Visual Processing of Real and Virtual Environments With and Without Visual Stress

Xin Chen Cai

PhD

University of York
Biology
January 2025

ABSTRACT

This thesis investigates visual perception across real-world and digital environments through an interdisciplinary approach that integrates neuroscience, psychology, mathematics, and immersive technology. It begins by outlining the fundamental science of light and vision, followed by a review of contemporary research on reading and visual perception. The thesis then examines signal processing techniques commonly applied to electroencephalography (EEG), addressing both their mathematical foundations and practical implementation for analysing non-stationary neural signals.

Building on this framework, a generalised EEG analysis pipeline integrating established time–frequency and spatial filtering methods is developed and validated using a public motor imagery dataset. The validated pipeline is subsequently applied to analyse behavioural and EEG data from a visual perception experiment comparing stimulus presentation across real-world, two-dimensional screen-based, augmented reality, and virtual reality environments. Reaction time measures are used to assess perceptual efficiency, while EEG recordings provide insight into associated neural dynamics.

The thesis concludes with an exploratory investigation of visual stress through a behavioural case study on Irlen Syndrome, examining how digital colour filtering may influence reaction time under varying task demands. Although constrained by sample size, the work demonstrates the feasibility of combining EEG, behavioural measures, and immersive technologies to study visual perception. Overall, this thesis contributes methodological tools and preliminary empirical insights that inform future research on visual processing in modern, technology-mediated environments.

DECLARATION

This is to certify that the thesis work entitled *"Visual Processing of Real and Virtual Environments With and Without Visual Stress"* has been carried out by *Xin Chen Cai* in the Department of *Biology*, University of York, York, England, UK. I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

ACKNOWLEDGEMENT

Thank you to my supervisors for their guidance throughout my PhD

Supervisors

1. Professor Adar Pelah Department of Physics, Engineering and Technology University of York	Supervisor
2. Professor David Halliday Professor Department of Physics, Engineering and Technology University of York	Supervisor
3. Professor Gavin Kearney Professor Department of Physics, Engineering and Technology University of York	TAP advisor
4. Professor Stephen Smith Professor Department of Physics, Engineering and Technology University of York	Progression panel

Special mention

1. Professor Alex Wade
   Department of Psychology
   University of York
   For your invaluable guidance on neuroscience and EEG signal processing

ACKNOWLEDGEMENT

To my partner, Xin Ying Lim,

Thank you for being my constant — for making space where I could be wholly myself, for turning every return into coming home, and for walking beside me as my equal partner through every uncertainty.

January 2025 ***Xin Chen Cai***

Contents

LIST OF FIGURES

LIST OF TABLES

‘‘‘

Chapter 1

Introduction

Visual perception is the process by which the brain interprets and organises visual input from the environment, enabling humans to perceive, understand, and interact with the world around them [18]. This process begins with the capture of light by the eyes, where photoreceptors in the retina convert incoming photons into electrical signals [19]. These signals are transmitted through the visual pathways to the brain, where they are processed across multiple cortical regions, including the occipital, parietal, and temporal lobes [20]. Through this distributed and hierarchical processing, the brain transforms raw sensory input into coherent representations of objects, motion, depth, and colour [18].

Vision is the dominant sensory modality for humans, shaping how we recognise objects [21], faces [22], and spatial relationships [23], and influencing behaviour across a wide range of contexts, from everyday decision-making [24] to scientific reasoning and learning [25]. Understanding how visual information is processed is therefore fundamental to neuroscience and psychology, and has practical implications for domains such as extended reality (XR) technologies [26, 27], driving safety [28, 29, 30], and digital media and art [31].

While visual perception underpins interaction with both physical and digital environments, the mechanisms by which the brain processes visual information are not invariant to how stimuli are presented. Differences in luminance, depth cues, motion parallax, latency, and sensory congruence can all influence perceptual processing. In recent years, the rapid proliferation of screen-based and immersive technologies—ranging from conventional two-dimensional displays to augmented reality (AR) and virtual reality (VR) systems—has raised important questions about whether visually identical stimuli are processed equivalently when encountered in different display modalities. These questions are particularly relevant for individuals who experience visual stress or atypical visual processing, for whom changes in stimulus format may disproportionately affect perceptual efficiency, cognitive load, and behavioural performance.

1.1 Research context and motivation

Extended reality (XR) systems, including VR and AR, have emerged as powerful tools for studying visual perception [26, 27]. Unlike traditional 2D displays, XR platforms allow the presentation of three-dimensional stimuli with precise control over visual parameters such as depth, motion, lighting, and viewpoint [32, 33]. This enables the creation of controlled yet ecologically relevant experimental environments that more closely resemble real-world visual experiences.

