of the trials, we presented a flash at the fovea, at ≈100 ms after the fixation-point offset (Figure 4.1 C). When presented, this flash would drag attention back to the fovea and force a restart of the relevant processes. If the eccentricity-dependent SRT increase is indeed due to attentional orienting from the fovea to the saccade target, the additional flash stimulus would maintain this increase, albeit with a constant delay in SRT at all target locations.

Figure 4.5 A shows the SRT distributions in trials with foveal flash from the same monkey (KL) presented in Figure 4.4. To facilitate comparison, we plotted the ex-Gaussian fits obtained in corresponding conditions without flashes as dashed lines. As expected, the addition of the flash prolonged the SRT [296, 297]. Importantly, the eccentricity-related SRT increase persisted with the addition of the flash.

We quantified these observations with linear regression. For monkey KL, SRT mode increased with target eccentricity for both Equal and Scaled targets (Equal: $CI_{\beta_1,95\%} = [4.34, 5.49]$, $p < .001$; Scaled: $CI_{\beta_1,95\%} = [3.98, 5.10]$, $p < .001$) Notably, after introducing the resetting foveal flash, which presumably further eliminated the differential activity build-up during the delay period, there is now no scaling-dependent change in slope (Figure 4.5 B, $CI_{\beta_2,95\%} = [-.22, .71]$, $p =$

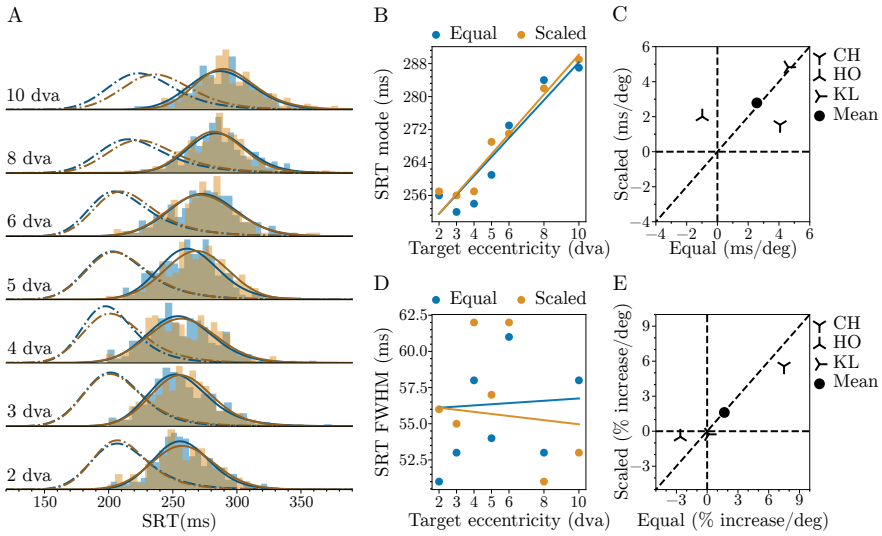


Figure 4.5: Eccentricity-dependent SRT in the Delayed task with foveal flash. (A) SRT distributions from monkey KL for Equal (Blue) and Scaled (Orange) targets. Dashed lines represent the ex-Gaussian fits for SRT distributions in corresponding conditions without the foveal flash (as in Figure 4). (B) Multiple linear regression of SRT mode on target eccentricity and scaling for monkey KL. (C) Scatter plot of SRT mode-eccentricity slopes for Scaled vs. Equal targets in all tested monkeys. For each monkey, the slopes are derived from the regression coefficient, as demonstrated in (B). The filled circle shows the mean slopes (D) Same as (B) but for FWHM. (E) Same as (C) but for FWHM. The raw slopes were converted to percent increase per degree of visual angle.

.365). Intriguingly, the foveal flash also eliminated the eccentricity-dependent FWHM increase (Figure 4.5 D, Equal: $CI_{\beta_1,95\%} = [-.32, 1.45]$, $p = .223$; Scaled: $CI_{\beta_1,95\%} = [-1.40, .26]$, $p = .163; CI_{\beta_2,95\%} = [.91, .60]$, $p = .586$). Fixed effect analysis combining the three monkeys shows similar pattern as described above for Monkey KL (Figure 4.5 C, the mode: Equal, $CI_{\overline{\beta_1},95\%} = [2.63, 3.50]$, $p < .001$; Scaled, $CI_{\overline{\beta_1},95\%} = [1.88, 2.87], p = .016\, CI_{\overline{\beta_2},95\%} = [-.18, .58]$, $p = .274$; Figure 4.5 E, FWHM: Equal, $CI_{\overline{\beta_1},95\%} = [1.14, 3.32]$, $p < .001$; Scaled, $CI_{\overline{\beta_1},95\%} = [.16, 2.15], p = .016$; $CI_{\overline{\beta_2},95\%} = [-.90, .81]$, $p = .876$).

4.4 Discussion

Saccadic eye movements play a fundamental role in organizing our spatial and temporal visual input, to the extent that SRT has become a crucial psychophysical measurement enabling us to characterize the underlying cognitive processes

that would otherwise remain unseen. However, despite the importance of this measurement, some fundamental questions regarding the determinants of SRT remain, even in relatively simple situations. In this study, we examined the dependence of SRT on target eccentricity in macaque monkeys.

When designing the experiment, we based our hypotheses on previous insights into the role of the SC in generating the response saccades: Once the SC decision unit passes the response threshold, there is a relatively fixed delay in the downstream motor nuclei before the actual eye movement occurs [113, 298] Consequently, any observed dependence of SRT on target eccentricity can be attributed to the spatiotemporal activation pattern of the SC, and reflects the properties of the neural processes associated with that particular task. It is these properties that have clear behavioral correlates, namely SRT, the primary interest of the current study.

The SRT reflects a combination of the time required for target detection, response initiation and execution. A dissociation of these components, respectively their dependence on eccentricity, might have been achieved by, e.g., including a condition with a manual response that remains identical across eccentricities. Such a dissociation is beyond the scope of the present study, but is an interesting target for future research.

4.4.1 SRT in the Step Task

In the Step task, afferent input to the SC was presumably dominated by the transient onset of the target stimulus. We first confirmed previous findings that, for medium-sized saccades, SRT increased with target eccentricity. Notably, this SRT increase was primarily explained by the foveal magnification factor in the SC (Figure 4.2 and Figure 4.3). The observed effect of foveal magnification on SRT can be explained by the increase of the accompanied RF size from the foveal to the peripheral representation in the SC. This increasing RF size compensates for the foveal magnification, thus equalizing the size of the active population for a point stimulus presented at different eccentricities, achieving a constant point image in the SC [140, 141]. However, larger RFs in the periphery also mean that an equal-sized stimulus occupies a smaller proportion of the corresponding SC neurons' RFs. The smaller occupied RF proportions likely drive the active population less strongly, leading to longer SRT [282]. When we scaled the stimulus size according to the SC magnification factor, the above-

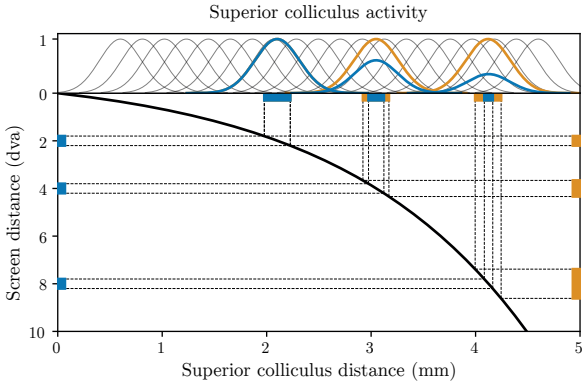


Figure 4.6: Proposed mechanism of how stimulus scaling reduces SRT in the Step task. The foveal magnification is illustrated by the thick black curve mapping points on the screen (Y-axis) to points on the SC (X-axis). Blue (orange) bars on the left (right) represent the Equal (Scaled) targets at 2, 4, and 8 degrees, respectively. The corresponding image of these stimuli in the SC is represented by horizontal blue (orange) bars at the top. These stimulus images represent the foveally-magnified afferent input to the SC. The thin gray Gaussian curves on the top illustrate example visual connectivity functions of SC neurons integrating the afferent inputs. Notice that homogenous connectivity functions, as we assumed here, correspond to RFs of increasing size with higher eccentricity when projected to the visual space. The activity level of a given SC neuron is modeled as the dot product between its Gaussian connectivity function and the afferent input. The activity profile across the SC-neuron population is calculated as the convolution between the connectivity function and the afferent signal. The thick colored curves on the top illustrate the resulting activity profiles evoked by each stimulus. Note that equally-sized stimuli (blue) activate the SC more weakly in the periphery, whereas scaled stimuli (orange) activate the SC equally strongly at all positions. Note also that the sizes of the activated neuronal populations are approximately equal because the stimuli are small compared to the SC neurons' RFs.

mentioned mechanism would predict an equal SRT at all eccentricities, which is what we found for the Scaled targets. This process is illustrated in Figure 4.6. A similar effect of foveal magnification neutralizing the eccentricity effect has been reported for a covert visual search task [299].

Furthermore, we showed that the foveal magnification factor interacted with stimulus contrast in determining the SRT. This interaction also aligns with the known properties of neurons in the SC. Specifically, the observation that, in monkey HO, where we also modulated the target contrast, stimulus scaling modulated SRT most effectively for low-contrast stimuli can be explained by

the visual response properties of SC neurons. Early recordings measuring the size tuning of SC visual responses found that with stimuli of luminance contrast 1-3 log unit higher than the response threshold, presumably corresponding to the high contrast condition in the current study, many neurons' response magnitudes are insensitive to stimulus size. Recently, Chen and Hafed [282] provided direct evidence for such interaction between stimulus contrast and size in driving the SC visual response. They showed that for a given SC neuron, a large stimulus evokes a larger response only if the contrast is low. When the contrast is high, the response magnitude is similar for both small and large stimuli. This interaction pattern can explain why scaling was most effective for low-contrast targets in eliminating the eccentricity-dependent SRT increase in our data (Figure 4.3). Together, these findings imply a direct translation of saccade target visibility (combining the effect of contrast and size) to response urgency mediated by the SC [300].

Our analysis has been focused on comparing SRT across different eccentricities to test whether a particular kind of scaling, i.e., scaling according to the foveal magnification in the SC, affected this SRT-eccentricity relationship. Conversely, one could compare SRT at fixed target eccentricity across different scaling factors. If our hypothesis is true, that the relative afferent input strength determines the SRT-eccentricity relationship in the Step task, and that the relative size of the saccade target to the corresponding neurons' RF can approximate this afferent input strength, one would expect that increasing the size of the saccade target at a fixed eccentricity shortens the SRT. Indeed, this expected effect of varying saccade target size on SRT has been reported before for targets that are small relative to the SC neuron's RF size [301]. Different processes appear to be involved in saccading to very large targets [302].

Scaling the target has an incidental consequence: a scaled target is of a different size than the fixation dot and all the non-scaled saccade targets, making it a relatively novel stimulus. Stimulus novelty can decrease reaction time [303]. However, for the difference in stimulus novelty to cause the observed effect of target scaling, i.e., flattening the SRT-eccentricity function, an additional hypothesis would be required to assign quantitatively appropriate levels of novelty to target stimuli of different sizes. Nevertheless, future research would benefit from using neutral fixation points of different types than the saccade targets to alleviate this concern [304].

4.4.2 SRT in the Delayed Task

In the Delayed task, when the monkey was about to make the response saccade, the target had already been on the screen for at least 800 ms, long beyond the target-onset-related visual burst in the SC neurons, which diminishes around 100 ms after the stimulus onset [305, 306]. As a result, the afferent input relevant to the response saccade was no longer expected to be dominated by the transient onset of the target stimulus. In agreement with this prediction, in the Delayed task, the dependence of SRT on eccentricity was minimally affected by scaling. Yet, the simple dependence of SRT on eccentricity was still present, leading to the question: What is the mechanism underlying this SRT increase? One potential candidate is pre-saccadic attentional shifting. Results from trials with foveal flashes provided further evidence supporting this hypothesis, as explained in the Result. If the observed SRT increase is indeed related to pre-saccadic attentional shifting, the regression slopes will inform us about how fast attention was moving from the fovea to the peripheral targets: ≈ 2.6 ms/dva as the average from three monkeys (Figure 4.5). This value, corresponding to ≈17.7 cm/s in the SC, is in a similar range as the speed of proposed attentional scanning estimated in a recent opinion paper by Fries [295]. The reported estimate was obtained partly based on neurophysiological observations and partly on a covert attention task containing a periphery reset signal. Covert attention and pre-saccadic attention are dissociable processes that can exhibit different temporal dynamics (for a recent reviewer, see Li, Hanning, and Carrasco [307]). For a meaningful comparison between our results and the reported estimate, it is necessary to point out that a salient transient signal, e.g., the foveal or peripheral flash, captures attention to the spatial location of that transient in both processes [308, 309], and the moment-by-moment spatial location of attention in both processes can be determined by probing behavioral performance. These considerations help to align the observed attentional shifting speed with the estimate in Fries [295], despite the inherent differences between covert and pre-saccadic attention.

Alternatively, the SRT increase in the Delayed task might reflect a speed-accuracy tradeoff. One way to describe the speed-accuracy tradeoff is to use the classical Fitts's Law [310].According to Fitts's Law, $M = A + B * \log_2(\frac{2S}{D})$, the movement time (M) increases linearly with the logarithm of the ratio between movement distance (S) and target diameter (D). If we consider M to be the SRT, S to be the target eccentricity, and D to be the target size, Fitts's Law, however,

cannot explain the SRT values observed across conditions. Specifically, the fact that at the same target eccentricity, Scaled targets yielded similar or even longer SRT compared to the Equal targets is not consistent with Fitss's Law. One might instead take D to be the size of the invisible target window, which was the same in all conditions. But then, it is difficult to explain the effect of the scaling observed in the Step task. In both cases, Fitts's Law struggles to explain the variability of SRT in our data. Note that, Wu, Kwon, and Kowler [311] found that Fitts's Law describes the speed-accuracy tradeoffs in a saccade sequence only when the time for the secondary, corrective saccades is included. When only the saccade latency for the primary saccade is considered as what we did in our analysis, Fitts's law no longer holds. Furthermore, within the context of a diffusion model, the speed-accuracy tradeoff is typically modeled by adjusting the decision threshold [312]. With other parameters staying unchanged, a higher threshold implies both a longer response time and a larger variance in the response time distribution. However, we do not find that SRT mode relates consistently to FWHM. For the respective example monkey, in the Step task, scaling eliminated the eccentricity-dependent SRT mode increase but not the FWHM increase (Figure 4.2); in the Delayed task, adding a foveal flash eliminated the FWHM increase but not the SRT mode increase (Figure 4.5). All these observations indicate that a speed-accuracy tradeoff, at least in isolation, is unlikely to explain the observed SRT variation.

If the pre-saccadic attentional shift caused the SRT increase in the Delayed task, one would expect to see a similar SRT increase for medium-sized saccades in the delayed visually guided saccades task of Hafed and Goffart [268]. Indeed, although the SRT increase in their delayed task is less prominent than in their Step task, it is clearly trending. The relatively small SRT increase in the delayed task is also apparent in the present study (comparing Figure 4.5 and Figure 4.6 to Figure 4.2 and Figure 4.4). In addition, it is worth noting that our SRT values are generally longer than those reported by Hafed and Goffart [268]. This discrepancy is likely due to the smaller target windows used in our study (1.0 dva vs. 2-2.5 dva in radius). The size of the fixation window may be related to the speed-accuracy tradeoff as discussed earlier. Although the process involved in the speed-accuracy tradeoff may not be the direct cause of the SRT increase, it is still worth exploring how these factors interact.

Pre-saccadic attentional shifting is not expected to significantly affect SRT in

the Step task. In the Step task, attention is automatically captured by the onset of the peripheral target, a process that is presumably not dependent on eccentricity. In contrast, the Delayed task involves triggering saccades by the offset of the foveal fixation dot, requiring a voluntary movement of attention to the peripheral target, which we hypothesize to be eccentricity-dependent.

4.4.3 Variability of SRT

As noted in previous studies, SRT exhibited significant trial-by-trail variation in both the Step and Delayed tasks. Within the framework of a rise-to-threshold model, a larger variance in SRT means either a higher threshold or a more variable rate of rise [225, 276, 284, 285]. A consistent finding in our study is the increased variability of SRT for targets located in the periphery. This pattern held true for both Equal and Scaled targets in both the Step and Delayed tasks. Notably, in the Step task with Scaled targets, SRT variance increased with target eccentricity, despite the effective flattening of the SRT mode by scaling (Figure 4.2 and Figure 4.3). These findings suggest that factors beyond foveal magnification, and to some degree invariant to the task, contributed to the observed SRT variance increase with target eccentricity. One potential mechanism could be that neuronal response fields (both motor and visual) of larger eccentricity typically have larger size. Larger response field sizes likely lead to less specific, or in other words, noisier contributions to the ensuing saccade. When cortical magnification is controlled, this leads to higher noise in the recruited neuronal population. Interestingly, Hafed and Chen [313] reported a weaker and more variable visual response associated with less accurate saccades in SC neurons with a larger response field. Note that these authors compared neurons representing lower versus upper visual fields, yet similar properties are conceivable comparing neurons with response fields centered at different eccentricities.

4.4.4 Individual Differences

We observed considerable variations in SRT distributions among the four tested monkeys. These variations likely have multiple sources. First, some variance in SRT distributions could be related to the task design. Most prominently, the maximum SRT allowed was 500 ms at all locations in all tasks. This setting is essential for assuring the same task structure across all conditions despite the difference in their typical SRTs. However, this setting is not optimal for

eliminating the influence of task-irrelevant processes. Indeed, the monkeys whose SRTs showed the largest variance, suggesting extra noise in their decision process, were also the ones that seemed to be outliers (HU in Figure 4.2 C and HO in Figure 4.4 C). Since using a long response window is required to compare SRT distributions that are expected to shift significantly, future studies can use variable reward amounts to encourage the animal to make as fast a response as possible and potentially reduce the variability among the subjects [212]. Second, variations associated with scaling might be partially explained by the individual differences in foveal magnification factors [314]. Finally, it is conceivable that some of the monkeys, e.g., monkey HU, had comparatively poorer vision. Overall, the vision was sufficient for all tested monkeys to produce the observed, consistent behavior. However, subtle deficiencies in vision might have contributed to the observed individual differences, and the individual visual acuity was not tested here.

Increasing the number of subjects is crucial for obtaining a more comprehensive understanding of whether the observed effect is present in the population. However, in the current study, the number of subjects was limited, partly because of the chosen animal model, i.e., macaques. Macaque monkeys were used due to the availability of published data on the foveal magnification factor in the SC, which was essential for the experimental design. Additionally, there is a rich history of using macaque monkeys to study the underlying neural mechanisms of oculomotor control and attention. Choosing macaque monkeys thus offers the potential to use these findings to motivate future circuit investigations. In human studies, it is generally easier to include a larger number of subjects; however, there is no direct measurement of the SC foveal magnification in Humans. Nevertheless, it is not unreasonable to utilize the same magnification factor as observed in macaque monkeys. Alternatively, one can consider using the magnification factor measured for the primary visual cortex, as evidence indicates that the magnification factor of the primary visual cortex and SC are similar [245].

4.5 Materials and Methods

4.5.1 Subjects

Four male adult monkeys (Macaca mulatta) participated in the current study (referred to as CH, HO, KL and HU). All experimental procedures were conducted in compliance with the German and European animal protection laws. The experiments were approved by the responsible local authority, the Regierungspräsidium Darmstadt. All animals were implanted with a titanium headpost [219]. Additionally, Monkey KL and HO had recording chambers implanted for addressing other scientific questions. The animals' water intake was regulated to ensure their motivation during the behavioral task in which a small juice reward was provided after each correct trial.

4.5.2 Visual Stimuli

All stimuli were controlled by custom software (`https://github.com/esi-neuroscience/ARCADE`), and presented on an LCD monitor (LG 32GK850G-B) at 143.9Hz. Viewing distance was 78 cm for all monkeys. Precise stimulus presentation time was validated with a photodiode attached to the screen. Parameters for all stimuli used are listed below.

Fixation dot. Trials in all tasks started with presenting a fixation dot. The fixation dot in all trials was a white filled cycle of 0.1 dva diameter, 243.7 cd/m^2, at the center of the screen.