XR technologies also make it possible to investigate multisensory integration in dynamic contexts, allowing researchers to examine how the brain combines visual input with auditory, haptic, and motor information to construct a coherent perceptual representation of the environment [32, 34, 35]. As such, XR occupies a unique methodological space between tightly controlled laboratory experiments and naturalistic real-world observation.

Understanding visual processing across different display modalities is therefore both timely and important. As digital and immersive technologies become increasingly embedded in education [36], work [37], healthcare [38], and entertainment[39], it is critical to determine whether and how these platforms alter perceptual and cognitive processing relative to real-world viewing. Reaction time and neural measures offer complementary and sensitive indices of perceptual efficiency and cognitive load, making them well-suited for probing such differences [40, 41].

This thesis is positioned at the intersection of visual neuroscience, cognitive psychology, EEG signal analysis, and immersive digital technologies. By combining behavioural and neural measures within controlled XR paradigms, it seeks to bridge methodological advances with theoretical questions about perception in modern visual environments.

1.2 Visual stress and atypical visual processing

The study of visual perception also provides insight into what occurs when visual processing is disrupted. Research in this area has informed understanding of conditions such as dyslexia [42, 43], visual agnosia [44], and stereoblindness [45]. In dyslexia, for example, difficulties in visual attention, motion processing, or letter discrimination have been linked to reading challenges [46].

Visual stress refers to a collection of perceptual symptoms, including visual discomfort, distortions, headaches, and reduced reading performance, that are triggered or exacerbated by visual stimuli, particularly high-contrast or repetitive patterns. Irlen Syndrome is one condition commonly associated with visual stress, although its mechanisms remain debated and are not fully understood [47].

In this thesis, dyslexia and Irlen Syndrome are considered as motivating case examples rather than the sole focus of investigation. The aim is not to provide clinical diagnosis or treatment, but to use visual stress as a lens through which to examine how differences between digital and non-digital visual stimuli may influence perceptual processing. These populations are particularly relevant for probing such differences because they may exhibit heightened sensitivity to stimulus characteristics that are otherwise subtle in neurotypical individuals.

1.3 Research aims and hypotheses

The overarching aim of this thesis is to investigate how visual stimuli presented across real-world, 2D screen-based, augmented reality, and virtual reality environments are processed by the human visual system. Behavioural reaction time measures and electroencephalography (EEG) are used to examine differences in perceptual and neural responses under conditions with and without visual stress.

The central hypothesis is that the mode of stimulus presentation influences perceptual efficiency and neural processing, such that reaction time and EEG measures will differ between real-world, 2D, AR, and VR environments. It is further hypothesised that these differences may be amplified under conditions associated with visual stress.

These hypotheses are tested through controlled reaction time experiments and EEG analyses that compare responses across display modalities while holding stimulus content constant.

1.4 Experimental approach and methodological rationale

To address these aims, this thesis employed a visual perception experiment combining behavioural reaction time measures with EEG recordings. Participants were tested across four presentation environments: real-world viewing, 2D screens, augmented reality, and virtual reality. In each environment, participants were required to respond as quickly as possible to visual stimuli.

Two experimental paradigms were used: a simple reaction time task, involving detection of a single target stimulus, and a choice reaction time task, requiring discrimination between target and non-target stimuli. These paradigms were chosen to probe perceptual processing under different levels of cognitive demand. EEG was recorded throughout the tasks to capture neural responses associated with stimulus perception and processing.

Reaction time was selected as a behavioural metric due to its sensitivity to perceptual efficiency and cognitive load, while EEG was chosen for its high temporal resolution and ability to capture rapid neural dynamics associated with visual processing [48]. Although techniques such as fMRI offer superior spatial resolution [49], EEG is better suited for studying the timing of perceptual processes and for integration with XR environments.

To ensure accurate interpretation of behavioural and neural data, a series of hardware characterisation experiments were conducted to quantify key technical confounds across presentation environments. These included measurements of luminance, system latency, and electromagnetic noise. This step was essential to distinguish genuine perceptual effects from artefacts introduced by display hardware.

EEG data were analysed using a novel signal processing pipeline that integrated wavelet-based time–frequency analysis with Common Spatial Pattern (CSP) spatial filtering. These two techniques have both been widely used in EEG research, and prior studies have also combined these techniques for feature extraction and classification [50]. Building on this existing work, this thesis integrates wavelet and CSP methods within a unified and generalised EEG analysis pipeline designed for non-stationary neural signals and comparative perceptual experiments across multiple display modalities. The pipeline was validated using a public motor imagery dataset, demonstrating its ability to capture task-related neural activity while also revealing limitations related to spatial resolution and inter-individual variability.

1.5 Scope and limitations of the work

This thesis does not attempt to provide clinical diagnoses, establish definitive EEG biomarkers, or draw strong population-level conclusions about visual stress or immersive perception. Practical constraints, including limited sample sizes, EEG signal sensitivity, and XR hardware limitations, restrict the generalisability of the findings.

Accordingly, inconclusive or null results are interpreted within an exploratory framework. Rather than being viewed as failures, these outcomes provide important information about methodological constraints and inform directions for future research.

1.6 Original contributions of the thesis

This thesis makes several methodological, technical, and empirical contributions to the study of visual perception in digital and immersive environments. These include the development of a modular 2D and 3D stimulus generation framework, the design and validation of a novel EEG signal processing pipeline, systematic hardware characterisation across display modalities, experimental comparison of real-world and immersive visual stimuli, integration of behavioural and neural measures in XR contexts, and exploratory investigation of visual stress effects across display modalities.

Methodological contributions are distinguished from empirical findings, with the former providing tools and frameworks for future research and the latter offering preliminary insights into perceptual processing in immersive environments.

1.7 Structure of the thesis

This thesis is structured as follows. Chapter 2 introduces the basic science of light and visual processing. Chapter 3 reviews literature on reading and visual perception. Chapters 8.1 and 5 present the mathematical foundations and practical implementation of EEG signal processing techniques, culminating in the development and validation of a signal processing pipeline. Chapter 7 describes the hardware characterisation experiments across presentation environments. Chapter 8 presents the main visual perception experiments and their results. Chapter 9 explores visual stress and Irlen Syndrome through an exploratory behavioural case study. Finally, Chapter 10 synthesises the findings, discusses limitations, and outlines directions for future work.

Chapter 2

Theoretical Framework on light and vision

2.1 Introduction

The study of visual processing requires a comprehensive understanding of the complex interactions between physical, biological, and psychological mechanisms that enable perception. Visual perception begins with the fundamental properties of light; thus, it is essential to first examine the physical nature of light and its behaviour. Understanding how light is generated, transmitted, and interacts with various surfaces provides the foundation for interpreting how visual information is encoded in the environment.

Following the discussion of the physical principles of light, it is necessary to explore the biological mechanisms that allow the visual system to capture and process this information. The human visual system, comprising the eyes, optic pathways, and visual centres of the brain, translates light stimuli into neural signals. These signals are then integrated and interpreted, enabling the perception of shapes, colours, motion, and spatial relationships. An understanding of this biological process is crucial for linking the physics of light to the perceptual experiences that arise from it.

Finally, the psychological aspects of visual processing must be considered. Perception is not a direct reflection of sensory input but a complex cognitive process that involves interpretation, organization, and contextual understanding. Psychological theories of visual processing address how the brain constructs meaning from visual stimuli, including how attention, memory, and prior knowledge influence perception.

This chapter presents the theoretical framework underlying visual processing by examining the physics of light, the biological mechanisms of visual perception, and the psychological processes involved in interpreting visual information. Building on this foundation, the thesis aims to explore how the brain processes visual input from different sources, namely a simple LED, a computer screen, and mixed reality (MR) systems. A strong understanding of the emission, capture, and processing of light provides the basis for explaining why visually similar stimuli from different sources can be perceived and processed differently by the brain.

2.2 Physics of light

The concept of light may sound intuitive. Ask any lay person what is light, and they may point to a light source near them, such as a light bulb or a candle flame. However, light is an abstract concept which scientist have studied for centuries, which led to some of the amazing inventions that we use today, such as the light bulb and the camera. This section briefly discusses the important historical milestones in the study of light.

2.2.1 What is light?

In the 18th century, the dominant view of light was Isaac Newton's corpuscular (particle) theory. This theory was based on observations that light travels in straight lines, bounces off surfaces (reflection), and changes direction when passing through different media (refraction), behaviours analogous to those of particles [51]. However, in 1801, Thomas Young's famous double-slit experiment demonstrated the constructive and destructive interference of light, providing strong evidence for its wave-like nature and challenging the particle-only view.

Figure 2.1 shows the modern version of Young's double slit experiment [1]. A distinctive pattern of bright and dark fringes is seen when a monochromatic light is shown on a distant screen after passing through two small slits of openings. The superposition of overlapping light waves coming from the two slits results in this interference pattern. Bright fringes represent regions of constructive interference while dark fringes represent regions of destructive interference.