Target stimuli. Each trial contained a single target stimulus chosen randomly from a predetermined set. As explained in the introduction, we focused on medium-sized saccades between 2 and 10 dva. In addition, to increase the trail counts per condition, we limited the saccade targets to the right visual field. Specifically, the stimulus set in the Step and the Delayed task contained targets centered at 2, 3, 4, 5, 6, 8, and 10 dva along the horizontal median to the right of the screen. At each target location, the set contained one target belonging to the Equal group and one target belonging to the Scaled group. For the Equal group, the target was a white filled circle of 243.7 cd/m^2 and 0.1 dva diameter at all eccentricities. For the Scaled group, the target was a white filled circle of 243.7 cd/m^2 and 0.1, 0.13, 0.17, 0.20, 0.24, 0.31, 0.38 dva diameters at 2, 3, 4, 5, 6, 8, and 10 dva, respectively. All target stimuli are small relative to the

spacing between them. In the modified Step task with variable target contrast, the target locations were limited to 2, 4, and 6 dva, and at each eccentricity and for both Scaled and Equal targets, the possible target set contained targets of three different luminance levels: namely low contrast, 74.57 cd/m^2; medium contrast, 124.66 cd/m^2 and high contrast 243.7 cd/m^2.

Screen background. Gray background was used in all tasks. Except for the modified Step task with variable target contrast, the screen background luminance was 60.0 cd/m^2. In the modified Step task, the screen background luminance was 30.7 cd/m^2.

Foveal flash. The foveal flash used in the Delayed task was a white filled circle of 1.0 dva and 243.7 cd/m^2.

4.5.3 Behavioral Task

We recorded the binocular gaze data at 500 Hz with an Eyelink-1000 system (SR Research Ltd.). Real-time gaze data from one eye was used for online behavioral control. The real-time monitoring of response saccade initiation used a virtual fixation window (r≈0.8 dva) centered on the fixation dot. Saccade landing on the target was determined by gaze entering and staying in a virtual target window (r≈1.0 dva) centered on the saccade target. The target window radius was the same in all conditions. Additionally, the saccade flight time between leaving the fixation window and entering the target window was monitored and limited to less than ≈80 ms. In all tasks, trials of different conditions were randomly interleaved.

In the Step task (Figure 4.1 B), after fixation was acquired and maintained for 800 to 1500 ms, the fixation dot was removed, and in the same video frame, a single target was presented in the periphery. The animal needed to make a saccade, within 500 ms, to the target and hold its gaze on the target for another 800 to 1500 ms to obtain the juice reward.

In the Delayed task (Figure 4.1 C), after the fixation was acquired and maintained for 800 to 1000 ms, without removal of the fixation point, a single target was presented in the periphery for 800 to 1500 ms after which the fixation dot was removed. Only after fixation dot removal, the monkey was allowed to make a saccade to the target, and it needed to make a saccade towards the target within

500 ms and hold its gaze on the target for another 800 to 1500 ms to successfully complete the trial and obtain the reward. Additionally, in 50% of the trials in the Delayed task, we included a ≈30 ms flash at the fovea (white-filled circle of 1.0 dva diameter) presented ≈100 ms after the fixation dot offset. The flash had no task relevance, and the monkey needed to complete the trial as in trials without flash.

4.5.4 Trial Inclusion Criteria

Only correct trials were included. As described above, in these trials, the monkey made timely (SRT<500 ms) and relatively precise (error below ≈1.0 dva) primary saccades to the target. Additionally, 1) Trials with SRTs less than 100 ms were excluded from all tasks. 2) In Delayed task trials containing a foveal flash, any trial with the response saccade occurring within 100 ms after the flash was also excluded.

4.5.5 Saccade Detection and SRT Definition

Saccades were detected using a two-dimensional velocity threshold, following the methodology described in Engbert and Kliegl [223]. In our analysis, we set lambda to 10. Saccade onset time was determined as the moment when the velocity threshold was crossed. In all tasks, SRT was defined as the time elapsed between the offset of the fixation dot and the onset of the response saccade.

4.5.6 Regression Analysis

The regression analysis was performed with the Python module statsmodels [315]. Further details can be found in the corresponding Results section. We used a permutation test, permuting the respective regressor label ($N = 1000$), to test against the null hypothesis that the coefficient equals zero. The two-tailed p-value calculated from the resulting permutation distribution are reported in the main text. Bootstrapping ($N = 1000$) was used to calculate the confidence interval. Each bootstrap sample was obtained by resampling trials within each group of trials corresponding to individual data points in the respective regression analysis.

CHAPTER 5

Gamma-Band Neuronal Synchronization Tracks Prior Updates of Stimulus Occurrence Frequency[1]

5.1 Abstract

Integrating prior knowledge with sensory evidence is fundamental for visual perception. But how prior knowledge is stored, updated, and used in the brain remains unclear. This chapter proposes that gamma-band neuronal synchronization, as reflected in gamma-band local field potential (LFP) power, tracks prior updates of stimulus occurrence frequency in a repetition task. Specifically, we modeled the repetition task as a series of Bernoulli trials with a beta-distributed prior over stimulus occurrence frequency, and we hypothesized that the gamma-

[1]The results presented in this chapter are based on data from a separate project detailed in the following manuscript in preperation: Eleni Psarou et al. "Repetition-related gamma plasticity in macaque V1 and V2 is highly stimulus specific and robust to stimulus set size". 2024. The conception of this chapter was inspired by the findings of Eleni Psarou and Dr. Mohsen Parto Dezfouli reported in the aforementioned manuscript. The author of this thesis received the preprocessed data, performed the analyses, and wrote this chapter. Special thanks to Eleni Psarou and Dr. Mohsen Parto Dezfouli for preparing the data, and to Dr. Pascal Fries, Christini Katsanevaki, and Dr. Tianlu Wang for the inspiring discussions.

band LFP power indexes the mode of updating prior distribution. With only three parameters, this model captured over 80% of the variability of the seemingly complex pattern of gamma-band LFP power over the session, including the non-linear increment of power with repetition, the partial persistency of the repetition effect over an intervening break, and the decreasing power with inter-leaved repetition, only after adaptation. Notably, if gamma-band neuronal synchronization indeed tracks prior updates, it also offers a straightforward mechanism to update the prior information through Hebbian plasticity and compute with such prior information through coherence.

5.2 Introduction

Natural visual scenes are rich with statistical regularities. These regularities range from low-level spatial frequency [317] and orientation contour distributions [117, 318] to high-level context-dependent object occurrence and co-occurrence frequencies [319–321]. Furthermore, we often actively select stimuli with specific features relevant to the current behavioral goal or simply revisit the same object [88, 264, 322]. As a result, our moment-to-moment visual inputs are often repeated.

Conversely, the repetition frequency of a specific visual feature or feature constellation reflects the statistical regularities in the environment. For optimal analysis of future visual inputs, especially under conditions of uncertainty, it is beneficial to learn from these repetition patterns such that knowledge of the current environment can be combined with incoming sensory evidence [265, 323–326]. A notable example of this learning process is reported in kittens raised in environments where only one orientation is visible [327]. In such a setting, this single orientation is repeated with a 100% occurrence frequency for an extended period. As a result, the kitten's visual cortex adapted significantly to the environment, with almost no neurons in area 17 responding to the orthogonal orientation. Similarly, in the adult primate brain, temporally increased stimulus occurrence frequency, achieved through repeated stimulus presentations, can also modulate neuronal response properties. Two distinct effects of repeated stimulation have been noted. The first, termed repetition suppression, is characterized by a reduced neuronal firing rate or hemodynamic response to the repeated stimulus [60, 328, 329]. The second effect involves enhanced gamma-band synchronization, noticeable in both macaque monkey local field

potentials and human magnetoencephalography (MEG) signals, particularly evident with a high number of repetitions of the same stimulus [61, 62, 257].

In this chapter, I propose that the enhanced gamma-band synchronization during stimulus repetition tracks the prior update of stimulus occurrence frequency. This is inspired by one key observation described in Psarou et al. [316]: gamma-band LFP power stopped increasing when the stimulus occurrence frequency was down-modulated while the number of repetitions still increased. We modeled how the prior distribution over stimulus occurrence frequency is updated in an experimental session and tested eight predictions, from qualitative to more quantitative ones, if gamma-band LFP power indeed tracks the prior updates. The results supported our hypothesis, and the observed dynamics of gamma-band neuronal synchronization can largely be accounted for by how the prior distribution over stimulus occurrence frequency is updated as the session progresses.

5.3 Results

We analyzed local field potentials (LFP) recorded from the primary visual cortex of an awake rhesus macaque (*Macaca mulatta*) doing a passive fixation task. The task contained three epochs. The first (pre-adaptation) epoch consisted of 10 mini-blocks, each with 18 trials displaying a full contrast grating stimulus at one of eighteen unique orientations (10 to 180 degrees, in 10-degree increments). The sequence of these 18 orientations was randomized for each mini-block. There was a short break of around 20 minutes, after which the task entered the adaptation epoch. The adaptation epoch presented 90 trials of a single orientation (the adaptor grating), randomly chosen from the 18 orientations for each session. The adaptation epoch was immediately followed by the post-adaptation epoch, which mirrors the structure of the pre-adaptation epoch. The grating's positions was fixed for every single session, overlapping the receptive field of the recorded neuronal population.

The grating stimuli induced clear gamma-band neuronal synchronization in the LFP power spectrum (Figure 5.1 B). As detailed in Psarou et al. [316], for the adaptor orientation, we observed that gamma-band LFP power varied characteristically throughout the session (Figure 5.1 D). Gamma-band LFP power increased before and during adaptation, but showed a noticeable decrease

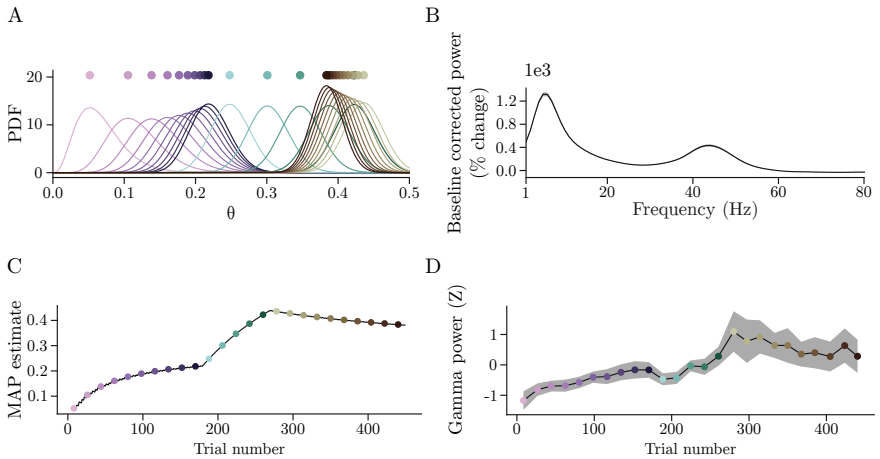


Figure 5.1: The model and observed gamma power. (A) Prior updates over the session. Purple, pre-adaptaiton epoch. Green, adaptation epoch. Brown, post-adaptation epoch. (B) Average power spectrum over all trials and orientations during the pre-adaptation epoch. (C) Trajectory of MAP estimate. Same color as in (A). (D) Observed gamma-band LFP power. Same color convention as in (A). The shaded region represents two times the SEM.

during the post-adaptation epoch. This decrement contrasts with the expected increment due to stimulus repetition, as reported earlier in both macaque monkeys and human participants [61, 62, 257]. To resolve this discrepancy, we propose that instead of reflecting the number of stimulus repetitions, gamma-band neuronal synchronization tracks the updates of inferred stimulus occurrence frequency according to Bayes' rule. In the following sections, we will describe in detail how this inference process can be modeled and then examine eight predictions it generates, ranging from qualitative to quantitative.

5.3.1 The Model

Our model posits that the brain observes the stimulus being presented in each trial and draws inferences about its occurrence frequency based on prior beliefs and the session's trial history. Specifically, for a given grating orientation S, the session can be modeled as a series of Bernoulli trials with Θ being the occurrence frequency of S. This frequency is presumed to follow a Beta prior distribution of the form:

$$f(\theta;\alpha,\beta) = c \cdot \theta^{\alpha-1}(1-\theta)^{\beta-1};\ \text{where } c = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)}$$

With Beta distribution being the conjugate prior for Bernoulli trials, each new presentation of S adjusts the Beta distribution by incrementing the parameter α, while presentations of other orientations increment the parameter β [330]. Based on the bandwidth of orientation tunning reported in the primary visual cortex, we further assumed that gratings less than 20 degrees apart drive the underlying neurons similarly [331]. However, this assumption is not critical to our overall findings.

Figure 5.1 A illustrates the prior distribution updates for the adaptor grating during an example session. For clarity, we plotted one distribution per 18 trials, aligning with the mini-block length in our task. We plotted the prior distributions in the pre-adaptation, adaptation, and post-adaptation epochs in purple, green, and brown, respectively. The progression within each epoch is represented by the shading of color from light to dark. It can be observed that 1) in the pre-adaptation epoch, the distribution shifts rightward; 2) during adaptation, this rightward shift continues with a higher speed; 3) in the post-adaptation epoch, the distribution moves back but with smaller steps. For this illustration and all model predictions that are presented in Figures 5.1 to 5.4, we have initialized the model with $\alpha = 1$ and $\beta = 50$ before the first trial. Although this pattern is robust over a large parameter set, having β considerably larger than α is necessary, possibly reflecting the low frequency of occurrence for any specific grating stimulus in everyday visual experience, even for a monkey used in a vision lab (discussed further in the Parameter Selection section).

Having updated the prior distribution in each trial, the inference about the stimulus occurrence frequency is assumed to be done by selecting the maximum a posteriori (MAP). For the Beta distribution, the MAP estimate $\hat{\Theta}_{MAP}$ has a simple form, $\frac{\alpha-1}{\alpha+\beta-2}$. In Figure 5.1 A, we plotted $\hat{\Theta}_{MAP}$ for each trial with a dot on top of the corresponding prior distribution. Figure 5.1 C shows how this MAP estimator evolves over the course of the session. This trajectory closely mirrors the observed gamma-band LFP power trajectory (Figure 5.1 D), suggesting that gamma-band neuronal oscillations may track the prior updates of stimulus occurrence frequency. Notice that the calculation of gamma-band LFP power trajectory depicted in Figure 5.1 D used the same mini-block binning

as in Figure 5.1 A. Because prior to the adaptation epoch, all orientations are expected to give the same response (when averaged over recording sessions), the gamma responses to all orientations are averaged to estimate the gamma-band LFP power in the pre-adaptation epoch.

If our hypothesis is true, that gamma-band neuronal synchronization tracks the prior update of stimulus occurrence frequency, then we can predict how gamma-band LFP power changes during a session based on how the prior distribution is updated. In the subsequent sections, I test eight such predictions, ranging from qualitative predictions that are robust to parameter selection to more quantitative predictions that involve fitting model parameters to the observed data.

5.3.2 Predictions I-IV: Trends of Gamma-Band LFP Power

Prediction I posits that gamma power for all grating orientations would increase in the pre-adaptation epoch (Figure 5.2 A). According to the model, if the initial prior belief of stimulus occurrence frequency is lower than the actual observed frequency in the mini-block, then the MAP estimate is expected to rise. To verify this, we calculated the Spearman rank correlation between the trial number and average gamma-band LFP power across mini-blocks (Figure 5.2 E). Our results confirm an increase in gamma-band LFP power during the pre-adaptation period ($\rho = .98$, $CI_{95\%} = [.89, 1.00]$, $p < .001$).

Prediction II posits that gamma power induced by the adaptor orientation would increase in the adaptation epoch (Figure 5.2 B). This prediction is expected to be robust. Because the actual stimulus occurrence frequency in this period is one, the maxium possible value. The model will thus always predict an increase in the MAP estimate with each new observation of the adaptor orientation. To validate this, we calculated the Spearman rank correlation between gamma power and trial number (Figure 5.2 F). Our findings confirm this prediction. There is a significant increase in gamma power during the adaptation epoch ($\rho = .85$, $CI_{95\%} = [.58, .92]$, $p = .001$). Previous studies have noted a transient decrease in gamma power during the initial 3 to 10 repetitions, which might slightly affect the correlation coefficient. To enhance statistical power, for this analysis in the adaptation epoch, 9-trial bins are used instead of the 18-trial bins. This can be done because the adaptation epoch lacks the 18-trial mini-block structure as in the pre- and post-adaptation epoch.

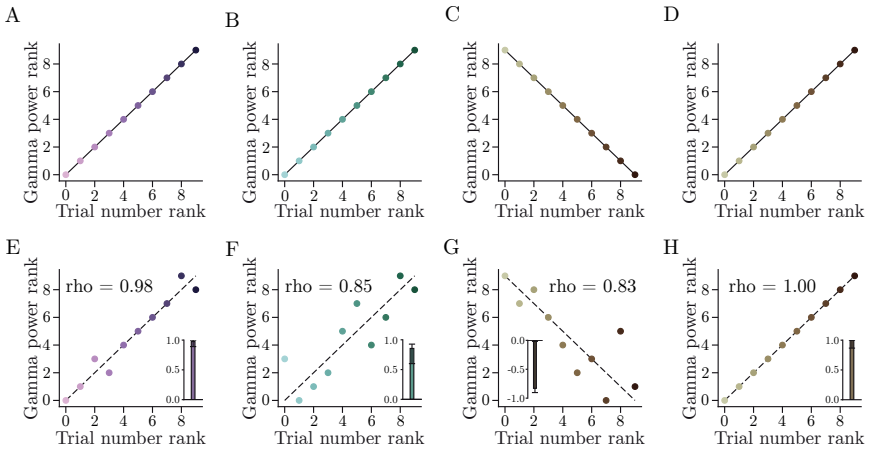


Figure 5.2: Prediction I-IV. (A-D) Predicted gamma-band LFP power rank during pre-adaptation epoch (A), adaptation epoch (B), and post-adaptation epoch (C, adaptor; D, non-adaptor). (E-H) Corresponding observed gamma-band LFP power rank. Insets show the 95% bootstrap confidence intervals of Spearman's rho.

Prediction III posits that gamma power for the adaptor orientation would decrease in the post-adaptation epoch (Figure 5.2 C). This is expected because 90 consecutive repetitions would likely elevate the belief about the stimulus occurrence frequency above the stimulus occurrence frequency in the mini block which is merely one over eighteen. The observed negative Spearman correlation coefficient (Figure 5.2 G, $\rho = .83$, $CI_{95\%} = [-.90, -.02]$, $p = .006$) agrees with the prediction. The relatively large confidence interval is because the analysis needs to be applied specifically to the adaptor orientation (the only orientation repeated during the adaptation epoch). Focusing on the adaptor orientation results in a smaller sample size for this analysis compared to the pre-adaptation epoch. The smaller sample size is also reflected in the larger error bars for post-adaptation gamma power in Figure 5.1 C.

Prediction IV posits that gamma power for non-adaptor orientations would rise in the post-adaptation epoch (Figure 5.2 D). This is because these orientations are absent during the adaptation epoch. The lack of exposure to these non-adaptor orientations for 90 trials is expected to lower the MAP estimate of their occurrence frequency, leading to an increase in MAP estimate when they reappear. To test this, we grouped data from all gratings more than 30 degrees away from the adaptor orientation. In line with our prediction, an increase in

gamma power for these orientations was observed in the post-adaptation epoch (Figure 5.2 H, $\rho = 1.00$, $CI_{95\%} = [.87, 1.00]$, $p = .006$).

5.3.3 Prediction V: The Speed of Gamma-Band LFP Power Increment Decreases in the Pre-Adaptation Epoch

Prediction V posits that the speed of gamma power increment would decrease over time in the pre-adaptation epoch (Figure 5.3 A). In the pre-adaptation epoch, as new observations accumulate, the mode of the prior distribution gradually approximates the mini-blocks' true stimulus occurrence frequency. This implies a progressively slower MAP increment speed over time. This is tested by first calculating the rate of gamma power change between consecutive mini-blocks and then assessing the correlation of these rates with trial numbers using Pearson's correlation coefficient. In agreement with the model prediction, a negative correlation was found, indicating that the speed of increment decreased (Figure 5.3 D, $\rho = -.52$, $CI_{95\%} = [-.64, -.14]$, $p = .015$). The decrement of gamma power increment speed has been reported before. Specifically, prior studies have shown that gamma power increases logarithmically with the stimulus repetition numbers, which also implies a decreasing rate of change [61, 257]. While the usage of a logarithm function in these studies was based on empirical observations, our prior update model makes quantitative predictions about the rate of change. Incidentally, within a certain parameter range, logarithm transformation approximates the speed of change predicted by the proposed model (Figure 5.3 A, with $\alpha = 1$ and $\beta = 50$ at the beginning of the session). However, it's important to note that the log transformation is only effective within a certain parameter range for explaining the rate of MAP estimate change and may only explain the local speed of change. Accordingly, the logarithmic function may not always effectively explain the speed of gamma power change. The proposed model may explain why the log transformation sometimes works and sometimes does not.