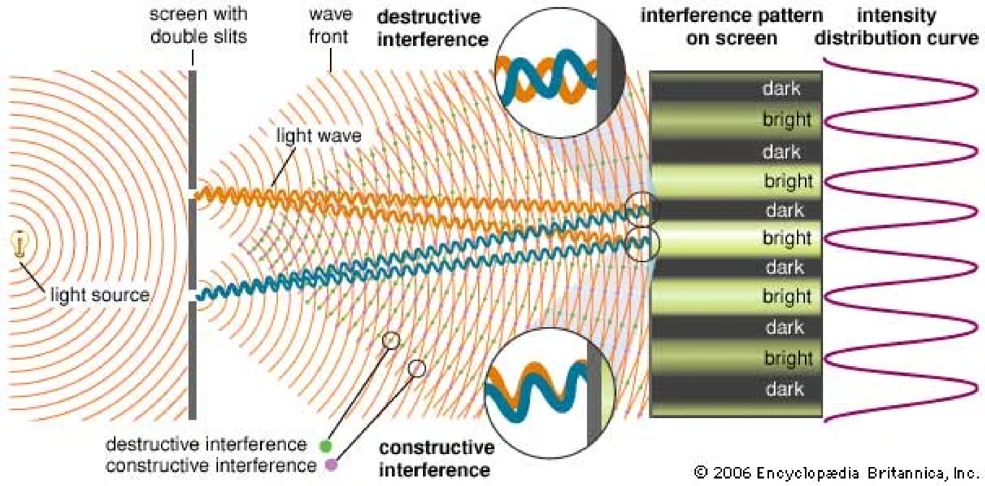


Figure 2.1: Young's double slit experiment. Image obtained from [1].

However, in the late 19th and early 20th centuries, a new theory emerged that challenged the classical wave theory of light. This theory, known as the quantum or photon theory, proposed that light is made up of tiny particles called photons [52]. It was supported by a number of experiments, including the photoelectric effect, which was first observed by Heinrich Hertz in 1887 [53]. This understanding of light to be both particles and wave, also known as, the wave-particle duality of light, persisted till today.

Figure 2.2 shows a simulation of the photoelectric effect, developed by [2]. When light is shown on a metallic surface, electrons are ejected from the surface, and the energy of the emitted electrons is proportional to the frequency of the light. The photoelectric effect played a key role in the development of quantum mechanics, which describes the behaviour of matter and energy at the atomic and subatomic levels.

An understanding in the mechanism of light will aid the understanding of how light is being captured by the eye, and how this light energy is transmitted from the eye to the brain to convey visual information about our world. The next

Subsection will discuss the concept of lens, with is a crucial instrument used by the eye to focus light onto the right spot.

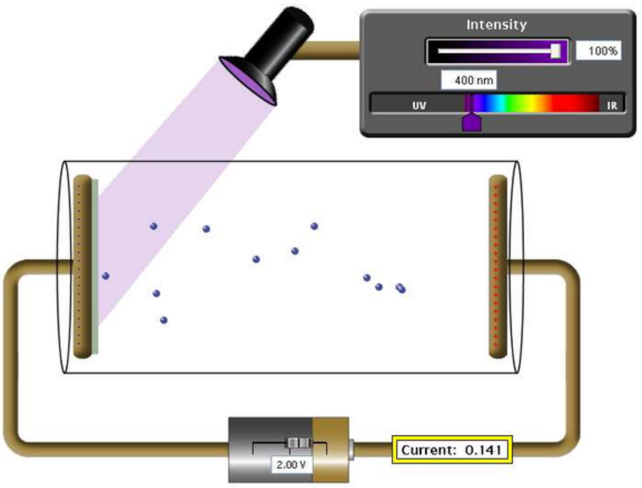


Figure 2.2: Simulation of the photoelectric effect. Image obtained from [2].

2.2.2 How do lenses affect light?

Lenses are transparent objects that refract light and are commonly used in various optical devices such as cameras, telescopes, and eyeglasses. They also exist in eyes, to focus light on the retina.

To understand the physics of lens, the laws of geometrical optics must first be understood. According to [3], there are three laws of geometrical optics:

- The law of rectilinear propagation (transmission): Light travel in a straight line in a medium with a constant refractive index.
- The law of reflection: When a ray of light is reflected at an interface dividing two optical media, the reflected ray remains within the plane of incidence, and the angle of reflection(θ_r) equals the angle of incidence(θ_i).
- The law of refraction (Snell's law). When a ray of light is refracted at an interface dividing two transparent media, the transmitted ray remains within the plane of incidence and the sine of the angle of refraction(θ_t) is directly proportional to the sine of the angle of incidence(θ_i).

Figure 2.3 shows an illustration of the laws of geometrical optics.

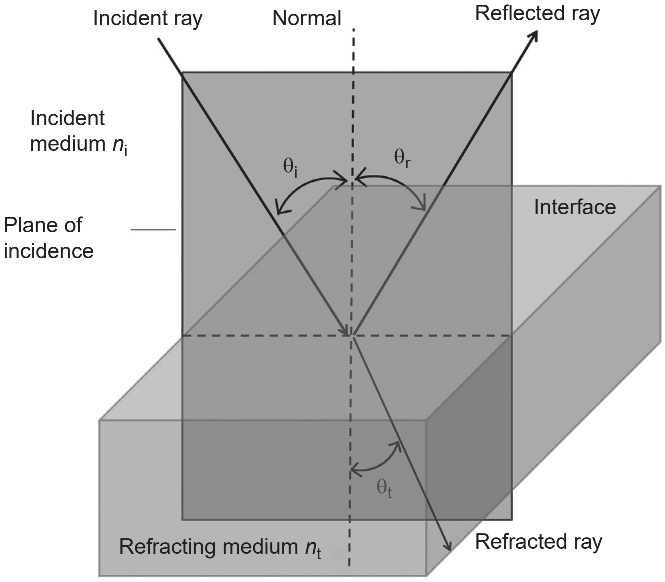


Figure 2.3: Illustration of the laws of geometrical optics. Image obtained from [3].

A lens works by refracting light, or bending its path, as it passes through the curved surface of the lens. The amount of bending depends on the shape of the lens and the angle of incidence of the light. As shown in Figure 2.4, a converging lens, also known as a convex lens, bulges outward in the middle and causes parallel rays of light to converge at a single point, known as the focal point. A diverging lens, also known as a concave lens, bulges inward in the middle and causes parallel rays of light to diverge. The focal point of a concave lens is the point where the diverging rays appear to come from after passing through the lens.

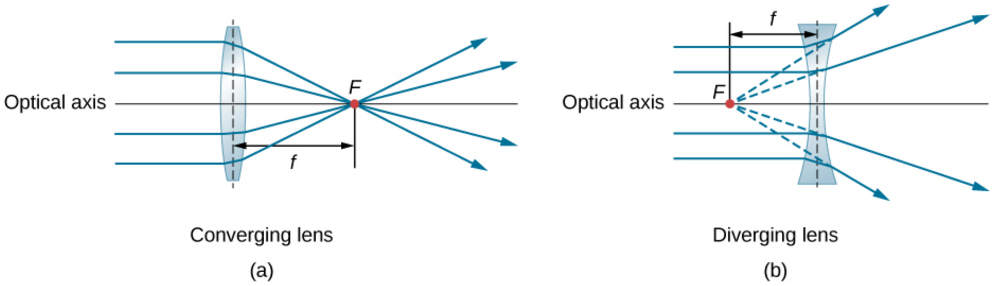


Figure 2.4: Converging and diverging lens. Image obtained from [4].

The mechanism of how the lens in the eye work will be examined in Subsection 2.3.1, where the anatomy and physiology of the eye is discussed.

2.2.3 Colour and wavelength of light

In Subsection 2.2.1, we discussed how light can behave as a wave. This means that light can travel at different wavelengths, and this will result in light of different colours. Thus, in a physical sense, colour is a visual sensation perceived by the human eye and is produced by the different wavelengths of light. A blue object, for example, will reflect blue light while absorbing light of other wavelengths, and a red object will reflect red light while absorbing light of other wavelengths.

Light, according to physics, is a form of electromagnetic radiation, and occupies a small portion of the electromagnetic spectrum. Figure 2.5 shows the full electromagnetic wave spectrum, and the portion which visible light occupies.

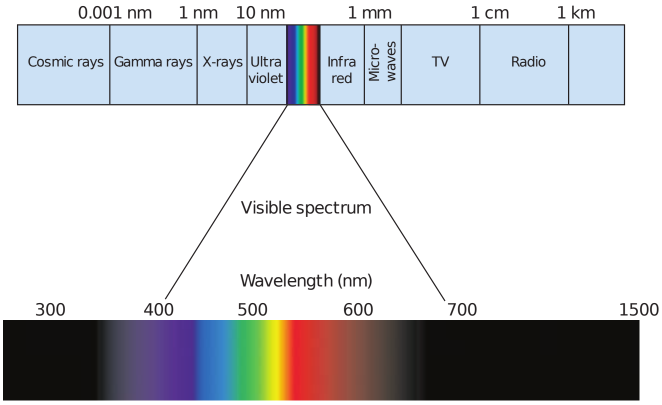


Figure 2.5: Electromagnetic spectrum and visible light. Image obtained from [5].