5.3.4 Prediction VI: Faster Gamma Power Increment with Non-Interleaving Repetition

Prediction VI posits that gamma power induced by the adaptor orientation would increase more rapidly with non-interleaving repetition during the adap-

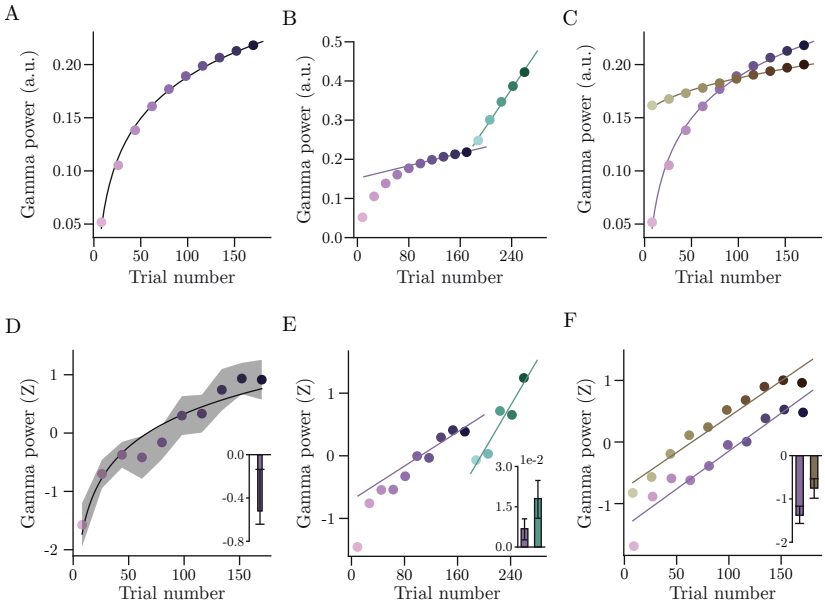


Figure 5.3: Prediction V-VII. (A) Predicted speed of gamma power increment in the pre-adaptation epoch. The solid black line shows the linear fit on the logarithm of the trial number. (B)Predicted faster gamma power increment with non-interleaving repetition. Purple lines, linear fit with the last five bins in the pre-adaptation epoch. Green line, linear fit during adaptation. (C) Predicted persistency of gamma power increment. (D) Observed data for (A). The solid line indicates the logarithm fit on the observed gamma-band LFP power. The shaded region indicates two times SEM. Insets show the bootstrapped confidence interval of person correlation of pair-wise slopes. (E) Observed data for (B). Insert shows the bootstrap confidence interval for the respective slope. (F) Observed data for (C). Insert shows the confidence interval for respective intercept.

tation epoch compared to interleaved repetition in the pre-adaptation epoch, particularly in the later phase, as depicted in Figure 5.3 C. This is expected, because, during the later phase of the pre-adaptation epoch, the MAP estimate closely aligns with the stimulus occurrence frequency within the mini-block and will thus not learn much from the new observations. However, at the onset of the adaptation epoch, the stimulus occurrence probability abruptly rises to its maximum value of 1 and will consequently drive the MAP estimate to increase based on these new observations.

To test this prediction, the rate of gamma power change in the last five mini-blocks in the pre-adaptation epoch was calculated. This rate of change was then compared to the rate of gamma power change observed during the adaptation period (Figure 5.3 B). As anticipated, both rates were positive, with the adaptation period showing a significantly faster power increment ($Slope_{pre\text{-}adapt} = .007$, $CI_{95\%} = [.003, .010]$; $Slope_{adapt} = .018$, $CI_{95\%} = [.011, .025]$; permutation test, $p = .045$) When calculating the rate of gamma power change in the pre-adaptation epoch, data from all orientations are included to increase the statistical power. This can be done because each stimulus orientation has an equal probability of occurrence in every mini-block before the adaptation epoch, and the sequence of presentations was fully randomized such that the rate of gamma-power change for different orientations are expected to be the same.

Notably, there is a decrement of gamma power in the first trial bin of the adaptation epoch compared to the last bin of the pre-adaptation epoch. This is likely due to the short, approximately 20-minute break taken by the monkey between these two epochs. During this break, no grating stimulus of any orientation was presented. If the gamma response to the adaptor orientation is of concern, having a break is thus analogous to having a series of non-adaptor orientation trials. Importantly, despite of the break, gamma power did not revert to the baseline level at the session start, indicating a persistence of the gamma power increment. Such persistency has been previously reported in both monkeys and human participants (See also Prediction VII) [61, 62]. Further exploration of this observation is done in Prediction VIII, where I attempt to model gamma power changes across all three epochs using a unified model.

5.3.5 Prediction VII: The Persistency of Gamma Power Increment

Prediction VII posits that the gamma power increment observed in one epoch would be partially maintained in subsequent epochs. In the proposed model, this persistency stems from the cumulative impact of past observations on the current inferences. In the preceding section, it was noted that gamma power for the adaptor orientation did not revert to baseline levels following a short break, suggesting that the increment in gamma power is maintained over this period. I next extend this observation to test whether the gamma power increment persists for the non-adaptor orientations even after the $\approx$ 20 min break plus the 90 repetitions of the adaptor orientation.

For the non-adaptor orientation, the proposed model predicts a partial retention of the gamma power increment at the onset of the post-adaptation epoch. Moreover, it predicts a slower rate of gamma power increase in the post-adaptation epoch compared to the pre-adaptation epoch. Specifically, when fitting a linear model to the gamma power trajectory in each epoch, the proposed model predicts a higher intercept but a lower slope for the post-adaptation epoch in non-adaptor orientations (Figure 5.3 C). The observed data align with this prediction, with a significantly larger intercept observed in the post-adaptation epoch ($Intercept_{pre\text{-}adapt} = 1.38, CI_{95\%} = [-1.57, -1.17]$; $Intercept_{post\text{-}adapt} = -.752, CI_{95\%} = [-.984, -.534]$; $p = .004$). However, the slope difference between these epochs is marginal and not statistically significant. Several factors may explain why the expected change in slope was not observed. Firstly, the initial prediction did not account for the influence of the break, which may have diminished the expected slope difference. Secondly, as the slope is predicted to decrease within the pre-adaptation epoch (Figure 5.3 A), calculating the slope over the entire epoch may underestimate the slope difference.

5.3.6 Prediction VIII: Gamma-Band LFP Power Correlates with MAP Estimate

Prediction VIII posits that the MAP estimator ($\hat{\Theta}_{MAP}$), derived from the updated prior distribution, can quantitatively predict variations in gamma power throughout a session. If this is true, with appropriate α and β values and given the trial sequence, the proposed model should predict not only the

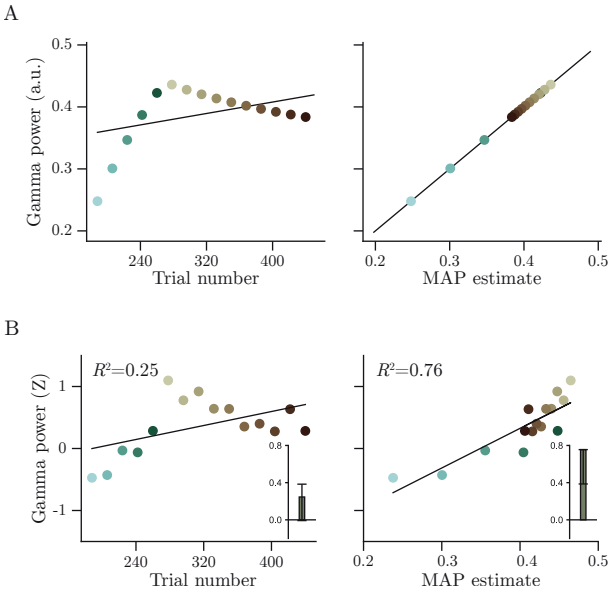


Figure 5.4: Prediction VIII (Part 1/2): Correlation between MAP estimates and gamma-band LFP power during adaptation and post-adaptation epochs. (A) Initially, a non-linear relationship is observed between gamma-band LFP power and trial number (left). This relationship is expected to linearize when gamma-band LFP power is plotted against the MAP estimate for stimulus occurrence frequency in each trial (right), highlighting the influence of neural adaptation. (B) The observed gamma-band LFP power is shown as a function of trial number (left) and MAP estimate (right) in each trial. Insets illustrate the bootstrap confidence intervals of the explained variance, with models fitted to the observed data based on the least square error criterion

trends (Predictions I to IV) but also the non-linear aspects (Predictions V to VII) of gamma power dynamics throughout the experimental session.

I first focus on how gamma-band LFP power, induced by the adaptor orientation, fluctuates over the adaptation and post-adaptation epochs. Ignoring the pre-adaptation epoch (for now) circumvents complexities introduced by the short break between the pre-adaptation and adaptation epochs. If gamma-band LFP power indeed tracks prior updating, then the complex dynamics of gamma-power throughout the session should transform into a more predictable, linear trajectory when plotted against the MAP estimate obtained for each trial, as depicted in Figure 5.4 A.

To assign a MAP estimate to each trial, it is necessary to determine the

appropriate values for α and β at the onset of the adaptation epoch. According to the proposed model, these parameters reflect the counts of adaptor orientation trials ($\pm$20 degrees) and non-adaptor orientation trials, respectively, up to the trial of interest. For simplification, α is set to one at the start of the pre-adaptation epoch, and let β be the only free parameter. This configuration assumes that the brain expects low occurrence frequency for any specific grating stimulus at the start of the session. Further, we assumed that the break after the pre-adaptation epoch primarily affects β (See discussions in Prediction VI). With such setup, we just need to find a suitable β value at the start of the adaptation epoch. With a one-dimensional parameter search maximizing the correlation between gamma-band LFP power and the resulting MAP estimates, the best β was found to be 207 at the adaptation epoch's onset. With β equals 207 at the start of the adaptation epoch and α equals one at the start of the pre-adaptation epoch, the resulting correlation between gamma power and MAP estimates across the adaptation and post-adaptation epochs ($R^2_{MAP} = .76$, $CI_{95\%} = [.39, .75]$) is significantly better than a correlation obtained through a simple linear regression on trial number ($R^2_{Lin} = .25$, $CI_{95\%} = [-.01, .38]$).

To expand our model to the pre-adaptation epoch, it is necessary to account for the break after the pre-adaptation epoch. The fact that we observed a drop in gamma power after the break (Prediction VI) suggests that we can model this interval as k trials of non-adaptor trials. Utilizing the previously determined parameters ($\alpha = 1$ at the beginning of pre-adaptation; $\beta = 207$ at the beginning of adaptation epoch), the optimal value of k, that best explains the gamma power changes in the pre-adaptation epoch, is found to be 56. This parameter set ($\alpha = 1$ at the beginning of the session; $\beta = 207$ at the beginning of the adaptation epoch; $k = 56$) led to an overall R^2 of 0.88 for predicting adaptor-induced gamma power across all epochs. The dashed line in Figure 5.5 B depicts the predicted gamma power as a function of the trial number. The predicted gamma power followed the trajectory of actual gamma power reasonably well.

5.4 Discussion

In this chapter, I analyzed local field potentials (LFP) recorded from the primary visual cortex of an awake macaque monkey (*Macaca mulatta*) during a passive fixation task. As detailed in Psarou et al. [316], stimulus-induced gamma-band

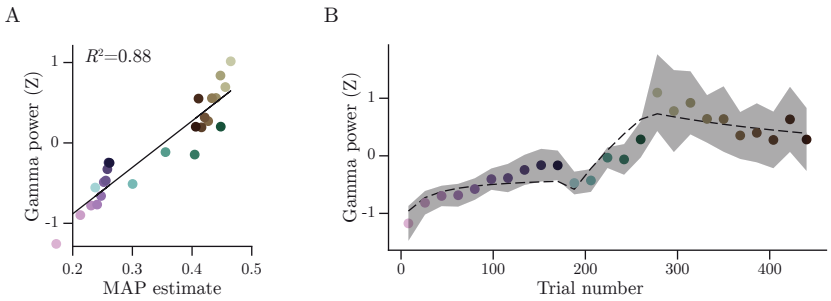


Figure 5.5: Prediction VIII (Part 2/2): Correlation between MAP estimates and gamma-band LFP power across all epochs. (A) Observed gamma power is plotted against the MAP estimate for each trial bin. (B) The fitted gamma power in (A) is plotted as a dashed line, overlaying the observed gamma power as a function of the trial number

LFP power shows characteristic dynamics throughout the session. To explain these dynamics, I propose that the gamma-band neuronal synchronization may track the updates of the inferred stimulus occurrence frequency according to Bayes' rule. I made and tested eight predictions ranging from qualitative to more quantitative ones. The results supported the hypothesis, and the observed dynamics of gamma-band neuronal synchronization can largely be accounted for by how the prior distribution over stimulus occurrence frequency was updated as the session progressed. Several studies have documented increased gamma response with consecutive and interleaved stimulus repetition [61, 62, 257]. Such repetition effect can be explained by the proposed model (Prediction I and II). Importantly, the proposed model predicts a reduction in the gamma power increment rate as the stimulus is repeated (Prediction V). In line with Prediction V, Brunet et al. [257]and Peter et al. [61] reported that gamma power increased logarithmically with the number of repetitions, also implying a decreasing increment speed. While the logarithmic transformation in these studies was empirically derived, the proposed model offers quantitative description about the rate of change. Notably, within certain parameter range, the logarithmic transformation approximates the rate of change our model predicts (Figure 5.3 A).

The repetition effect has been shown to be stimulus-specific [61, 62, 316]. Stimulus specificity is a built-in feature of the proposed model because the stimulus occurrence frequency is specific to the stimulus of interest. I have demonstrated such stimulus specificity by showing opposite trends of gamma-

power change for the adaptor and non-adaptor orientation in the post-adaptation epoch. The model does however not specify about the extent of the stimulus specificity. The current analysis combined gamma-response from gratings less then 20 degrees away from the adaptor orientation in the pre-adaptation epoch, where all 18 orientations are randomly mixed up. Given the width of orientation tuning for V1 neurons [331] this may be a reasonable choice. However, the decreasing trend in the post-adaptation epoch seems confined to only the adaptor orientation, not even the neighbouring orientations ($\pm$10 degrees). Such high stimulus specificity after adaptation may be caused by a more accurate representation of adaptor orientation as the consequence of adaptation [332–335]. Importantly, the width of binning appears to be not critical for our general findings. Less or more bining can still provide a good fit, albeit with different model parameters (β and k).

The predicted persistency of gamma power increment (Prediction VII) has also been corroborated by previous research. Stauch et al. [62] observed in human participants that gamma-power enhancement partially persisted when the inducing stimulus was repeated after more than 25 minutes of intervening presentation of other orientations. Similarly, Peter et al. [61] reported persistency in gamma power in macaque V1, both with grating stimuli and natural images. Notably, they also documented a renormalization of gamma power during the rest, akin to what we observed between the pre-adaptation and adaptation epochs (Figure 5.3 F; Figure S6 E in Peter et al. [61]). An intriguing question regarding the persistency of repetition-related change is its timescale: How long does it persist, and what is the precise temporal profile of its decay? Answering these questions is challenging since any attempt to probe the (de-)adaptation process inherently influences it. The proposed model offers insights into these latent processes. For example, the proposed model predicts a partial renormalization, observable if the adaptor orientation is presented less frequently than expected. This is seen in the post-adaptation epoch: adaptor-induced gamma decreased towards the baseline.

Integrating prior knowledge with sensory evidence is fundamental for visual perception within the predictive coding framework [265, 325]. This chapter notices that the strength of gamma-synchronization, as reflected in LFP power, tracks the prior updates of stimulus occurrence frequency. This observation suggests the potential role of gamma-band synchronization in both updating

and computing with the prior information. For updating prior information, it has been suggested that gamma-band synchronization facilitates Hebbian plasticity by providing temporally precise spike sequences [328, 336]. What follows then is an intriguing outcome: enhanced gamma-band synchronization, resulting from learning, is itself a mechanism for facilitating further learning. For computing with the stored prior, coherence could be used. Spikes that encode the predicted stimulus, coherent with the excitability cycles of post-synaptic neurons, can exert a greater influence [69]. This process of weighting incoming spikes according to their causes' predictability effectively enacts Bayes' rule, namely, multiplying the likelihood, represented by the spiking activity, with the prior, represented by the gamma-band coherence [337].

5.5 Materials and Methods

Subject. One adult rhesus macaque (Macaque mulatta) was used. All experimental procedures were conducted in compliance with German and European animal protection laws. The experiment was approved by the responsible local authority (Regierungspräsidium Darmstadt).

Behavioral task. The monkey performed a passive fixation task. Each trial commenced with a white fixation dot presented against a gray background. After acquiring fixation for 0.8 $\sim$ 0.9 second, an oriented grating stimulus was presented for 0.9$\sim$1.0 second, during which the monkey needed to continue fixation. Successfully maintaining fixation would result in a juice reward, indicating a correct trial. Each trial was followed by an inter-trial interval of 0.5 seconds. The task was structured into three epochs. In the first epoch, termed the pre-adaptation epoch, there were ten mini-blocks. Each mini-block comprised 18 trials, displaying each of the 18 unique grating stimuli (10 to 180 degrees, in 10-degree increments) once. The sequence of orientations was randomized in each mini-block. If the monkey broke fixation during a trial, the current and remaining orientations in the mini-block were reshuffled. After a 20-minute break, the task progressed to the adaptation epoch. This epoch presented 90 trials of a single orientation (the adaptor grating), randomly selected from the 18 orientations for each session. The final phase, the post-adaptation epoch, immediately followed the adaptation epoch without a break and mirrored the pre-adaptation epoch's trial-structure. At least ten mini-blocks were collected for the post-adaptation epoch.

Electrophysiological recording. On each recording day, an acute laminar recording was obtained from a single penetration perpendicular to V1 using V-probes (Plexon Inc, Dallas, Texas, US). Electrophysiological signals were recorded with Tucker Davis Technologies systems (TDT, Alachua, Florida, United States). The raw data were recorded at 24414.0625 Hz. For the analysis of Local Field Potentials (LFP), this raw data was then low-pass filtered and down-sampled to 1000 Hz offline. Spectral analysis Only correct trials entered the analysis pipeline. We first calculated bipolar-derived LFP (bLFP) by re-referenceing the signal of each channel to the immediate lower channel. From the bLFP signal, we first calculated the bLFP power spectrum during stimulus presentation (270 to 770 ms after stimulus onset) using Welch's methods (five overlapping 300 ms epochs, linear-detrending, Hann tapering, and zero-padding to 1 s). This stimulus-period power for each trial was then subtracted by the average baseline power per session (estimated from 530 to 30 ms before stimulus onset) to get the baseline corrected power spectra. Gamma-band LFP power is calculated using the frequency band 33-49 Hz, which corresponds to $\approx 30\%$ power reduction around the peak frequency.

Statistical analysis Confidence intervals are based on bootstrap procedure over sessions ($N = 1000$). Hypothesis tests in Figures 5.2 and 5.3 are based on a permutation test permuting the mini-block order within each session.

CHAPTER 6

General Discussion

There is a discrepancy between how vision is studied in the lab and how vision is used in real life. In real life, an observer actively obtains information about their surroundings by shifting their gaze, turning their head, and changing their point of observation. By contrast, a participant in the lab usually passively receives stimulation that impinges on their retina. This is especially true when neuroscientists study the neural correlates of visual perception using macaque monkeys. More often than not, macaque monkeys are found in the lab head-posted, looking fiercely at an uninteresting fixation dot. This is not the type of vision we aim to understand.

Following James J. Gibson, ecological psychologists have long recognized this discrepancy and suggested adopting an ecological approach to address it. Taking an ecological approach means considering the information available to the active observer and the process by which the active observer picks up this information. The work presented in this thesis advocates for a similar ecological approach in neuroscience.

The observer's visual perceptual system is attuned to information that is meaningful to him. Such information can appear in various forms, including those specified by two-dimensional images. Stimulation that specifies the same information is inherently linked. For example, distance and depth information can be specified through binocular disparity, texture density, or occluding edges. Establishing this link is crucial as it opens the door to understanding natural vision by studying two-dimensional images. Two-dimensional images are considered ecological stimulation if the information they specify is natural.

Notably, ecological stimulation does not need to be realistic. In fact, because the real environment is information-rich, identifying relevant information necessarily requires stripping away the irrelevant. Chapter 2 reports on how attentional cues can be adapted to maintain information about overlapping surfaces while gradually removing other aspects of the cue.

Meaningful information for the observer can be statistically defined. For example, resource richness can be measured by fruit density in forest canopies or insect abundance on tree bark. For an active observer that samples the environment, the density or abundance of food translates into the frequency of stimulus occurrence. Chapter 5 explores how the brain detects information that specifies the stimulus occurrence frequency. A notable drawback of the methods applied in Chapter 5 is that, unlike in real life, the transformation of spatial regularity into temporal regularity is artificially imposed by the experimental trial structure. This problem arises from neglecting how observers pick up information in the real world. Addressing this issue calls for new behavioral tasks that allow the observers to explore their environment naturally. The controlled foraging task presented in Chapter 3 represents progress in this direction.

Understanding the neural mechanisms supporting the observer's information pickup is the main goal of the envisaged ecological neuroscience. For this, it is important to identify the changes and persistence in neuronal activity as the observer actively explores his surroundings. This thesis suggests three mechanisms by which environmental persistence may be registered in the brain during frequent gaze shifts. The observer's active exploration in real life involves more than just saccades. It also includes head movements, locomotion, and object manipulation for inspection. The current thesis did not explore these aspects or their interactions.

The development of virtual, augmented, and mixed reality techniques is making it possible to engineer the world that surrounds the observer directly [338]. Better tools are taking behavioral analysis to an unprecedented level [339]. Large-scale recordings simultaneously from multiple brain areas and thousands of neurons are becoming possible [340]. Consciously or not, the field is moving towards ecological neuroscience. The goal of this thesis is to contribute to this exciting momentum.

APPENDIX A

Data Management

Ethical Approval

All experimental procedures, including surgical interventions and animal housing, were conducted in full compliance with the European Union Directive 2010/63/EU on the protection of animals used for scientific purposes, as well as the relevant German regulations. Studies contained in this thesis were approved by the responsible local authority, the Regierungspräsidium Darmstadt.

Findability and Accessibility

All data presented in this thesis were collected at the Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with the Max Planck Society in Frankfurt am Main, Germany. All data and relevant metadata presented in this thesis have been archived on a private server located at ESI. Requests for access to the data presented in this thesis can be directed to Pascal Fries. The data analysis scripts, which were developed using MATLAB and Python, are available upon

request. For access to the corresponding analysis scripts, please contact Yufeng Zhang.

APPENDIX B

Curriculum Vitae

Yufeng Zhang was born on January 28, 1992, in Shanxi, China. He graduated with a bachelor's degree in Biological Sciences from Zhejiang University, China, in 2014. Following his undergraduate studies, he worked for two years as a research assistant in Prof. Anna W. Roe's lab. In 2016, he was admitted to the International Max Planck Research School (IMPRS) for Neural Circuits, where he soon began his doctoral studies under the supervision of Prof. Pascal Fries.

APPENDIX C

List of Publications

1. Yufeng Zhang and Pascal Fries. "Eccentricity-Dependent Saccadic Reaction Time: The Roles of Foveal Magnification and Attentional Orienting." In: *bioRxiv* (Aug. 2023). DOI: http://doi.org/10.1101/2023.08.08.552339

2. Tim Näher*, Yufeng Zhang*, Martina Pandinelli, and Pascal Fries. "Primate Saccade Rhythmicity." In: *bioRxiv* (June. 2024). DOI: http://doi.org/10.1101/2023.09.27.559710. (*Contributed equally)

3. Eleni Psarou, Julien Vezoli, Marieke L Schölvinck, Pierre-Antoine Ferracci, Yufeng Zhang, Iris Grothe, Rasmus Roese and Pascal Fries. "Modular, cement-free, customized headpost and connector-chamber implants for macaques". In: *Journal of Neuroscience Methods* 393 (June 2023), p. 109899. ISSN: 1872-678X. DOI: https://doi.org/10.1016/j.jneumeth.2023.109899.

APPENDIX D

Donders Graduate School

For a successful research Institute, it is vital to train the next generation of scientists. To achieve this goal, the Donders Institute for Brain, Cognition and Behaviour established the Donders Graduate School in 2009. The mission of the Donders Graduate School is to guide our graduates to become skilled academics who are equipped for a wide range of professions. To achieve this, we do our utmost to ensure that our PhD candidates receive support and supervision of the highest quality.

Since 2009, the Donders Graduate School has grown into a vibrant community of highly talented national and international PhD candidates, with over 500 PhD candidates enrolled. Their backgrounds cover a wide range of disciplines, from physics to psychology, medicine to psycholinguistics, and biology to artificial intelligence. Similarly, their interdisciplinary research covers genetic, molecular, and cellular processes at one end and computational, system-level neuroscience with cognitive and behavioural analysis at the other end. We ask all PhD candidates within the Donders Graduate School to publish their PhD thesis in

de Donders Thesis Series. This series currently includes over 600 PhD theses from our PhD graduates and thereby provides a comprehensive overview of the diverse types of research performed at the Donders Institute. A complete overview of the Donders Thesis Series can be found on our website: `https://www.ru.nl/donders/donders-series`

The Donders Graduate School tracks the careers of our PhD graduates carefully. In general, the PhD graduates end up at high-quality positions in different sectors, for a complete overview see `https://www.ru.nl/donders/destination-our-former-phd`. A large proportion of our PhD alumni continue in academia ($> 50\%$). Most of them first work as a postdoc before growing into more senior research positions. They work at top institutes worldwide, such as University of Oxford, University of Cambridge, Stanford University, Princeton University, UCL London, MPI Leipzig, Karolinska Institute, UC Berkeley, EPFL Lausanne, and many others. In addition, a large group of PhD graduates continue in clinical positions, sometimes combining it with academic research. Clinical positions can be divided into medical doctors, for instance, in genetics, geriatrics, psychiatry, or neurology, and in psychologists, for instance as healthcare psychologist, clinical neuropsychologist, or clinical psychologist. Furthermore, there are PhD graduates who continue to work as researchers outside academia, for instance at non-profit or government organizations, or in pharmaceutical companies. There are also PhD graduates who work in education, such as teachers in high school, or as lecturers in higher education. Others continue in a wide range of positions, such as policy advisors, project managers, consultants, data scientists, web- or software developers, business owners, regulatory affairs specialists, engineers, managers, or IT architects. As such, the career paths of Donders PhD graduates span a broad range of sectors and professions, but the common factor is that they almost all have become successful professionals. For more information on the Donders Graduate School, as well as past and upcoming defences please visit: `http://www.ru.nl/donders/graduate-school/phd/`

Bibliography

[1] Davida Y. Teller and John Palmer. *Linking Vision and the Visual System*. Sept. 2018. URL: https://faculty.washington.edu/jpalmer/files/Teller/TellerBook2018/.

[2] D. Y. Teller. "Linking propositions". eng. In: *Vision Research* 24.10 (1984), pp. 1233–1246. ISSN: 0042-6989. DOI: 10.1016/0042-6989(84)90178-0.

[3] James J. Gibson. *The Ecological Approach to Visual Perception: Classic Edition*. en. 1st ed. Psychology Press, Nov. 2014. ISBN: 978-1-315-74021-8. DOI: 10.4324/9781315740218. URL: https://www.taylorfrancis.com/books/9781315740218 (visited on 12/17/2023).

[4] Edward G. Jones. "Neuroanatomy: Cajal and after Cajal". eng. In: *Brain Research Reviews* 55.2 (Oct. 2007), pp. 248–255. ISSN: 0165-0173. DOI: 10.1016/j.brainresrev.2007.06.001.

[5] J. S. Lund. "Organization of neurons in the visual cortex, area 17, of the monkey (Macaca mulatta)". eng. In: *The Journal of Comparative Neurology* 147.4 (Feb. 1973), pp. 455–496. ISSN: 0021-9967. DOI: 10.1002/cne.901470404.

[6] Hongkui Zeng and Joshua R. Sanes. "Neuronal cell-type classification: challenges, opportunities and the path forward". eng. In: *Nature Reviews.*

Neuroscience 18.9 (Sept. 2017), pp. 530–546. ISSN: 1471-0048. DOI: 10.1038/nrn.2017.85.

[7] W. Waldeyer. "Ueber einige neuere Forschungen im Gebiete der Anatomie des Centralnervensystems". de. In: *DMW - Deutsche Medizinische Wochenschrift* 17.44 (Nov. 1891), pp. 1213–1218. ISSN: 0012-0472, 1439-4413. DOI: 10.1055/s-0029-1206824. URL: http://www.thieme-connect.de/DOI/DOI?10.1055/s-0029-1206824 (visited on 07/27/2024).

[8] Rafael Yuste. "From the neuron doctrine to neural networks". eng. In: *Nature Reviews. Neuroscience* 16.8 (Aug. 2015), pp. 487–497. ISSN: 1471-0048. DOI: 10.1038/nrn3962.

[9] Nikola T. Markov et al. "Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex". eng. In: *The Journal of Comparative Neurology* 522.1 (Jan. 2014), pp. 225–259. ISSN: 1096-9861. DOI: 10.1002/cne.23458.

[10] Ehud Kaplan. "The M, P, and K Pathways of the Primate Visual System". en. In: *The Visual Neurosciences, 2-vol. set.* Ed. by Leo M. Chalupa and John S. Werner. The MIT Press, Nov. 2003, pp. 481–493. ISBN: 978-0-262-27012-0. DOI: 10.7551/mitpress/7131.003.0036. URL: https://direct.mit.edu/books/book/5395/chapter/3948190/The-M-P-and-K-Pathways-of-the-Primate-Visual (visited on 07/28/2024).

[11] D. H. Hubel and T. N. Wiesel. "Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey". eng. In: *The Journal of Comparative Neurology* 146.4 (Dec. 1972), pp. 421–450. ISSN: 0021-9967. DOI: 10.1002/cne.901460402.

[12] S. H. Hendry and T. Yoshioka. "A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus". eng. In: *Science (New York, N.Y.)* 264.5158 (Apr. 1994), pp. 575–577. ISSN: 0036-8075. DOI: 10.1126/science.8160015.

[13] Mortimer Mishkin, Leslie G. Ungerleider, and Kathleen A. Macko. "Object vision and spatial vision: two cortical pathways". en. In: *Trends in Neurosciences* 6 (Jan. 1983), pp. 414–417. ISSN: 01662236. DOI: 10.1016/0166-2236(83)90190-X. URL: https://linkinghub.elsevier.com/retrieve/pii/016622368390190X (visited on 07/28/2024).

[14] K. S. Rockland and D. N. Pandya. "Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey". eng.

In: *Brain Research* 179.1 (Dec. 1979), pp. 3–20. ISSN: 0006-8993. DOI: 10.1016/0006-8993(79)90485-2.

[15] D. J. Felleman and D. C. Van Essen. "Distributed hierarchical processing in the primate cerebral cortex". eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 1.1 (1991), pp. 1–47. ISSN: 1047-3211. DOI: 10.1093/cercor/1.1.1-a.

[16] C. C. Hilgetag, M. A. O'Neill, and M. P. Young. "Indeterminate organization of the visual system". eng. In: *Science (New York, N.Y.)* 271.5250 (Feb. 1996), pp. 776–777. ISSN: 0036-8075. DOI: 10.1126/science.271.5250.776.

[17] S. Zeki and S. Shipp. "The functional logic of cortical connections". eng. In: *Nature* 335.6188 (Sept. 1988), pp. 311–317. ISSN: 0028-0836. DOI: 10.1038/335311a0.

[18] D. H. Hubel and T. N. Wiesel. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". eng. In: *The Journal of Physiology* 160.1 (Jan. 1962), pp. 106–154. ISSN: 0022-3751. DOI: 10.1113/jphysiol.1962.sp006837.

[19] R. C. Reid and J. M. Alonso. "Specificity of monosynaptic connections from thalamus to visual cortex". eng. In: *Nature* 378.6554 (Nov. 1995), pp. 281–284. ISSN: 0028-0836. DOI: 10.1038/378281a0.

[20] Nicholas J. Priebe. "Mechanisms of Orientation Selectivity in the Primary Visual Cortex". eng. In: *Annual Review of Vision Science* 2 (Oct. 2016), pp. 85–107. ISSN: 2374-4650. DOI: 10.1146/annurev-vision-111815-114456.

[21] Anna W. Roe et al. "Toward a unified theory of visual area V4". eng. In: *Neuron* 74.1 (Apr. 2012), pp. 12–29. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.03.011.

[22] Richard T. Born and David C. Bradley. "Structure and function of visual area MT". eng. In: *Annual Review of Neuroscience* 28 (2005), pp. 157–189. ISSN: 0147-006X. DOI: 10.1146/annurev.neuro.26.041002.131052.

[23] N. K. Logothetis, J. Pauls, and T. Poggio. "Shape representation in the inferior temporal cortex of monkeys". eng. In: *Current biology: CB* 5.5 (May 1995), pp. 552–563. ISSN: 0960-9822. DOI: 10.1016/s0960-9822(95)00108-4.

[24] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". eng. In: *Nature* 521.7553 (May 2015), pp. 436–444. ISSN: 1476-4687. DOI: 10.1038/nature14539.

[25] Caitlin Siu et al. “A direct interareal feedback-to-feedforward circuit in primate visual cortex”. eng. In: *Nature Communications* 12.1 (Aug. 2021), p. 4911. ISSN: 2041-1723. DOI: 10.1038/s41467-021-24928-6.

[26] N. T. Markov et al. “A weighted and directed interareal connectivity matrix for macaque cerebral cortex”. eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 24.1 (Jan. 2014), pp. 17–36. ISSN: 1460-2199. DOI: 10.1093/cercor/bhs270.

[27] Jean Bullier. “What Is Fed Back?” en. In: *23 Problems in Systems Neuroscience*. Ed. by J. Leo Van Hemmen and Terrence J. Sejnowski. 1st ed. Oxford University PressNew York, Jan. 2006, pp. 103–132. ISBN: 978-0-19-514822-0 978-0-19-986467-6. DOI: 10.1093/acprof:oso/9780195148220.003.0006. URL: https://academic.oup.com/book/4769/chapter/147024008 (visited on 07/28/2024).

[28] Tirin Moore and Marc Zirnsak. “Neural Mechanisms of Selective Visual Attention”. eng. In: *Annual Review of Psychology* 68 (Jan. 2017), pp. 47–72. ISSN: 1545-2085. DOI: 10.1146/annurev-psych-122414-033400.

[29] Eric R. Kandel et al. “The Constructive Nature of Visual Processing”. In: *Principles of Neural Science, 6e*. New York, NY: McGraw Hill, 2021. URL: neurology.mhmedical.com/content.aspx?aid=1180641791 (visited on 07/28/2024).

[30] Stewart Shipp. “Computational components of visual predictive coding circuitry”. eng. In: *Frontiers in Neural Circuits* 17 (2023), p. 1254009. ISSN: 1662-5110. DOI: 10.3389/fncir.2023.1254009.

[31] Andre M. Bastos et al. “Canonical microcircuits for predictive coding”. eng. In: *Neuron* 76.4 (Nov. 2012), pp. 695–711. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.10.038.

[32] André Moraes Bastos et al. “Visual areas exert feedforward and feedback influences through distinct frequency channels”. eng. In: *Neuron* 85.2 (Jan. 2015), pp. 390–401. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2014.12.018.

[33] D. H. Hubel. “Tungsten Microelectrode for Recording from Single Units”. eng. In: *Science (New York, N.Y.)* 125.3247 (Mar. 1957), pp. 549–550. ISSN: 0036-8075. DOI: 10.1126/science.125.3247.549.

[34] D. H. Hubel and T. N. Wiesel. “Receptive fields and functional architecture of monkey striate cortex”. eng. In: *The Journal of Physiology* 195.1 (Mar. 1968), pp. 215–243. ISSN: 0022-3751. DOI: 10.1113/jphysiol.1968.sp008455.

[35] V. Mountcastle. *Topographic organization and modality representation in first somatic area of cat's cerebral cortex by method of single unit analysis*. Pages: 646 Publication Title: American Journal of Physiology Volume: 183. 1955. URL: https://cir.nii.ac.jp/crid/1370285705048729476.

[36] G. G. Blasdel and G. Salama. "Voltage-sensitive dyes reveal a modular organization in monkey striate cortex". eng. In: *Nature* 321.6070 (June 1986), pp. 579–585. ISSN: 0028-0836. DOI: 10.1038/321579a0.

[37] L. C. Sincich and G. G. Blasdel. "Oriented axon projections in primary visual cortex of the monkey". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 21.12 (June 2001), pp. 4416–4426. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.21-12-04416.2001.

[38] Jennifer S. Lund, Alessandra Angelucci, and Paul C. Bressloff. "Anatomical substrates for functional columns in macaque monkey primary visual cortex". eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 13.1 (Jan. 2003), pp. 15–24. ISSN: 1047-3211. DOI: 10.1093/cercor/13.1.15.

[39] N. V. Swindale et al. "Visual cortex maps are optimized for uniform coverage". eng. In: *Nature Neuroscience* 3.8 (Aug. 2000), pp. 822–826. ISSN: 1097-6256. DOI: 10.1038/77731.

[40] D. H. Hubel and M. S. Livingstone. "Segregation of form, color, and stereopsis in primate area 18". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 7.11 (Nov. 1987), pp. 3378–3415. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.07-11-03378.1987.

[41] Haidong D. Lu et al. "A motion direction map in macaque V2". eng. In: *Neuron* 68.5 (Dec. 2010), pp. 1002–1013. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2010.11.020.

[42] Hisashi Tanigawa, Haidong D. Lu, and Anna W. Roe. "Functional organization for color and orientation in macaque V4". eng. In: *Nature Neuroscience* 13.12 (Dec. 2010), pp. 1542–1548. ISSN: 1546-1726. DOI: 10.1038/nn.2676.

[43] Rendong Tang et al. "Curvature-processing domains in primate V4". eng. In: *eLife* 9 (Nov. 2020), e57502. ISSN: 2050-084X. DOI: 10.7554/eLife.57502.

[44] T. D. Albright. "Direction and orientation selectivity of neurons in visual area MT of the macaque". eng. In: *Journal of Neurophysiology* 52.6 (Dec.

1984), pp. 1106–1130. ISSN: 0022-3077. DOI: 10.1152/jn.1984.52.6.1106.

[45] G. C. DeAngelis and W. T. Newsome. "Organization of disparity-selective neurons in macaque area MT". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 19.4 (Feb. 1999), pp. 1398–1415. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.19-04-01398.1999.

[46] Pinglei Bao et al. "A map of object space in primate inferotemporal cortex". eng. In: *Nature* 583.7814 (July 2020), pp. 103–108. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2350-5.

[47] Jakob Heinzle, Klaus Hepp, and Kevan A. C. Martin. "A microcircuit model of the frontal eye fields". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27.35 (Aug. 2007), pp. 9341–9353. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0974-07.2007.

[48] C. J. Bruce et al. "Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements". eng. In: *Journal of Neurophysiology* 54.3 (Sept. 1985), pp. 714–734. ISSN: 0022-3077. DOI: 10.1152/jn.1985.54.3.714.

[49] M. T. Schmolesky et al. "Signal timing across the macaque visual system". eng. In: *Journal of Neurophysiology* 79.6 (June 1998), pp. 3272–3278. ISSN: 0022-3077. DOI: 10.1152/jn.1998.79.6.3272.

[50] Yu Chen et al. "Population burst propagation across interacting areas of the brain". eng. In: *Journal of Neurophysiology* 128.6 (Dec. 2022), pp. 1578–1592. ISSN: 1522-1598. DOI: 10.1152/jn.00066.2022.

[51] Thomas Zhihao Luo and John H. R. Maunsell. "Neuronal Modulations in Visual Cortex Are Associated with Only One of Multiple Components of Attention". eng. In: *Neuron* 86.5 (June 2015), pp. 1182–1188. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2015.05.007.

[52] Jasper Poort et al. "The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex". eng. In: *Neuron* 75.1 (July 2012), pp. 143–156. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.04.032.

[53] Yin Yan, Li Zhaoping, and Wu Li. "Bottom-up saliency and top-down learning in the primary visual cortex of monkeys". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 115.41 (Oct. 2018), pp. 10499–10504. ISSN: 1091-6490. DOI: 10.1073/pnas.1803854115.

[54] D. J. Heeger. "Normalization of cell responses in cat striate cortex". eng. In: *Visual Neuroscience* 9.2 (Aug. 1992), pp. 181–197. ISSN: 0952-5238. DOI: 10.1017/s0952523800009640.

[55] Jingyang Zhou et al. "Predicting neuronal dynamics with a delayed gain control model". eng. In: *PLoS computational biology* 15.11 (Nov. 2019), e1007484. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1007484.

[56] Beren Millidge, Anil Seth, and Christopher L Buckley. *Predictive Coding: a Theoretical and Experimental Review*. Version Number: 4. 2021. DOI: 10.48550/ARXIV.2107.12979. URL: https://arxiv.org/abs/2107.12979 (visited on 07/28/2024).