The visible spectrum of light ranges from approximately 400 to 700 nm in wavelength. As shown in Figure 2.5, the longest wavelength visible to the human eye is red, while the shortest is violet. Colours that fall outside of this range, such as ultraviolet and infrared, are not visible to the human eye but can be detected by specialized instruments.

Beside from wavelength, the perception of colour can also be influenced by factors such as lighting conditions, surrounding colours, and individual differences in colour perception. Additionally, some people may have colour vision deficiencies, which can affect their ability to distinguish certain colors.

In Subsection 2.3.1, we will discuss how receptors in our eyes distinguish between different wavelengths of light.

2.3 Anatomy and physiology of our visual-motor system

The biology of how humans see things around us has fascinated many researchers for centuries. Intuitively, it may seem obvious as to how we can see the objects around us. However, as Snowden et. al. have noted, to see is easy, to be able to interpret what you see, that's vision, and it's complicated. Many books on physiology and neurophysiology have documented the anatomy and physiology of our visual pathways, including the book "Basic Vision" by Snowden et. al [5]. This section will attempt to summarised what Snowden et. al. have written about the anatomy and physiology of our visual pathways, along with any additional research on the topic which are published after the publication of the book.

2.3.1 The eye

Humans have the ability to perceive visual stimuli in the world around them thanks to the presence of light. When light is reflected off an object and towards us, it enters our body through our eye. Before the visual information is even processed, the eye place a crucial role in ensuring that the image we see is as sharp as possible, so that it can be more easily processed downstream. Sagittal cross-section of the human eye is shown in Figure 2.6, and the pathway light moves through the light is shown in Figure 2.7.

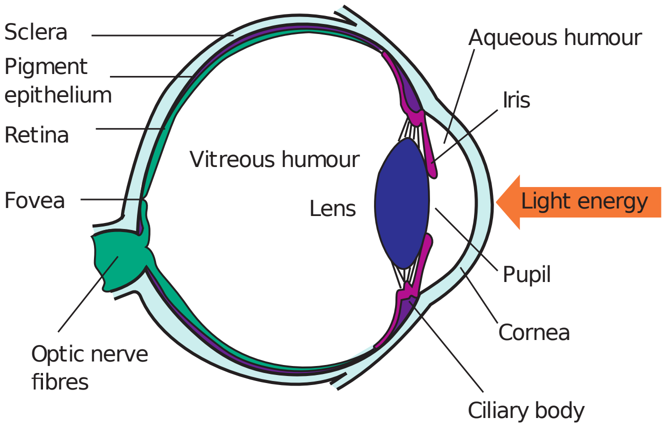


Figure 2.6: Sagittal cross-section of the human eye. Image obtained from [5].

Light first enters our eye through the cornea, which is a transparent outermost

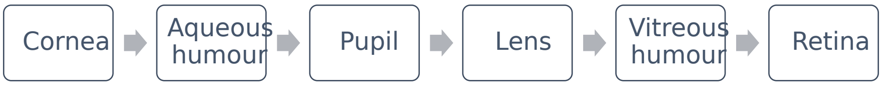


Figure 2.7: Light pathway through the human eye

layer of the eye. It acts as protective layer and refracts the light so that the light is focused on the lens. This refractive power of the cornea accounts for approximately two-thirds of the eye's total optical power. Besides being a refractive layer, the cornea also contains nerve endings which are sensitive to touch, which closes the eyelid when activated.

Next, the light reaches the aqueous humor, which is a watery liquid that flows from the ciliary body. This fluid maintains the structure of the globe of the eye, by exerting hydrostatic pressure on the structure around it to maintain the spherical shape of the eye.

After the aqueous humour, the light passes through the iris, which is a coloured structure with an adjustable aperture, known as the pupil. The size of the pupil can be adjusted by constricting or relaxing the iris.

Behind the iris, the lens, a clear, flexible, and biconvex structure, receives the light and refracts it so that it is focused on the retina. The lens allows us to focus on objects at different distances, ensuring that we can see clearly both up close and far away. This is done by contracting or relaxing the ciliary muscles, improving close vision or distance vision respectively.

The next structure light passes through is the vitreous humour, which is a gelationous subtance that keeps the eyeball in shape and the retina pinned to the back of the eye. The vitreous humour allows for the diffusion of metabolic solutes, and allows the light to reach the retina [54].

Finally, light will reach the retina, which is a thin later of tissue located at the back of the eye. The retina is made up of several structures crucial for sight, including photoreceptor cells. These cells have either a large (magno) or small (parvo) cell body, hence they are called magnocellular (M) and parvocellular (P) cells respectively. Both types of cells are also called rods and cones respectively.