[57] Kalanit Grill-Spector, Richard Henson, and Alex Martin. "Repetition and the brain: neural models of stimulus-specific effects". en. In: *Trends in Cognitive Sciences* 10.1 (Jan. 2006), pp. 14–23. ISSN: 13646613. DOI: 10.1016/j.tics.2005.11.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S1364661305003232 (visited on 12/16/2023).

[58] J. R. Müller et al. "Rapid adaptation in visual cortex to the structure of images". eng. In: *Science (New York, N.Y.)* 285.5432 (Aug. 1999), pp. 1405–1408. ISSN: 0036-8075. DOI: 10.1126/science.285.5432.1405.

[59] Iris I. A. Groen et al. "Temporal Dynamics of Neural Responses in Human Visual Cortex". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 42.40 (Oct. 2022), pp. 7562–7580. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.1812-21.2022.

[60] L. Li, E. K. Miller, and R. Desimone. "The representation of stimulus familiarity in anterior inferior temporal cortex". en. In: *Journal of Neurophysiology* 69.6 (June 1993), pp. 1918–1929. ISSN: 0022-3077, 1522-1598. DOI: 10.1152/jn.1993.69.6.1918. URL: https://www.physiology.org/doi/10.1152/jn.1993.69.6.1918 (visited on 12/16/2023).

[61] Alina Peter et al. "Stimulus-specific plasticity of macaque V1 spike rates and gamma". eng. In: *Cell Reports* 37.10 (Dec. 2021), p. 110086. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2021.110086.

[62] Benjamin J Stauch et al. "Stimulus-specific plasticity in human visual gamma-band activity and functional connectivity". en. In: *eLife* 10 (Aug. 2021), e68240. ISSN: 2050-084X. DOI: 10.7554/eLife.68240. URL: https://elifesciences.org/articles/68240 (visited on 12/16/2023).

[63] M. N. Shadlen and W. T. Newsome. "The variable discharge of cortical neurons: implications for connectivity, computation, and information coding". eng. In: *The Journal of Neuroscience: The Official Journal*

of the Society for Neuroscience 18.10 (May 1998), pp. 3870–3896. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.18-10-03870.1998.

[64] Z. F. Mainen and T. J. Sejnowski. "Reliability of spike timing in neocortical neurons". eng. In: *Science (New York, N.Y.)* 268.5216 (June 1995), pp. 1503–1506. ISSN: 0036-8075. DOI: 10.1126/science.7770778.

[65] Peter Cariani and Janet M. Baker. "Time Is of the Essence: Neural Codes, Synchronies, Oscillations, Architectures". eng. In: *Frontiers in Computational Neuroscience* 16 (2022), p. 898829. ISSN: 1662-5188. DOI: 10.3389/fncom.2022.898829.

[66] C. Koch, M. Rapp, and I. Segev. "A brief history of time (constants)". eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 6.2 (1996), pp. 93–101. ISSN: 1047-3211. DOI: 10.1093/cercor/6.2.93.

[67] C. M. Gray et al. "Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties". eng. In: *Nature* 338.6213 (Mar. 1989), pp. 334–337. ISSN: 0028-0836. DOI: 10.1038/338334a0.

[68] W. Singer and C. M. Gray. "Visual feature integration and the temporal correlation hypothesis". eng. In: *Annual Review of Neuroscience* 18 (1995), pp. 555–586. ISSN: 0147-006X. DOI: 10.1146/annurev.ne.18.030195.003011.

[69] Pascal Fries. "Rhythms for Cognition: Communication through Coherence". en. In: *Neuron* 88.1 (Oct. 2015), pp. 220–235. ISSN: 08966273. DOI: 10.1016/j.neuron.2015.09.034. URL: https://linkinghub.elsevier.com/retrieve/pii/S0896627315008235 (visited on 12/16/2023).

[70] Conrado A. Bosman et al. "Attentional stimulus selection through selective synchronization between monkey visual areas". eng. In: *Neuron* 75.5 (Sept. 2012), pp. 875–888. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2012.06.037.

[71] Stefano Panzeri et al. "The structures and functions of correlations in neural population codes". eng. In: *Nature Reviews. Neuroscience* 23.9 (Sept. 2022), pp. 551–567. ISSN: 1471-0048. DOI: 10.1038/s41583-022-00606-4.

[72] Marlene R. Cohen and John H. R. Maunsell. "Attention improves performance primarily by reducing interneuronal correlations". eng. In: *Nature Neuroscience* 12.12 (Dec. 2009), pp. 1594–1600. ISSN: 1546-1726. DOI: 10.1038/nn.2439.

[73] Jude F. Mitchell, Kristy A. Sundberg, and John H. Reynolds. "Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4". eng. In: *Neuron* 63.6 (Sept. 2009), pp. 879–888. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2009.09.013.

[74] Douglas A. Ruff and Marlene R. Cohen. "Attention can either increase or decrease spike count correlations in visual cortex". eng. In: *Nature Neuroscience* 17.11 (Nov. 2014), pp. 1591–1597. ISSN: 1546-1726. DOI: 10.1038/nn.3835.

[75] Sunny Nigam, Sorin Pojoga, and Valentin Dragoi. "Synergistic Coding of Visual Information in Columnar Networks". eng. In: *Neuron* 104.2 (Oct. 2019), 402–411.e4. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2019.07.006.

[76] Thilo Womelsdorf et al. "Orientation selectivity and noise correlation in awake monkey area V1 are modulated by the gamma cycle". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 109.11 (Mar. 2012), pp. 4302–4307. ISSN: 1091-6490. DOI: 10.1073/pnas.1114223109.

[77] Alexander S. Ecker et al. "State dependence of noise correlations in macaque primary visual cortex". eng. In: *Neuron* 82.1 (Apr. 2014), pp. 235–248. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2014.02.006.

[78] Bilal Haider, Michael Häusser, and Matteo Carandini. "Inhibition dominates sensory responses in the awake cortex". eng. In: *Nature* 493.7430 (Jan. 2013), pp. 97–100. ISSN: 1476-4687. DOI: 10.1038/nature11665.

[79] György Buzsáki. *The Brain from Inside Out*. en. 1st ed. Oxford University PressNew York, June 2019. ISBN: 978-0-19-090538-5 978-0-19-090541-5. DOI: 10.1093/oso/9780190905385.001.0001. URL: https://academic.oup.com/book/35081 (visited on 07/28/2024).

[80] Marisa Carrasco. "Visual attention: the past 25 years". eng. In: *Vision Research* 51.13 (July 2011), pp. 1484–1525. ISSN: 1878-5646. DOI: 10.1016/j.visres.2011.04.012.

[81] Wayne Wu. "We know what attention is!" eng. In: *Trends in Cognitive Sciences* 28.4 (Apr. 2024), pp. 304–318. ISSN: 1879-307X. DOI: 10.1016/j.tics.2023.11.007.

[82] M. I. Posner. "Orienting of attention". eng. In: *The Quarterly Journal of Experimental Psychology* 32.1 (Feb. 1980), pp. 3–25. ISSN: 0033-555X. DOI: 10.1080/00335558008248231.

[83] Michael J. Hautus, Neil A. Macmillan, and C. Douglas Creelman. *Detection Theory: A User's Guide*. en. 3rd ed. New York: Routledge, Aug. 2021. ISBN: 978-1-00-320363-6. DOI: 10.4324/9781003203636. URL: https://www.taylorfrancis.com/books/9781003203636 (visited on 07/28/2024).

[84] Alexandre Zénon and Richard J. Krauzlis. "Attention deficits without cortical neuronal deficits". eng. In: *Nature* 489.7416 (Sept. 2012), pp. 434–437. ISSN: 1476-4687. DOI: 10.1038/nature11497.

[85] Jan-Mathijs Schoffelen, Robert Oostenveld, and Pascal Fries. "Neuronal coherence as a mechanism of effective corticospinal interaction". eng. In: *Science (New York, N.Y.)* 308.5718 (Apr. 2005), pp. 111–113. ISSN: 1095-9203. DOI: 10.1126/science.1107027.

[86] Ac Nobre, A. Correa, and Jt Coull. "The hazards of time". eng. In: *Current Opinion in Neurobiology* 17.4 (Aug. 2007), pp. 465–470. ISSN: 0959-4388. DOI: 10.1016/j.conb.2007.07.006.

[87] Geoffrey M. Ghose and John H. R. Maunsell. "Attentional modulation in visual cortex depends on task timing". eng. In: *Nature* 419.6907 (Oct. 2002), pp. 616–620. ISSN: 0028-0836. DOI: 10.1038/nature01057.

[88] Jeremy M. Wolfe. "Guided Search 6.0: An updated model of visual search". en. In: *Psychonomic Bulletin & Review* 28.4 (Aug. 2021), pp. 1060–1092. ISSN: 1069-9384, 1531-5320. DOI: 10.3758/s13423-020-01859-9. URL: https://link.springer.com/10.3758/s13423-020-01859-9 (visited on 12/15/2023).

[89] J. Moran and R. Desimone. "Selective attention gates visual processing in the extrastriate cortex". eng. In: *Science (New York, N.Y.)* 229.4715 (Aug. 1985), pp. 782–784. ISSN: 0036-8075. DOI: 10.1126/science.4023713.

[90] P. Fries et al. "Modulation of oscillatory neuronal synchronization by selective visual attention". eng. In: *Science (New York, N.Y.)* 291.5508 (Feb. 2001), pp. 1560–1563. ISSN: 0036-8075. DOI: 10.1126/science.1055465.

[91] Jacqueline R. Hembrook-Short, Vanessa L. Mock, and Farran Briggs. "Attentional Modulation of Neuronal Activity Depends on Neuronal Feature Selectivity". eng. In: *Current biology: CB* 27.13 (July 2017), 1878–1887.e5. ISSN: 1879-0445. DOI: 10.1016/j.cub.2017.05.080.

[92] Eric Lowet et al. "Enhanced Neural Processing by Covert Attention only during Microsaccades Directed toward the Attended Stimulus".

eng. In: *Neuron* 99.1 (July 2018), 207–214.e3. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2018.05.041.

[93] John M. Findlay. *Active vision the psychology of looking and seeing*. eng. OCLC: 1227278813. Oxford: Oxford University Press, 2003. ISBN: 978-0-19-171181-7.

[94] I. D. Gilchrist, V. Brown, and J. M. Findlay. "Saccades without eye movements". eng. In: *Nature* 390.6656 (Nov. 1997), pp. 130–131. ISSN: 0028-0836. DOI: 10.1038/36478.

[95] Narcisse P. Bichot, Andrew F. Rossi, and Robert Desimone. "Parallel and serial neural mechanisms for visual search in macaque area V4". eng. In: *Science (New York, N.Y.)* 308.5721 (Apr. 2005), pp. 529–534. ISSN: 1095-9203. DOI: 10.1126/science.1109676.

[96] Kendrick Kay et al. "Tasks and their role in visual neuroscience". eng. In: *Neuron* 111.11 (June 2023), pp. 1697–1713. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2023.03.022.

[97] J.J. Gibson. *The senses considered as perceptual systems*. The senses considered as perceptual systems. Oxford, England: Houghton Mifflin, 1966.

[98] Alfred L. Yarbus. *Eye Movements and Vision*. en. Boston, MA: Springer US, 1967. ISBN: 978-1-4899-5381-0 978-1-4899-5379-7. DOI: 10.1007/978-1-4899-5379-7. URL: http://link.springer.com/10.1007/978-1-4899-5379-7 (visited on 07/29/2024).

[99] Sasha Archibald. *Ways of Seeing — Sasha Archibald*. en. URL: https://cabinetmagazine.org/issues/30/archibald.php (visited on 07/30/2024).

[100] William T. Powers. *Making sense of behavior: the meaning of control*. New Canaan, Conn: Benchmark Publications, 1998. ISBN: 978-0-9647121-5-7.

[101] Karl Friston et al. "Active Inference: A Process Theory". eng. In: *Neural Computation* 29.1 (Jan. 2017), pp. 1–49. ISSN: 1530-888X. DOI: 10.1162/NECO_a_00912.

[102] Ryan Smith, Karl J. Friston, and Christopher J. Whyte. "A step-by-step tutorial on active inference and its application to empirical data". eng. In: *Journal of Mathematical Psychology* 107 (Apr. 2022), p. 102632. ISSN: 0022-2496. DOI: 10.1016/j.jmp.2021.102632.

[103] P. Gao, B. O. Ploog, and H. P. Zeigler. "Whisking as a "voluntary" response: operant control of whisking parameters and effects of whisker

denervation". eng. In: *Somatosensory & Motor Research* 20.3-4 (2003), pp. 179–189. ISSN: 0899-0220. DOI: 10.1080/08990220310001623031.

[104] Rune W. Berg and David Kleinfeld. "Rhythmic whisking by rat: retraction as well as protraction of the vibrissae is under active muscular control". eng. In: *Journal of Neurophysiology* 89.1 (Jan. 2003), pp. 104–117. ISSN: 0022-3077. DOI: 10.1152/jn.00600.2002.

[105] Nicholas J. Sofroniew et al. "Natural whisker-guided behavior by head-fixed mice in tactile virtual reality". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34.29 (July 2014), pp. 9537–9550. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0712-14.2014.

[106] Jeffrey D. Moore et al. "Hierarchy of orofacial rhythms revealed through whisking and breathing". eng. In: *Nature* 497.7448 (May 2013), pp. 205–210. ISSN: 1476-4687. DOI: 10.1038/nature12076.

[107] Jun Takatoh et al. "The whisking oscillator circuit". eng. In: *Nature* 609.7927 (Sept. 2022), pp. 560–568. ISSN: 1476-4687. DOI: 10.1038/s41586-022-05144-8.

[108] V. Khatri et al. "Whisking in air: encoding of kinematics by VPM neurons in awake rats". eng. In: *Somatosensory & Motor Research* 27.3 (2010), pp. 111–120. ISSN: 1369-1651. DOI: 10.3109/08990220.2010.502381.

[109] Brian R. Isett and Daniel E. Feldman. "Cortical Coding of Whisking Phase during Surface Whisking". eng. In: *Current biology: CB* 30.16 (Aug. 2020), 3065–3074.e5. ISSN: 1879-0445. DOI: 10.1016/j.cub.2020.05.064.

[110] John C. Curtis and David Kleinfeld. "Phase-to-rate transformations encode touch in cortical neurons of a scanning sensorimotor system". eng. In: *Nature Neuroscience* 12.4 (Apr. 2009), pp. 492–501. ISSN: 1546-1726. DOI: 10.1038/nn.2283.

[111] Tim Näher et al. *Primate Saccade Rhythmicity*. en. Pages: 2023.09.27.559710 Section: New Results. June 2024. DOI: 10.1101/2023.09.27.559710. URL: https://www.biorxiv.org/content/10.1101/2023.09.27.559710v3 (visited on 07/29/2024).

[112] Jorge Otero-Millan et al. "Triggering mechanisms in microsaccade and saccade generation: a novel proposal". eng. In: *Annals of the New York Academy of Sciences* 1233 (Sept. 2011), pp. 107–116. ISSN: 1749-6632. DOI: 10.1111/j.1749-6632.2011.06177.x.

[113] Jeffrey D. Schall and Martin Paré. "The unknown but knowable relationship between Presaccadic Accumulation of activity and Saccade initiation". eng. In: *Journal of Computational Neuroscience* 49.3 (Aug. 2021), pp. 213–228. ISSN: 1573-6873. DOI: 10.1007/s10827-021-00784-7.

[114] Kush Paul and James W. Gnadt. "Activity of omnipause neurons during "staircase saccades" elicited by persistent microstimulation of the superior colliculus". eng. In: *Vision Research* 46.20 (Oct. 2006), pp. 3430–3442. ISSN: 0042-6989. DOI: 10.1016/j.visres.2006.05.014.

[115] Jorge Otero-Millan et al. "Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator". eng. In: *Journal of Vision* 8.14 (Dec. 2008), pp. 21.1–18. ISSN: 1534-7362. DOI: 10.1167/8.14.21.

[116] Jeffrey D. Schall. "Accumulators, Neurons, and Response Time". eng. In: *Trends in Neurosciences* 42.12 (Dec. 2019), pp. 848–860. ISSN: 1878-108X. DOI: 10.1016/j.tins.2019.10.001.

[117] Martin Rolfs and Richard Schweitzer. "Coupling perception to action through incidental sensory consequences of motor behaviour". en. In: *Nature Reviews Psychology* 1.2 (Feb. 2022), pp. 112–123. ISSN: 2731-0574. DOI: 10.1038/s44159-021-00015-x. URL: https://www.nature.com/articles/s44159-021-00015-x (visited on 12/15/2023).

[118] Paola Binda and Maria Concetta Morrone. "Vision During Saccadic Eye Movements". eng. In: *Annual Review of Vision Science* 4 (Sept. 2018), pp. 193–213. ISSN: 2374-4650. DOI: 10.1146/annurev-vision-091517-034317.

[119] Raymond Dodge. "Visual perception during eye movement." en. In: *Psychological Review* 7.5 (Sept. 1900), pp. 454–465. ISSN: 1939-1471, 0033-295X. DOI: 10.1037/h0067215. URL: https://doi.apa.org/doi/10.1037/h0067215 (visited on 07/29/2024).

[120] F. W. Campbell and R. H. Wurtz. "Saccadic omission: why we do not see a grey-out during a saccadic eye movement". eng. In: *Vision Research* 18.10 (1978), pp. 1297–1303. ISSN: 0042-6989. DOI: 10.1016/0042-6989(78)90219-5.

[121] D. C. Burr, M. C. Morrone, and J. Ross. "Selective suppression of the magnocellular visual pathway during saccadic eye movements". eng. In: *Nature* 371.6497 (Oct. 1994), pp. 511–513. ISSN: 0028-0836. DOI: 10.1038/371511a0.

[122] D. C. Burr, M. J. Morgan, and M. C. Morrone. "Saccadic suppression precedes visual motion analysis". eng. In: *Current biology: CB* 9.20 (Oct. 1999), pp. 1207–1209. ISSN: 0960-9822. DOI: 10.1016/S0960-9822(00)80028-7.

[123] Kai V. Thilo et al. "The site of saccadic suppression". eng. In: *Nature Neuroscience* 7.1 (Jan. 2004), pp. 13–14. ISSN: 1097-6256. DOI: 10.1038/nn1171.

[124] K. Matsumiya and K. Uchikawa. "Apparent size of an object remains uncompressed during presaccadic compression of visual space". eng. In: *Vision Research* 41.23 (Oct. 2001), pp. 3039–3050. ISSN: 0042-6989. DOI: 10.1016/s0042-6989(01)00174-2.

[125] Markus Lappe et al. "The fate of object features during perisaccadic mislocalization". eng. In: *Journal of Vision* 6.11 (Nov. 2006), pp. 1282–1293. ISSN: 1534-7362. DOI: 10.1167/6.11.11.

[126] Markus Wittenberg, Frank Bremmer, and Thomas Wachtler. "Perceptual evidence for saccadic updating of color stimuli". eng. In: *Journal of Vision* 8.14 (Oct. 2008), pp. 9.1–9. ISSN: 1534-7362. DOI: 10.1167/8.14.9.

[127] Guido M. Cicchini et al. "Transient spatiotopic integration across saccadic eye movements mediates visual stability". eng. In: *Journal of Neurophysiology* 109.4 (Feb. 2013), pp. 1117–1125. ISSN: 1522-1598. DOI: 10.1152/jn.00478.2012.

[128] James W. Bisley, Koorosh Mirpour, and Yelda Alkan. "The functional roles of neural remapping in cortex". eng. In: *Journal of Vision* 20.9 (Sept. 2020), p. 6. ISSN: 1534-7362. DOI: 10.1167/jov.20.9.6.

[129] Paola Binda et al. "Spatiotemporal distortions of visual perception at the time of saccades". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29.42 (Oct. 2009), pp. 13147–13157. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3723-09.2009.

[130] Zhong-Lin Lu and Barbara Anne Dosher. "Current directions in visual perceptual learning". eng. In: *Nature Reviews Psychology* 1.11 (Nov. 2022), pp. 654–668. ISSN: 2731-0574. DOI: 10.1038/s44159-022-00107-2.

[131] E.J. Gibson. *Principles of perceptual learning and development*. Principles of perceptual learning and development. East Norwalk, CT, US: Appleton-Century-Crofts, 1969.

[132] Connor Shorten and Taghi M. Khoshgoftaar. "A survey on Image Data Augmentation for Deep Learning". en. In: *Journal of Big Data* 6.1 (Dec.

2019), p. 60. ISSN: 2196-1115. DOI: 10.1186/s40537-019-0197-0. URL: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0 (visited on 07/29/2024).

[133] Mace William M. "James J. Gibson's strategy for perceiving : Ask not what's inside your head, but what's your head inside of". In: *Perceiving, acting, and knowing : Towards an ecological psychology* (1977). Publisher: Lawrence Erlbaum. URL: https://cir.nii.ac.jp/crid/1573668924783445888.

[134] Jarrod R. Dowdall. "The role of V1 spike rates, gamma oscillations, and inter-areal coherence in visual selective attention". PhD thesis. Radboud Universiteit Nijmegen, 2023. URL: https://repository.ubn.ru.nl/handle/2066/298504.

[135] S. Kornblum, T. Hasbroucq, and A. Osman. "Dimensional overlap: cognitive basis for stimulus-response compatibility–a model and taxonomy". eng. In: *Psychological Review* 97.2 (Apr. 1990), pp. 253–270. ISSN: 0033-295X. DOI: 10.1037/0033-295x.97.2.253.

[136] Jarrod Robert Dowdall et al. "Rapid learning of a fixation-based visual selective attention task and its effects on V1 firing rates". In: 2019. URL: https://www.abstractsonline.com/pp8/?_gl=1*143qom*_gcl_au*MTYwODQyNjcyNi4xNzIzOTk0MjUx*_ga*MTQ1NDUyMTE3MC4xNzIzOTk0MjUx*_ga_T09K3Q2WDN*MTcyMzk5NDI1MS4xLjEuMTcyMzk5NDI5Mi4xOS4wLjA.#!/7883/presentation/56373.

[137] Jarrod Robert Dowdall et al. "Rapid learning of a foveation-based attention task is tracked by V1 spike rates". Unpublished manuscript. 2023.

[138] T. Imai et al. "Interaction of the body, head, and eyes during walking and turning". eng. In: *Experimental Brain Research* 136.1 (Jan. 2001), pp. 1–18. ISSN: 0014-4819. DOI: 10.1007/s002210000533.

[139] D. H. Hubel and T. N. Wiesel. "Uniformity of monkey striate cortex: a parallel relationship between field size, scatter, and magnification factor". eng. In: *The Journal of Comparative Neurology* 158.3 (Dec. 1974), pp. 295–305. ISSN: 0021-9967. DOI: 10.1002/cne.901580305.

[140] F. P. Ottes, J. A. Van Gisbergen, and J. J. Eggermont. "Visuomotor fields of the superior colliculus: a quantitative model". eng. In: *Vision Research* 26.6 (1986), pp. 857–873. ISSN: 0042-6989. DOI: 10.1016/0042-6989(86)90144-6.

[141] Ben M. Harvey and Serge O. Dumoulin. "The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31.38 (Sept. 2011), pp. 13604–13612. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2572-11.2011.

[142] Jeremy Freeman and Eero P. Simoncelli. "Metamers of the ventral stream". eng. In: *Nature Neuroscience* 14.9 (Aug. 2011), pp. 1195–1201. ISSN: 1546-1726. DOI: 10.1038/nn.2889.

[143] L. Chelazzi et al. "Responses of neurons in inferior temporal cortex during memory-guided visual search". eng. In: *Journal of Neurophysiology* 80.6 (Dec. 1998), pp. 2918–2940. ISSN: 0022-3077. DOI: 10.1152/jn.1998.80.6.2918.

[144] M. Land, N. Mennie, and J. Rusted. "The roles of vision and eye movements in the control of activities of daily living". eng. In: *Perception* 28.11 (1999), pp. 1311–1328. ISSN: 0301-0066. DOI: 10.1068/p2935.

[145] A. T. Bahill, D. Adler, and L. Stark. "Most naturally occurring human saccades have magnitudes of 15 degrees or less". eng. In: *Investigative Ophthalmology* 14.6 (June 1975), pp. 468–469. ISSN: 0020-9988.

[146] Ying Zhang et al. "Object decoding with attention in inferior temporal cortex". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 108.21 (May 2011), pp. 8850–8855. ISSN: 1091-6490. DOI: 10.1073/pnas.1100999108.

[147] Shaul Hochstein and Merav Ahissar. "View from the top: hierarchies and reverse hierarchies in the visual system". eng. In: *Neuron* 36.5 (Dec. 2002), pp. 791–804. ISSN: 0896-6273. DOI: 10.1016/s0896-6273(02)01091-7.

[148] H. Deubel and W. X. Schneider. "Saccade target selection and object recognition: evidence for a common attentional mechanism". eng. In: *Vision Research* 36.12 (June 1996), pp. 1827–1837. ISSN: 0042-6989. DOI: 10.1016/0042-6989(95)00294-4.

[149] J. E. Hoffman and B. Subramaniam. "The role of visual attention in saccadic eye movements". eng. In: *Perception & Psychophysics* 57.6 (Aug. 1995), pp. 787–795. ISSN: 0031-5117. DOI: 10.3758/bf03206794.

[150] A.Terry Bahill, Michael R. Clark, and Lawrence Stark. "The main sequence, a tool for studying human eye movements". en. In: *Mathematical Biosciences* 24.3-4 (Jan. 1975), pp. 191–204. ISSN: 00255564. DOI:

10.1016/0025-5564(75)90075-9. URL: https://linkinghub.elsevier.com/retrieve/pii/0025556475900759 (visited on 08/05/2024).

[151] Theodoros P. Zanos et al. "A sensorimotor role for traveling waves in primate visual cortex". eng. In: *Neuron* 85.3 (Feb. 2015), pp. 615–627. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2014.12.043.

[152] D. P. Munoz, D. Pélisson, and D. Guitton. "Movement of neural activity on the superior colliculus motor map during gaze shifts". eng. In: *Science (New York, N.Y.)* 251.4999 (Mar. 1991), pp. 1358–1360. ISSN: 0036-8075. DOI: 10.1126/science.2003221.

[153] S. J. Luck. "The Spatiotemporal Dynamics of Visual-Spatial Attention". en. In: *From Attention to Goal-Directed Behavior*. Ed. by Francisco Aboitiz and Diego Cosmelli. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 51–66. ISBN: 978-3-540-70572-7 978-3-540-70573-4. DOI: 10.1007/978-3-540-70573-4_3. URL: http://link.springer.com/10.1007/978-3-540-70573-4_3 (visited on 07/29/2024).

[154] K. G. Thompson and J. D. Schall. "The detection of visual signals by macaque frontal eye field during masking". eng. In: *Nature Neuroscience* 2.3 (Mar. 1999), pp. 283–288. ISSN: 1097-6256. DOI: 10.1038/6398.

[155] L. Chelazzi et al. "Responses of neurons in macaque area V4 during memory-guided visual search". eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 11.8 (Aug. 2001), pp. 761–772. ISSN: 1047-3211. DOI: 10.1093/cercor/11.8.761.

[156] Daniel Guitton et al. "On the feedback control of orienting gaze shifts made with eye and head movements". eng. In: *Progress in Brain Research* 142 (2003), pp. 55–68. ISSN: 0079-6123. DOI: 10.1016/S0079-6123(03)42006-2.

[157] Robijanto Soetedjo, Chris R. S. Kaneko, and Albert F. Fuchs. "Evidence against a moving hill in the superior colliculus during saccadic eye movements in the monkey". eng. In: *Journal of Neurophysiology* 87.6 (June 2002), pp. 2778–2789. ISSN: 0022-3077. DOI: 10.1152/jn.2002.87.6.2778.

[158] Nima Ghitani et al. "Excitatory synaptic feedback from the motor layer to the sensory layers of the superior colliculus". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34.20 (May 2014), pp. 6822–6833. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3137-13.2014.

[159] Nicholas A. Steinmetz and Tirin Moore. "Eye movement preparation modulates neuronal responses in area V4 when dissociated from attentional demands". eng. In: *Neuron* 83.2 (July 2014), pp. 496–506. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2014.06.014.

[160] Lyle Muller et al. "The stimulus-evoked population response in visual cortex of awake monkey is a propagating wave". eng. In: *Nature Communications* 5 (Apr. 2014), p. 3675. ISSN: 2041-1723. DOI: 10.1038/ncomms4675.

[161] K. Rayner. "Eye movements in reading and information processing: 20 years of research". eng. In: *Psychological Bulletin* 124.3 (Nov. 1998), pp. 372–422. ISSN: 0033-2909. DOI: 10.1037/0033-2909.124.3.372.

[162] Grace Edwards, Rufin VanRullen, and Patrick Cavanagh. "Decoding Trans-Saccadic Memory". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 38.5 (Jan. 2018), pp. 1114–1123. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0854-17.2017.

[163] R. Desimone and J. Duncan. "Neural mechanisms of selective visual attention". eng. In: *Annual Review of Neuroscience* 18 (1995), pp. 193–222. ISSN: 0147-006X. DOI: 10.1146/annurev.ne.18.030195.001205.

[164] M. I. Posner, C. R. Snyder, and B. J. Davidson. "Attention and the detection of signals". eng. In: *Journal of Experimental Psychology* 109.2 (June 1980), pp. 160–174. ISSN: 0022-1015.

[165] C. Elizabeth Boudreau, Tori H. Williford, and John H. R. Maunsell. "Effects of task difficulty and target likelihood in area V4 of macaque monkeys". eng. In: *Journal of Neurophysiology* 96.5 (Nov. 2006), pp. 2377–2387. ISSN: 0022-3077. DOI: 10.1152/jn.01072.2005.

[166] Detlef Wegener et al. "Feature-based attention and the suppression of non-relevant object features". eng. In: *Vision Research* 48.27 (Dec. 2008), pp. 2696–2707. ISSN: 1878-5646. DOI: 10.1016/j.visres.2008.08.021.

[167] Marlene R. Cohen and John H. R. Maunsell. "Using neuronal populations to study the mechanisms underlying spatial and feature attention". eng. In: *Neuron* 70.6 (June 2011), pp. 1192–1204. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2011.04.029.

[168] Farran Briggs, George R. Mangun, and W. Martin Usrey. "Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits". eng. In: *Nature* 499.7459 (July 2013), pp. 476–480. ISSN: 1476-4687. DOI: 10.1038/nature12276.

[169] Guilhem Ibos, Jean-René Duhamel, and Suliann Ben Hamed. "A functional hierarchy within the parietofrontal network in stimulus selection and attention control". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 33.19 (May 2013), pp. 8359–8369. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.4058-12.2013.

[170] Tatiana A. Engel et al. "Selective modulation of cortical state during spatial attention". eng. In: *Science (New York, N.Y.)* 354.6316 (Dec. 2016), pp. 1140–1144. ISSN: 1095-9203. DOI: 10.1126/science.aag1420.

[171] J. Patrick Mayo and John H. R. Maunsell. "Graded Neuronal Modulations Related to Visual Spatial Attention". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 36.19 (May 2016), pp. 5353–5361. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0192-16.2016.

[172] Bram-Ernst Verhoef and John H. R. Maunsell. "Attention-related changes in correlated neuronal activity arise from normalization mechanisms". eng. In: *Nature Neuroscience* 20.7 (July 2017), pp. 969–977. ISSN: 1546-1726. DOI: 10.1038/nn.4572.

[173] Jacqueline R. Hembrook-Short et al. "Attention Enhances the Efficacy of Communication in V1 Local Circuits". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 39.6 (Feb. 2019), pp. 1066–1076. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2164-18.2018.

[174] Ramanujan Srinath, Douglas A. Ruff, and Marlene R. Cohen. "Attention improves information flow between neuronal populations without changing the communication subspace". eng. In: *Current biology: CB* 31.23 (Dec. 2021), 5299–5313.e4. ISSN: 1879-0445. DOI: 10.1016/j.cub.2021.09.076.

[175] B. C. Motter. "Neural correlates of attentive selection for color or luminance in extrastriate area V4". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 14.4 (Apr. 1994), pp. 2178–2189. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.14-04-02178.1994.

[176] S. Treue and J. H. Maunsell. "Attentional modulation of visual motion processing in cortical areas MT and MST". eng. In: *Nature* 382.6591 (Aug. 1996), pp. 539–541. ISSN: 0028-0836. DOI: 10.1038/382539a0.

[177] C. J. McAdams and J. H. Maunsell. "Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4". eng.

In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 19.1 (Jan. 1999), pp. 431–441. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.19-01-00431.1999.

[178] J. H. Reynolds, L. Chelazzi, and R. Desimone. "Competitive mechanisms subserve attention in macaque areas V2 and V4". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 19.5 (Mar. 1999), pp. 1736–1753. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.19-05-01736.1999.

[179] S. Treue and J. C. Martínez Trujillo. "Feature-based attention influences motion processing gain in macaque visual cortex". eng. In: *Nature* 399.6736 (June 1999), pp. 575–579. ISSN: 0028-0836. DOI: 10.1038/21176.

[180] S. Treue and J. H. Maunsell. "Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 19.17 (Sept. 1999), pp. 7591–7602. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.19-17-07591.1999.

[181] J. H. Reynolds, T. Pasternak, and R. Desimone. "Attention increases sensitivity of V4 neurons". eng. In: *Neuron* 26.3 (June 2000), pp. 703–714. ISSN: 0896-6273. DOI: 10.1016/s0896-6273(00)81206-4.

[182] Julio Martínez-Trujillo and Stefan Treue. "Attentional modulation strength in cortical area MT depends on stimulus contrast". eng. In: *Neuron* 35.2 (July 2002), pp. 365–370. ISSN: 0896-6273. DOI: 10.1016/s0896-6273(02)00778-x.

[183] Julio C. Martinez-Trujillo and Stefan Treue. "Feature-based attention increases the selectivity of population responses in primate visual cortex". eng. In: *Current biology: CB* 14.9 (May 2004), pp. 744–751. ISSN: 0960-9822. DOI: 10.1016/j.cub.2004.04.028.

[184] Carrie J. McAdams and R. Clay Reid. "Attention modulates the responses of simple cells in monkey primary visual cortex". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 25.47 (Nov. 2005), pp. 11023–11033. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2904-05.2005.

[185] K. Taylor et al. "Coherent oscillatory activity in monkey area v4 predicts successful allocation of attention". eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 15.9 (Sept. 2005), pp. 1424–1437. ISSN: 1047-3211. DOI: 10.1093/cercor/bhi023.

[186] Tori Williford and John H. R. Maunsell. "Effects of spatial attention on contrast response functions in macaque area V4". eng. In: *Journal of Neurophysiology* 96.1 (July 2006), pp. 40–54. ISSN: 0022-3077. DOI: 10.1152/jn.01207.2005.

[187] Thilo Womelsdorf et al. "Dynamic shifts of visual receptive fields in cortical area MT by spatial attention". eng. In: *Nature Neuroscience* 9.9 (Sept. 2006), pp. 1156–1160. ISSN: 1097-6256. DOI: 10.1038/nn1748.

[188] Thilo Womelsdorf et al. "Gamma-band synchronization in visual cortex predicts speed of change detection". eng. In: *Nature* 439.7077 (Feb. 2006), pp. 733–736. ISSN: 1476-4687. DOI: 10.1038/nature04258.

[189] Pascal Fries et al. "The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 28.18 (Apr. 2008), pp. 4823–4835. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.4499-07.2008.

[190] J. L. Herrero et al. "Acetylcholine contributes through muscarinic receptors to attentional modulation in V1". eng. In: *Nature* 454.7208 (Aug. 2008), pp. 1110–1114. ISSN: 1476-4687. DOI: 10.1038/nature07141.

[191] Georgia G. Gregoriou et al. "High-frequency, long-range coupling between prefrontal and visual cortex during attention". eng. In: *Science (New York, N.Y.)* 324.5931 (May 2009), pp. 1207–1210. ISSN: 1095-9203. DOI: 10.1126/science.1171402.

[192] Alexander Thiele et al. "Additive effects of attention and stimulus contrast in primary visual cortex". eng. In: *Cerebral Cortex (New York, N.Y.: 1991)* 19.12 (Dec. 2009), pp. 2970–2981. ISSN: 1460-2199. DOI: 10.1093/cercor/bhp070.

[193] Matthew Chalk et al. "Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in V1". eng. In: *Neuron* 66.1 (Apr. 2010), pp. 114–125. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2010.03.013.

[194] Lee P. Lovejoy and Richard J. Krauzlis. "Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments". eng. In: *Nature Neuroscience* 13.2 (Feb. 2010), pp. 261–266. ISSN: 1546-1726. DOI: 10.1038/nn.2470.

[195] Elizabeth A. Buffalo et al. "Laminar differences in gamma and alpha coherence in the ventral stream". eng. In: *Proceedings of the National*

Academy of Sciences of the United States of America 108.27 (July 2011), pp. 11262–11267. ISSN: 1091-6490. DOI: 10.1073/pnas.1011284108.

[196] Yuri B. Saalmann et al. "The pulvinar regulates information transmission between cortical areas based on attention demands". eng. In: *Science (New York, N.Y.)* 337.6095 (Aug. 2012), pp. 753–756. ISSN: 1095-9203. DOI: 10.1126/science.1223082.

[197] Jose L. Herrero et al. "Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors". eng. In: *Neuron* 78.4 (May 2013), pp. 729–739. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2013.03.029.

[198] Martin Vinck et al. "Attentional modulation of cell-class-specific gamma-band synchronization in awake monkey area v4". eng. In: *Neuron* 80.4 (Nov. 2013), pp. 1077–1089. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2013.08.019.

[199] Georgia G. Gregoriou et al. "Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4". eng. In: *Nature Neuroscience* 17.7 (July 2014), pp. 1003–1011. ISSN: 1546-1726. DOI: 10.1038/nn.3742.

[200] Benjamin Voloh et al. "Theta-gamma coordination between anterior cingulate and prefrontal cortex indexes correct attention shifts". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 112.27 (July 2015), pp. 8457–8462. ISSN: 1091-6490. DOI: 10.1073/pnas.1500438112.

[201] Guilhem Ibos and David J. Freedman. "Interaction between Spatial and Feature Attention in Posterior Parietal Cortex". eng. In: *Neuron* 91.4 (Aug. 2016), pp. 931–943. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2016.07.025.

[202] Alexander Thiele et al. "Attention Induced Gain Stabilization in Broad and Narrow-Spiking Cells in the Frontal Eye-Field of Macaque Monkeys". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 36.29 (July 2016), pp. 7601–7612. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0872-16.2016.

[203] Huihui Zhou, Robert John Schafer, and Robert Desimone. "Pulvinar-Cortex Interactions in Vision and Attention". eng. In: *Neuron* 89.1 (Jan. 2016), pp. 209–220. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2015.11.034.

[204] Craig G. Richter et al. "Top-Down Beta Enhances Bottom-Up Gamma". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 37.28 (July 2017), pp. 6698–6711. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3771-16.2017.

[205] Douglas A. Ruff and Marlene R. Cohen. "A normalization model suggests that attention changes the weighting of inputs between visual areas". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 114.20 (May 2017), E4085–E4094. ISSN: 1091-6490. DOI: 10.1073/pnas.1619857114.

[206] Iris Grothe et al. "Attention Selectively Gates Afferent Signal Transmission to Area V4". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 38.14 (Apr. 2018), pp. 3441–3452. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2221-17.2018.

[207] Georgios Spyropoulos, Conrado Arturo Bosman, and Pascal Fries. "A theta rhythm in macaque visual cortex and its attentional modulation". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 115.24 (June 2018), E5614–E5623. ISSN: 1091-6490. DOI: 10.1073/pnas.1719433115.

[208] Miguel Dasilva et al. "Cell class-specific modulation of attentional signals by acetylcholine in macaque frontal eye field". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 116.40 (Oct. 2019), pp. 20180–20189. ISSN: 1091-6490. DOI: 10.1073/pnas.1905413116.

[209] Demetrio Ferro et al. "Directed information exchange between cortical layers in macaque V1 and V4 and its modulation by selective attention". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 118.12 (Mar. 2021), e2022097118. ISSN: 1091-6490. DOI: 10.1073/pnas.2022097118.

[210] Jochem van Kempen et al. "Top-down coordination of local cortical state during selective attention". eng. In: *Neuron* 109.5 (Mar. 2021), 894–904.e8. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2020.12.013.

[211] Mark J. Prescott and Hannah M. Buchanan-Smith. "Training nonhuman primates using positive reinforcement techniques". eng. In: *Journal of applied animal welfare science: JAAWS* 6.3 (2003), pp. 157–161. ISSN: 1088-8705. DOI: 10.1207/S15327604JAWS0603_01.

[212] Benjamin Fischer and Detlef Wegener. "Emphasizing the "positive" in positive reinforcement: using nonbinary rewarding for training monkeys

on cognitive tasks". eng. In: *Journal of Neurophysiology* 120.1 (July 2018), pp. 115–128. ISSN: 1522-1598. DOI: 10.1152/jn.00572.2017.