M cells are responsible for vision in low light conditions and are most densely located in the periphery of the retina, while P cells are responsible for color vision and visual acuity and are most densely located in the center of the retina, called the macula. Both types of cells work together to transmit visual information to the brain via the optic nerve, allowing us to see the world around us.

The function of M and P cells are discussed in Subsection 2.3.3, where colour vision is examined.

2.3.2 Optic chiasm

Beyond the retina lies the optic chiasm. This is a where the optic fibers from both eyes meet and cross over to the opposite side of the brain. This crossing of fibers allows the information from both eyes to be integrated and processed together, which is critical for depth perception and other aspects of visual processing.

The two sides of each eye project differently through the optic chaism to the visual cortex. The side medial to the nose is known as nasal hemiretina, while the side lateral to the nose is known as temporal hemiretina. [55]. Figure 2.8 shows the visual pathway from the retina to the visual cortex.

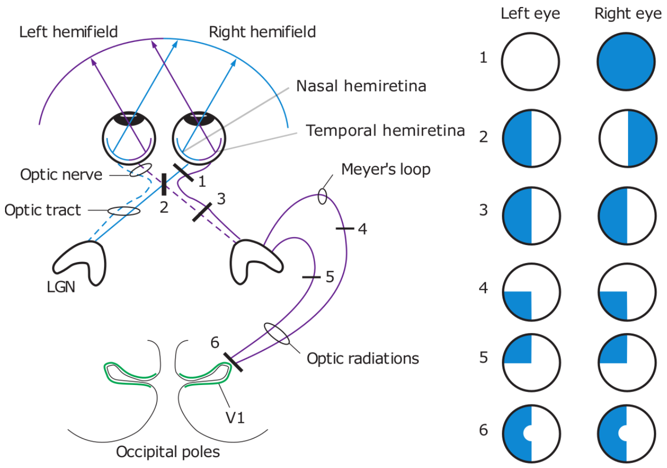


Figure 2.8: Visual pathway from the retina to the visual cortex. The circles on the right indicate the loss of vision which would result from a lesion at the marked region. Image modified from [5].

From Figure 2.8, the optic chiasm is marked by the number 2, where the optic nerves from each eye cross each other. This crossing only occurs with the medial optic nerves, connecting the medial optic nerves to the contralateral side of the brain, while the lateral optic nerves are connected to the ipsilateral side of the brain. As a result of this arrangement, the left half of each retina is connected to the left hemisphere of the brain, while the right half of each retina is connected to the right hemisphere of the brain.

Damage to the optic chiasm can result in a range of visual symptoms, depending on the location and extent of the injury. The circles on the right of Figure 2.8 indicate the resulting visual impairment from a lesion at the corresponding location in the visual pathway. For example, damage to the optic fibers anterior to the optic chaism can result in loss of vision on one eye, while damage to the optic fibers posterior to the optic chaism can result in loss of vision on one side of both eyes.

2.3.3 Lateral geniculate nucleus

The visual pathway connects the retina to the lateral geniculate nucleus (LGN), which is a layered structure located in the posterior part of the thalamus, that receives information from the retina and sends it to the primary visual cortex. Figure 2.10 shows the structure of the LGN.

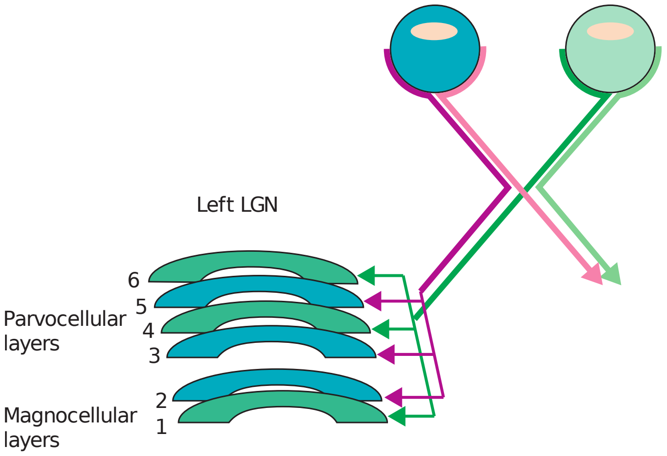


Figure 2.9: Structure of the lateral geniculate nucleus. Image obtained from [5].