[213] Bram-Ernst Verhoef and John Hr Maunsell. "Attention operates uniformly throughout the classical receptive field and the surround". eng. In: *eLife* 5 (Aug. 2016), e17256. ISSN: 2050-084X. DOI: 10.7554/eLife.17256.

[214] Anirvan S. Nandy, Jonathan J. Nassi, and John H. Reynolds. "Laminar Organization of Attentional Modulation in Macaque Visual Area V4". eng. In: *Neuron* 93.1 (Jan. 2017), pp. 235–246. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2016.11.029.

[215] A. M. Ni et al. "Learning and attention reveal a general relationship between population activity and behavior". eng. In: *Science (New York, N.Y.)* 359.6374 (Jan. 2018), pp. 463–465. ISSN: 1095-9203. DOI: 10.1126/science.aao0284.

[216] Douglas A. Ruff and Marlene R. Cohen. "Simultaneous multi-area recordings suggest that attention improves performance by reshaping stimulus representations". eng. In: *Nature Neuroscience* 22.10 (Oct. 2019), pp. 1669–1676. ISSN: 1546-1726. DOI: 10.1038/s41593-019-0477-1.

[217] Andrew F. Rossi et al. "Top down attentional deficits in macaques with lesions of lateral prefrontal cortex". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27.42 (Oct. 2007), pp. 11306–11314. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2939-07.2007.

[218] P. R. Roelfsema, V. A. Lamme, and H. Spekreijse. "Object-based attention in the primary visual cortex of the macaque monkey". eng. In: *Nature* 395.6700 (Sept. 1998), pp. 376–381. ISSN: 0028-0836. DOI: 10.1038/26475.

[219] Eleni Psarou et al. "Modular, cement-free, customized headpost and connector-chamber implants for macaques". eng. In: *Journal of Neuroscience Methods* 393 (June 2023), p. 109899. ISSN: 1872-678X. DOI: 10.1016/j.jneumeth.2023.109899.

[220] Alina Peter et al. "Surface color and predictability determine contextual modulation of V1 firing and gamma oscillations". eng. In: *eLife* 8 (Feb. 2019), e42101. ISSN: 2050-084X. DOI: 10.7554/eLife.42101.

[221] Iris Grothe et al. "Switching neuronal inputs by differential modulations of gamma-band phase-coherence". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 32.46 (Nov. 2012),

pp. 16172–16180. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.0890-12.2012.

[222] P. Trillenberg et al. "CNV and temporal uncertainty with 'ageing' and 'non-ageing' S1-S2 intervals". eng. In: *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology* 111.7 (July 2000), pp. 1216–1226. ISSN: 1388-2457. DOI: 10.1016/s1388-2457(00)00274-1.

[223] Ralf Engbert and Reinhold Kliegl. "Microsaccades uncover the orientation of covert attention". eng. In: *Vision Research* 43.9 (Apr. 2003), pp. 1035–1045. ISSN: 0042-6989. DOI: 10.1016/s0042-6989(03)00084-1.

[224] Heiko H. Schütt et al. "Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data". eng. In: *Vision Research* 122 (May 2016), pp. 105–123. ISSN: 1878-5646. DOI: 10.1016/j.visres.2016.02.002.

[225] T. P. Trappenberg et al. "A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus". eng. In: *Journal of Cognitive Neuroscience* 13.2 (Feb. 2001), pp. 256–271. ISSN: 0898-929X. DOI: 10.1162/089892901564306.

[226] Jeffrey D. Schall. "The neural selection and control of saccades by the frontal eye field". eng. In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 357.1424 (Aug. 2002), pp. 1073–1082. ISSN: 0962-8436. DOI: 10.1098/rstb.2002.1098.

[227] James W. Bisley and Michael E. Goldberg. "Neuronal activity in the lateral intraparietal area and spatial attention". eng. In: *Science (New York, N.Y.)* 299.5603 (Jan. 2003), pp. 81–86. ISSN: 1095-9203. DOI: 10.1126/science.1077395.

[228] L. Caitlin Elmore et al. "Visual short-term memory compared in rhesus monkeys and humans". eng. In: *Current biology: CB* 21.11 (June 2011), pp. 975–979. ISSN: 1879-0445. DOI: 10.1016/j.cub.2011.04.031.

[229] E. B. Cutrell and R. T. Marrocco. "Electrical microstimulation of primate posterior parietal cortex initiates orienting and alerting components of covert attention". eng. In: *Experimental Brain Research* 144.1 (May 2002), pp. 103–113. ISSN: 0014-4819. DOI: 10.1007/s00221-002-1032-x.

[230] Gustavo Rohenkohl, Conrado Arturo Bosman, and Pascal Fries. "Gamma Synchronization between V1 and V4 Improves Behavioral Performance".

eng. In: *Neuron* 100.4 (Nov. 2018), 953–963.e3. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2018.09.019.

[231] Philip L. Smith. "Spatial attention and the detection of weak visual signals". In: *Tutorials in visual cognition* (2010). Publisher: Psychology Press New York (NY), pp. 211–259. URL: https://books.google.com/books?hl=en&lr=&id=lIwBmKik4M8C&oi=fnd&pg=PA211&dq=Spatial+Attention+and+the+Detection+of+Weak+Visual+Signals&ots=kK9xH91Y_I&sig=Y6x4ov9PohTkyLj74cO5hoVgvoc (visited on 12/17/2023).

[232] Michael I. Posner and Yoav Cohen. "Components of visual orienting". In: *Attention and performance X: Control of language processes* 32 (1984). Publisher: Hilldale, NJ, pp. 531–556. URL: https://www.researchgate.net/profile/Yoav-Cohen/publication/203918232_Components_of_visual_orienting/links/0c960516ab4213e497000000/Components-of-visual-orienting.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail (visited on 12/17/2023).

[233] Sam Ling and Marisa Carrasco. "When sustained attention impairs perception". eng. In: *Nature Neuroscience* 9.10 (Oct. 2006), pp. 1243–1245. ISSN: 1097-6256. DOI: 10.1038/nn1761.

[234] F. A. Wichmann and N. J. Hill. "The psychometric function: I. Fitting, sampling, and goodness of fit". eng. In: *Perception & Psychophysics* 63.8 (Nov. 2001), pp. 1293–1313. ISSN: 0031-5117. DOI: 10.3758/bf03194544.

[235] Brian J. White and Douglas P. Munoz. "The superior colliculus". In: *Oxford handbook of eye movements* 1 (2011). Publisher: Oxford University Press New York, pp. 195–213. URL: https://books.google.com/books?hl=en&lr=&id=Hiof9gzHPNoC&oi=fnd&pg=PA195&dq=The+superior+colliculus.+&ots=zupjUVS9sG&sig=ZuHY3PYCPZbHj1GnB_0Njvjw3ic (visited on 12/17/2023).

[236] Patricia S. Churchland, Vilayanur S. Ramachandran, and Terrence J. Sejnowski. "A critique of pure vision". In: (1993). URL: https://philpapers.org/rec/CHUACO-2 (visited on 12/17/2023).

[237] Andreas K. Engel et al. "Where's the action? The pragmatic turn in cognitive science". eng. In: *Trends in Cognitive Sciences* 17.5 (May 2013), pp. 202–209. ISSN: 1879-307X. DOI: 10.1016/j.tics.2013.03.006.

[238] Alex Gomez-Marin and Asif A. Ghazanfar. "The Life of Behavior". eng. In: *Neuron* 104.1 (Oct. 2019), pp. 25–36. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2019.09.017.

[239] Philip R. L. Parker et al. "A dynamic sequence of visual processing initiated by gaze shifts". eng. In: *Nature Neuroscience* 26.12 (Dec. 2023), pp. 2192–2202. ISSN: 1546-1726. DOI: 10.1038/s41593-023-01481-7.

[240] R. H. Wurtz. "Visual receptive fields of striate cortex neurons in awake monkeys". eng. In: *Journal of Neurophysiology* 32.5 (Sept. 1969), pp. 727–742. ISSN: 0022-3077. DOI: 10.1152/jn.1969.32.5.727.

[241] Patrick Jendritza et al. "Visual Neuroscience Methods for Marmosets: Efficient Receptive Field Mapping and Head-Free Eye Tracking". en. In: *eNeuro* 8.3 (May 2021). Publisher: Society for Neuroscience Section: Research Article: Methods/New Tools. ISSN: 2373-2822. DOI: 10.1523/ENEURO.0489-20.2021. URL: https://www.eneuro.org/content/8/3/ENEURO.0489-20.2021 (visited on 12/17/2023).

[242] György Buzsáki and David Tingley. "Cognition from the Body-Brain Partnership: Exaptation of Memory". eng. In: *Annual Review of Neuroscience* 46 (July 2023), pp. 191–210. ISSN: 1545-4126. DOI: 10.1146/annurev-neuro-101222-110632.

[243] R. B. Tootell et al. "Deoxyglucose analysis of retinotopic organization in primate striate cortex". eng. In: *Science (New York, N.Y.)* 218.4575 (Nov. 1982), pp. 902–904. ISSN: 0036-8075. DOI: 10.1126/science.7134981.

[244] Brian A. Wandell and Jonathan Winawer. "Imaging retinotopic maps in the human brain". eng. In: *Vision Research* 51.7 (Apr. 2011), pp. 718–737. ISSN: 1878-5646. DOI: 10.1016/j.visres.2010.08.004.

[245] Chih-Yang Chen et al. "The Foveal Visual Representation of the Primate Superior Colliculus". eng. In: *Current biology: CB* 29.13 (July 2019), 2109–2119.e7. ISSN: 1879-0445. DOI: 10.1016/j.cub.2019.05.040.

[246] J. D. Schall et al. "Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 15.6 (June 1995), pp. 4464–4487. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.15-06-04464.1995.

[247] Reza Rajimehr et al. "Retinotopy versus face selectivity in macaque visual cortex". eng. In: *Journal of Cognitive Neuroscience* 26.12 (Dec. 2014), pp. 2691–2700. ISSN: 1530-8898. DOI: 10.1162/jocn_a_00672.

[248] Fabrizio Santini et al. "EyeRIS: a general-purpose system for eye-movement-contingent display control". eng. In: *Behavior Research Methods* 39.3 (Aug. 2007), pp. 350–364. ISSN: 1554-351X. DOI: 10.3758/bf03193003.

[249] Daniel R. Saunders and Russell L. Woods. "Direct measurement of the system latency of gaze-contingent displays". en. In: *Behavior Research Methods* 46.2 (June 2014), pp. 439–447. ISSN: 1554-3528. DOI: 10.3758/s13428-013-0375-5. URL: https://doi.org/10.3758/s13428-013-0375-5 (visited on 12/17/2023).

[250] George W. McConkie. "Eye Movement Contingent Display Control: Personal Reflections and Comments". en. In: *Scientific Studies of Reading* 1.4 (Oct. 1997), pp. 303–316. ISSN: 1088-8438, 1532-799X. DOI: 10.1207/s1532799xssr0104_1. URL: http://www.tandfonline.com/doi/abs/10.1207/s1532799xssr0104_1 (visited on 12/17/2023).

[251] George W. McConkie and Lester C. Loschky. "Perception onset time during fixations in free viewing". en. In: *Behavior Research Methods, Instruments, & Computers* 34.4 (Nov. 2002), pp. 481–490. ISSN: 1532-5970. DOI: 10.3758/BF03195477. URL: https://doi.org/10.3758/BF03195477 (visited on 12/17/2023).

[252] Richard L. Gregory. *Eye and brain: the psychology of seeing*. eng. Fifth edition. Princeton science library. Princeton Oxford: Princeton University Press, 2015. ISBN: 978-0-691-04837-6 978-0-691-16516-5 978-0-691-04840-6.

[253] Beizhen Zhang et al. "Transforming absolute value to categorical choice in primate superior colliculus during value-based decision making". eng. In: *Nature Communications* 12.1 (June 2021), p. 3410. ISSN: 2041-1723. DOI: 10.1038/s41467-021-23747-z.

[254] Hans Supèr et al. "Correspondence of presaccadic activity in the monkey primary visual cortex with saccadic eye movements". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 101.9 (Mar. 2004), pp. 3230–3235. ISSN: 0027-8424. DOI: 10.1073/pnas.0400433101.

[255] T. Moore, A. S. Tolias, and P. H. Schiller. "Visual representations during saccadic eye movements". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 95.15 (July 1998), pp. 8981–8984. ISSN: 0027-8424. DOI: 10.1073/pnas.95.15.8981.

[256] Angela L. Gee, Anna E. Ipata, and Michael E. Goldberg. "Activity in V4 Reflects the Direction, But Not the Latency, of Saccades During Visual Search". In: *Journal of Neurophysiology* 104.4 (Oct. 2010). Publisher: American Physiological Society, pp. 2187–2193. ISSN: 0022-3077. DOI:

10.1152/jn.00898.2009. URL: https://journals.physiology.org/doi/full/10.1152/jn.00898.2009 (visited on 12/17/2023).

[257] Nicolas M. Brunet et al. "Stimulus repetition modulates gamma-band synchronization in primate visual cortex". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 111.9 (Mar. 2014), pp. 3626–3631. ISSN: 1091-6490. DOI: 10.1073/pnas.1309714111.

[258] Mario Fiorani et al. "Automatic mapping of visual cortex receptive fields: a fast and precise algorithm". eng. In: *Journal of Neuroscience Methods* 221 (Jan. 2014), pp. 112–126. ISSN: 1872-678X. DOI: 10.1016/j.jneumeth.2013.09.012.

[259] Dario L. Ringach. "Mapping receptive fields in primary visual cortex". eng. In: *The Journal of Physiology* 558.Pt 3 (Aug. 2004), pp. 717–728. ISSN: 0022-3751. DOI: 10.1113/jphysiol.2004.065771.

[260] Anne M. Treisman and Garry Gelade. "A feature-integration theory of attention". en. In: *Cognitive Psychology* 12.1 (Jan. 1980), pp. 97–136. ISSN: 00100285. DOI: 10.1016/0010-0285(80)90005-5. URL: https://linkinghub.elsevier.com/retrieve/pii/0010028580900055 (visited on 12/17/2023).

[261] B. C. Motter and E. J. Belky. "The guidance of eye movements during active visual search". eng. In: *Vision Research* 38.12 (June 1998), pp. 1805–1815. ISSN: 0042-6989. DOI: 10.1016/s0042-6989(97)00349-0.

[262] Jude F. Mitchell, John H. Reynolds, and Cory T. Miller. "Active vision in marmosets: a model system for visual neuroscience". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 34.4 (Jan. 2014), pp. 1183–1194. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.3899-13.2014.

[263] R. Klein. "Inhibitory tagging system facilitates visual search". eng. In: *Nature* 334.6181 (Aug. 1988), pp. 430–431. ISSN: 0028-0836. DOI: 10.1038/334430a0.

[264] Niklas Wilming et al. "Saccadic Momentum and Facilitation of Return Saccades Contribute to an Optimal Foraging Strategy". en. In: *PLoS Computational Biology* 9.1 (Jan. 2013). Ed. by Olaf Sporns, e1002871. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1002871. URL: https://dx.plos.org/10.1371/journal.pcbi.1002871 (visited on 12/15/2023).

[265] Wei Ji Ma, Konrad Kording, and Daniel Goldreich. *Bayesian models of perception and action: an introduction*. Cambridge, MA: MIT Press, 2023. ISBN: 978-0-262-37283-1 978-0-262-37282-4.

[266] Yufeng Zhang and Pascal Fries. *Eccentricity-Dependent Saccadic Reaction Time: The Roles of Foveal Magnification and Attentional Orienting*. en. Aug. 2023. DOI: 10.1101/2023.08.08.552339. URL: http://biorxiv.org/lookup/doi/10.1101/2023.08.08.552339 (visited on 07/03/2024).

[267] A. E. Bartz. "Eye-movement latency, duration, and response time as a function of angular displacement". eng. In: *Journal of Experimental Psychology* 64 (Sept. 1962), pp. 318–324. ISSN: 0022-1015. DOI: 10.1037/h0043475.

[268] Ziad M. Hafed and Laurent Goffart. "Gaze direction as equilibrium: more evidence from spatial and temporal aspects of small-saccade triggering in the rhesus macaque monkey". eng. In: *Journal of Neurophysiology* 123.1 (Jan. 2020), pp. 308–322. ISSN: 1522-1598. DOI: 10.1152/jn.00588.2019.

[269] R. P. Kalesnykas and P. E. Hallett. "Retinal eccentricity and the latency of eye saccades". eng. In: *Vision Research* 34.4 (Feb. 1994), pp. 517–531. ISSN: 0042-6989. DOI: 10.1016/0042-6989(94)90165-1.

[270] R. P. Kalesnykas and P. E. Hallett. "Fixation conditions, the foveola and saccadic latency". eng. In: *Vision Research* 36.19 (Oct. 1996), pp. 3195–3203. ISSN: 0042-6989. DOI: 10.1016/0042-6989(96)00029-6.

[271] D. Wyman and R. M. Steinman. "Letter: Latency characteristics of small saccades". eng. In: *Vision Research* 13.11 (Nov. 1973), pp. 2173–2175. ISSN: 0042-6989. DOI: 10.1016/0042-6989(73)90195-8.

[272] A. H. Bell, S. Everling, and D. P. Munoz. "Influence of stimulus eccentricity and direction on characteristics of pro- and antisaccades in non-human primates". eng. In: *Journal of Neurophysiology* 84.5 (Nov. 2000), pp. 2595–2604. ISSN: 0022-3077. DOI: 10.1152/jn.2000.84.5.2595.

[273] Joan M. Dafoe, Irene T. Armstrong, and Doug P. Munoz. "The influence of stimulus direction and eccentricity on pro- and anti-saccades in humans". eng. In: *Experimental Brain Research* 179.4 (June 2007), pp. 563–570. ISSN: 0014-4819. DOI: 10.1007/s00221-006-0817-8.

[274] Antje Nuthmann et al. "No Evidence for a Saccadic Range Effect for Visually Guided and Memory-Guided Saccades in Simple Saccade-Targeting Tasks". eng. In: *PloS One* 11.9 (2016), e0162449. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0162449.

[275] M. Paré and D. P. Munoz. "Saccadic reaction time in the monkey: advanced preparation of oculomotor programs is primarily responsible for express saccade occurrence". eng. In: *Journal of Neurophysiology* 76.6 (Dec. 1996), pp. 3666–3681. ISSN: 0022-3077. DOI: 10.1152/jn.1996.76.6.3666.

[276] D. P. Hanes and J. D. Schall. "Neural control of voluntary movement initiation". eng. In: *Science (New York, N.Y.)* 274.5286 (Oct. 1996), pp. 427–430. ISSN: 0036-8075. DOI: 10.1126/science.274.5286.427.

[277] Simon P. Liversedge, Iain D. Gilchrist, and Stefan Everling, eds. *The Oxford handbook of eye movements*. Oxford library of psychology. OCLC: ocn713186814. Oxford ; New York: Oxford University Press, 2011. ISBN: 978-0-19-953978-9.

[278] J. P. Rijsdijk, J. N. Kroon, and G. J. van der Wildt. "Contrast sensitivity as a function of position on the retina". eng. In: *Vision Research* 20.3 (1980), pp. 235–241. ISSN: 0042-6989. DOI: 10.1016/0042-6989(80)90108-x.

[279] P. Azzopardi and A. Cowey. "The overrepresentation of the fovea and adjacent retina in the striate cortex and dorsal lateral geniculate nucleus of the macaque monkey". eng. In: *Neuroscience* 72.3 (June 1996), pp. 627–639. ISSN: 0306-4522. DOI: 10.1016/0306-4522(95)00589-7.

[280] J. M. Findlay. "The visual stimulus for saccadic eye movements in human observers". eng. In: *Perception* 9.1 (1980), pp. 7–21. ISSN: 0301-0066. DOI: 10.1068/p090007.

[281] M. Cynader and N. Berman. "Receptive-field organization of monkey superior colliculus". eng. In: *Journal of Neurophysiology* 35.2 (Mar. 1972), pp. 187–201. ISSN: 0022-3077. DOI: 10.1152/jn.1972.35.2.187.

[282] Chih-Yang Chen and Ziad M. Hafed. "Orientation and Contrast Tuning Properties and Temporal Flicker Fusion Characteristics of Primate Superior Colliculus Neurons". eng. In: *Frontiers in Neural Circuits* 12 (2018), p. 58. ISSN: 1662-5110. DOI: 10.3389/fncir.2018.00058.

[283] H. Doma and P. E. Hallett. "Rod-cone dependence of saccadic eye-movement latency in a foveating task". eng. In: *Vision Research* 28.8 (1988), pp. 899–913. ISSN: 0042-6989. DOI: 10.1016/0042-6989(88)90099-5.

[284] R. H. Carpenter and M. L. Williams. "Neural computation of log likelihood in control of saccadic eye movements". eng. In: *Nature* 377.6544 (Sept. 1995), pp. 59–62. ISSN: 0028-0836. DOI: 10.1038/377059a0.

[285] Roger Ratcliff and Gail McKoon. "The diffusion decision model: theory and data for two-choice decision tasks". eng. In: *Neural Computation* 20.4 (Apr. 2008), pp. 873–922. ISSN: 0899-7667. DOI: 10.1162/neco.2008.12-06-420.

[286] T. Van Zandt. "How to fit a response time distribution". eng. In: *Psychonomic Bulletin & Review* 7.3 (Sept. 2000), pp. 424–465. ISSN: 1069-9384. DOI: 10.3758/bf03214357.

[287] Pascal Fries and Eric Maris. "What to Do If N Is Two?" eng. In: *Journal of Cognitive Neuroscience* 34.7 (June 2022), pp. 1114–1118. ISSN: 1530-8898. DOI: 10.1162/jocn_a_01857.

[288] A. H. Bell et al. "Stimulus intensity modifies saccadic reaction time and visual response latency in the superior colliculus". eng. In: *Experimental Brain Research* 174.1 (Sept. 2006), pp. 53–59. ISSN: 0014-4819. DOI: 10.1007/s00221-006-0420-z.

[289] T. J. Gawne, T. W. Kjaer, and B. J. Richmond. "Latency: another potential code for feature binding in striate cortex". eng. In: *Journal of Neurophysiology* 76.2 (Aug. 1996), pp. 1356–1360. ISSN: 0022-3077. DOI: 10.1152/jn.1996.76.2.1356.

[290] Joonyeol Lee, Tori Williford, and John H. R. Maunsell. "Spatial attention and the latency of neuronal responses in macaque area V4". eng. In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27.36 (Sept. 2007), pp. 9632–9637. ISSN: 1529-2401. DOI: 10.1523/JNEUROSCI.2734-07.2007.

[291] Casimir J. H. Ludwig, Iain D. Gilchrist, and Eugene McSorley. "The influence of spatial frequency and contrast on saccade latencies". eng. In: *Vision Research* 44.22 (2004), pp. 2597–2604. ISSN: 0042-6989. DOI: 10.1016/j.visres.2004.05.022.

[292] Bijan Pesaran et al. "Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation". eng. In: *Nature Neuroscience* 21.7 (July 2018), pp. 903–919. ISSN: 1546-1726. DOI: 10.1038/s41593-018-0171-8.

[293] Tomohiro Tanaka et al. "Visual response of neurons in the lateral intraparietal area and saccadic reaction time during a visual detection task". eng. In: *The European Journal of Neuroscience* 37.6 (Mar. 2013), pp. 942–956. ISSN: 1460-9568. DOI: 10.1111/ejn.12100.

[294] K. Nakayama and M. Mackeben. “Sustained and transient components of focal visual attention”. eng. In: *Vision Research* 29.11 (1989), pp. 1631–1647. ISSN: 0042-6989. DOI: 10.1016/0042-6989(89)90144-2.

[295] Pascal Fries. “Rhythmic attentional scanning”. eng. In: *Neuron* 111.7 (Apr. 2023), pp. 954–970. ISSN: 1097-4199. DOI: 10.1016/j.neuron.2023.02.015.

[296] Antimo Buonocore and Robert D. McIntosh. “Saccadic inhibition underlies the remote distractor effect”. eng. In: *Experimental Brain Research* 191.1 (Oct. 2008), pp. 117–122. ISSN: 1432-1106. DOI: 10.1007/s00221-008-1558-7.

[297] Eyal M. Reingold and Dave M. Stampe. “Saccadic inhibition in voluntary and reflexive saccades”. eng. In: *Journal of Cognitive Neuroscience* 14.3 (Apr. 2002), pp. 371–388. ISSN: 0898-929X. DOI: 10.1162/089892902317361903.

[298] Neeraj J. Gandhi and Husam A. Katnani. “Motor functions of the superior colliculus”. eng. In: *Annual Review of Neuroscience* 34 (2011), pp. 205–231. ISSN: 1545-4126. DOI: 10.1146/annurev-neuro-061010-113728.

[299] M. Carrasco and K. S. Frieder. “Cortical magnification neutralizes the eccentricity effect in visual search”. eng. In: *Vision Research* 37.1 (Jan. 1997), pp. 63–82. ISSN: 0042-6989. DOI: 10.1016/s0042-6989(96)00102-2.

[300] Richard Veale, Ziad M. Hafed, and Masatoshi Yoshida. “How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling”. eng. In: *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 372.1714 (Feb. 2017), p. 20160113. ISSN: 1471-2970. DOI: 10.1098/rstb.2016.0113.

[301] R. Boch, B. Fischer, and E. Ramsperger. “Express-saccades of the monkey: reaction times versus intensity, size, duration, and eccentricity of their targets”. eng. In: *Experimental Brain Research* 55.2 (1984), pp. 223–231. ISSN: 0014-4819. DOI: 10.1007/BF00237273.

[302] J. P. De Vries, R. Azadi, and M. R. Harwood. “The saccadic size-latency phenomenon explored: Proximal target size is a determining factor in the saccade latency”. eng. In: *Vision Research* 129 (Dec. 2016), pp. 87–97. ISSN: 1878-5646. DOI: 10.1016/j.visres.2016.09.006.

[303] Raphael Koster et al. “Stimulus Novelty Energizes Actions in the Absence of Explicit Reward”. eng. In: *PloS One* 11.7 (2016), e0159120. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0159120.

[304] L. Thaler et al. “What is the best fixation target? The effect of target shape on stability of fixational eye movements”. eng. In: *Vision Research* 76 (Jan. 2013), pp. 31–42. ISSN: 1878-5646. DOI: 10.1016/j.visres.2012.10.012.

[305] M. E. Goldberg and R. H. Wurtz. “Activity of superior colliculus in behaving monkey. I. Visual receptive fields of single neurons”. eng. In: *Journal of Neurophysiology* 35.4 (July 1972), pp. 542–559. ISSN: 0022-3077. DOI: 10.1152/jn.1972.35.4.542.

[306] Ziad M. Hafed et al. “Visual Functions of the Primate Superior Colliculus”. eng. In: *Annual Review of Vision Science* 9 (Sept. 2023), pp. 361–383. ISSN: 2374-4650. DOI: 10.1146/annurev-vision-111022-123817.

[307] Hsin-Hung Li, Nina M. Hanning, and Marisa Carrasco. “To look or not to look: dissociating presaccadic and covert spatial attention”. eng. In: *Trends in Neurosciences* 44.8 (Aug. 2021), pp. 669–686. ISSN: 1878-108X. DOI: 10.1016/j.tins.2021.05.002.

[308] R. W. Remington, J. C. Johnston, and S. Yantis. “Involuntary attentional capture by abrupt onsets”. eng. In: *Perception & Psychophysics* 51.3 (Mar. 1992), pp. 279–290. ISSN: 0031-5117. DOI: 10.3758/bf03212254.

[309] J. Theeuwes et al. “Influence of attentional capture on oculomotor control”. eng. In: *Journal of Experimental Psychology. Human Perception and Performance* 25.6 (Dec. 1999), pp. 1595–1608. ISSN: 0096-1523. DOI: 10.1037//0096-1523.25.6.1595.

[310] P. M. Fitts. “The information capacity of the human motor system in controlling the amplitude of movement”. eng. In: *Journal of Experimental Psychology* 47.6 (June 1954), pp. 381–391. ISSN: 0022-1015.

[311] Chia-Chien Wu, Oh-Sang Kwon, and Eileen Kowler. “Fitts’s Law and speed/accuracy trade-offs during sequences of saccades: Implications for strategies of saccadic planning”. eng. In: *Vision Research* 50.21 (Oct. 2010), pp. 2142–2157. ISSN: 1878-5646. DOI: 10.1016/j.visres.2010.08.008.

[312] Roger Ratcliff et al. “Diffusion Decision Model: Current Issues and History”. eng. In: *Trends in Cognitive Sciences* 20.4 (Apr. 2016), pp. 260–281. ISSN: 1879-307X. DOI: 10.1016/j.tics.2016.01.007.

[313] Ziad M. Hafed and Chih-Yang Chen. "Sharper, Stronger, Faster Upper Visual Field Representation in Primate Superior Colliculus". eng. In: *Current biology: CB* 26.13 (July 2016), pp. 1647–1658. ISSN: 1879-0445. DOI: 10.1016/j.cub.2016.04.059.

[314] D. C. Van Essen, W. T. Newsome, and J. H. Maunsell. "The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability". eng. In: *Vision Research* 24.5 (1984), pp. 429–448. ISSN: 0042-6989. DOI: 10.1016/0042-6989(84)90041-5.

[315] Skipper Seabold and Josef Perktold. "statsmodels: Econometric and statistical modeling with python". In: *9th Python in Science Conference*. 2010.

[316] Eleni Psarou et al. "Repetition-related gamma plasticity in macaque V1 and V2 is highly stimulus specific and robust to stimulus set size". 2024.

[317] Antonio Torralba and Aude Oliva. "Statistics of natural image categories". eng. In: *Network (Bristol, England)* 14.3 (Aug. 2003), pp. 391–412. ISSN: 0954-898X.

[318] D. M. Coppola et al. "The distribution of oriented contours in the real world". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 95.7 (Mar. 1998), pp. 4002–4006. ISSN: 0027-8424. DOI: 10.1073/pnas.95.7.4002.

[319] Michael F. Bonner and Russell A. Epstein. "Object representations in the human brain reflect the co-occurrence statistics of vision and language". en. In: *Nature Communications* 12.1 (July 2021), p. 4081. ISSN: 2041-1723. DOI: 10.1038/s41467-021-24368-2. URL: https://www.nature.com/articles/s41467-021-24368-2 (visited on 12/15/2023).

[320] Keith Rayner et al. "The Effects of Frequency and Predictability on Eye Fixations in Reading: Implications for the E-Z Reader Model." en. In: *Journal of Experimental Psychology: Human Perception and Performance* 30.4 (2004), pp. 720–732. ISSN: 1939-1277, 0096-1523. DOI: 10.1037/0096-1523.30.4.720. URL: http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.30.4.720 (visited on 12/15/2023).

[321] Hsueh-Cheng Wang, Alex D. Hwang, and Marc Pomplun. "Object Frequency and Predictability Effects on Eye Fixation Durations in Real-World Scene Viewing". In: *Journal of Eye Movement Research* 3.3 (July 2010). ISSN: 1995-8692. DOI: 10.16910/jemr.3.3.3. URL: https://bop.unibe.ch/JEMR/article/view/2297 (visited on 12/15/2023).

[322] Ignace Th C. Hooge et al. “Inhibition of return is not a foraging facilitator in saccadic search and free viewing”. eng. In: *Vision Research* 45.14 (June 2005), pp. 1901–1908. ISSN: 0042-6989. DOI: 10.1016/j.visres.2005.01.030.

[323] Laurence T. Maloney and Hang Zhang. “Decision-theoretic models of visual perception and action”. en. In: *Vision Research* 50.23 (Nov. 2010), pp. 2362–2374. ISSN: 00426989. DOI: 10.1016/j.visres.2010.09.031. URL: https://linkinghub.elsevier.com/retrieve/pii/S0042698910004761 (visited on 12/15/2023).

[324] Rajesh P. N. Rao and Dana H. Ballard. “Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects”. en. In: *Nature Neuroscience* 2.1 (Jan. 1999), pp. 79–87. ISSN: 1097-6256, 1546-1726. DOI: 10.1038/4580. URL: https://www.nature.com/articles/nn0199_79 (visited on 12/15/2023).

[325] Wolf Singer. “Recurrent dynamics in the cerebral cortex: Integration of sensory evidence with stored knowledge”. en. In: *Proceedings of the National Academy of Sciences* 118.33 (Aug. 2021), e2101043118. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2101043118. URL: https://pnas.org/doi/full/10.1073/pnas.2101043118 (visited on 12/15/2023).

[326] M.W. Spratling. “A review of predictive coding algorithms”. en. In: *Brain and Cognition* 112 (Mar. 2017), pp. 92–97. ISSN: 02782626. DOI: 10.1016/j.bandc.2015.11.003. URL: https://linkinghub.elsevier.com/retrieve/pii/S027826261530035X (visited on 12/15/2023).

[327] Colin Blakemore and Grahame F. Cooper. “Development of the Brain depends on the Visual Environment”. en. In: *Nature* 228.5270 (Oct. 1970), pp. 477–478. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/228477a0. URL: https://www.nature.com/articles/228477a0 (visited on 12/16/2023).

[328] Ralf A. W. Galuske, Matthias H. J. Munk, and Wolf Singer. “Relation between gamma oscillations and neuronal plasticity in the visual cortex”. en. In: *Proceedings of the National Academy of Sciences* 116.46 (Nov. 2019), pp. 23317–23325. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1901277116. URL: https://pnas.org/doi/full/10.1073/pnas.1901277116 (visited on 12/16/2023).

[329] Nathaniel Williams and Carl R. Olson. “Independent repetition suppression in macaque area V2 and inferotemporal cortex”. en. In: *Journal*

of Neurophysiology 128.6 (Dec. 2022), pp. 1421–1434. ISSN: 0022-3077, 1522-1598. DOI: 10.1152/jn.00043.2022. URL: https://journals.physiology.org/doi/10.1152/jn.00043.2022 (visited on 12/16/2023).

[330] Peter M. Lee. *Bayesian statistics: an introduction*. 4th ed. Chichester, West Sussex ; Hoboken, N.J: Wiley, 2012. ISBN: 978-1-118-33257-3.

[331] Dario L. Ringach, Robert M. Shapley, and Michael J. Hawken. "Orientation Selectivity in Macaque V1: Diversity and Laminar Dependence". en. In: *The Journal of Neuroscience* 22.13 (July 2002), pp. 5639–5651. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.22-13-05639.2002. URL: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.22-13-05639.2002 (visited on 12/16/2023).

[332] Christian Beste and Hubert R. Dinse. "Learning without Training". en. In: *Current Biology* 23.11 (June 2013), R489–R499. ISSN: 09609822. DOI: 10.1016/j.cub.2013.04.044. URL: https://linkinghub.elsevier.com/retrieve/pii/S0960982213004855 (visited on 12/16/2023).

[333] Sam F. Cooke and Mark F. Bear. "Visual Experience Induces Long-Term Potentiation in the Primary Visual Cortex". en. In: *The Journal of Neuroscience* 30.48 (Dec. 2010), pp. 16304–16313. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.4333-10.2010. URL: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.4333-10.2010 (visited on 12/16/2023).

[334] T Poggio, M Fahle, and S Edelman. "Fast perceptual learning in visual hyperacuity". en. In: *Science* 256.5059 (May 1992), pp. 1018–1021. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1589770. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.1589770 (visited on 12/16/2023).

[335] Aniek Schoups et al. "Practising orientation identification improves orientation coding in V1 neurons". en. In: *Nature* 412.6846 (Aug. 2001), pp. 549–553. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/35087601. URL: https://www.nature.com/articles/35087601 (visited on 12/16/2023).

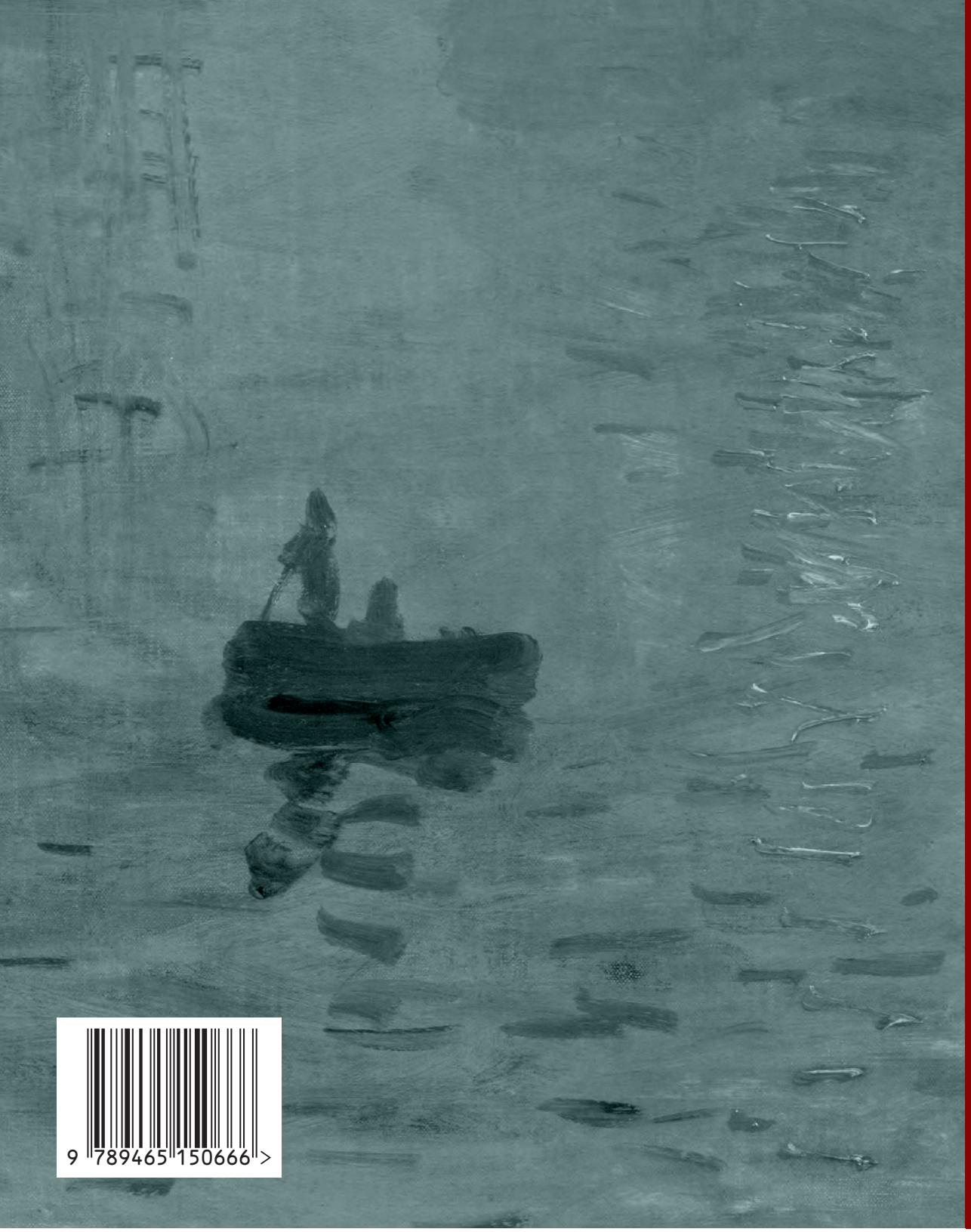
[336] Valérie Wespatat, Frank Tennigkeit, and Wolf Singer. "Phase Sensitivity of Synaptic Modifications in Oscillating Cells of Rat Visual Cortex". en. In: *The Journal of Neuroscience* 24.41 (Oct. 2004), pp. 9067–9075. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.2221-04.2004. URL: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2221-04.2004 (visited on 12/16/2023).

[337] Rens Van De Schoot et al. “Bayesian statistics and modelling”. en. In: *Nature Reviews Methods Primers* 1.1 (Jan. 2021), p. 1. ISSN: 2662-8449. DOI: 10.1038/s43586-020-00001-2. URL: https://www.nature.com/articles/s43586-020-00001-2 (visited on 12/16/2023).

[338] Obed Okwoli Apochi et al. “Virtual, mixed, and augmented realities: A commentary on their significance in cognitive neuroscience and neuropsychology”. eng. In: *Applied Neuropsychology. Adult* (July 2024), pp. 1–4. ISSN: 2327-9109. DOI: 10.1080/23279095.2024.2365870.

[339] Alexander Mathis et al. “DeepLabCut: markerless pose estimation of user-defined body parts with deep learning”. eng. In: *Nature Neuroscience* 21.9 (Sept. 2018), pp. 1281–1289. ISSN: 1546-1726. DOI: 10.1038/s41593-018-0209-y.

[340] Ian H. Stevenson and Konrad P. Kording. “How advances in neural recording affect data analysis”. eng. In: *Nature Neuroscience* 14.2 (Feb. 2011), pp. 139–142. ISSN: 1546-1726. DOI: 10.1038/nn.2731.



9 789465 150666

DONDERS INSTITUTE

Radboud University

Radboudumc