The LGN is made up of six layers, and these layers are divided into 4 dorsal parvocellular (P) layers and 2 ventral magnocellular (M) layers. The M layers(layer 1 and 2) are connected to the M cells in the retina, while the P layers (layers 3 to 6) receive information from the P cells in the retina. The retinal cells also project onto different layers depending on their position. Retinal cells afferent from the nasal hemiretina project to layers 1,4 and 6, while retinal cells afferent from the temporal hemiretina project to layers 2,3 and 5.

In 1982, Schiller studied the effects of using neuroblockers to inhibit different layers of the LGN in animals to study how it affects their colour and motion perception. He found that inhibiting neurons in the P layers blocked colour perception, but inhibiting neurons in the M layers blocked motion perception [56].

Today, the understanding about the LGN has expanded from being a relay center for visual signals to a complex structure responsible for high-level visual information processing, such as computational transformation and filtering of visual signals [55]. The LGN is moderately sensitive to orientation, and is also responsive to temporal changes. It is also crucial for regulating the visual signals before they reach the striate cortex [55].

2.3.4 Striate cortex

The striate cortex, also known as the primary visual cortex, V1 or Brodmann area 17, is the largest visual area located in the occipital lobe of the brain. From here, the striate cortex will be referred to as V1. V1 is the first area which receives visual information from the LGN, and is the main area where visual information is processed. It is organised into multiple layers, which projects to other visual areas in the brain. Figure 2.10 shows the input and output to each layer of V1.

In 1968, Hubel and Wiesel studied the V1 in monkeys by placing electrodes in different neurons in the striate cortex. They found that the neurons in different layers of V1 respond selectively to different visual information, and each neuron can be mapped to a specific retinal cell (also known as retinal mapping). These neurons can be classified into simple, complex and hypercomplex cells. Table 2.1

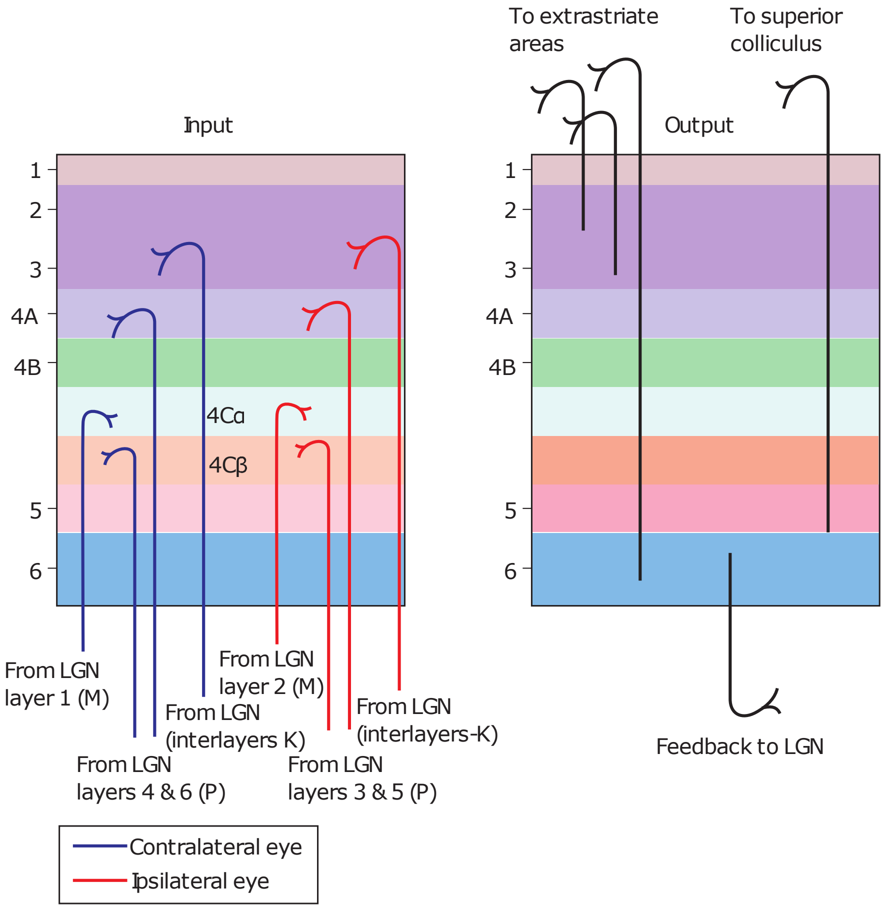


Figure 2.10: Input and output to each layer of the striate cortex. Image obtained from [5].

Table 2.1: Table of functions of each subsection of the occipital lobe.

Cell type	**Function**	**Location**
Simple cell	Detects the presence of stimuli in a specific spot	Layers 3, 4
Complex cell	Detects the orientation of a line	Layers 2, 3, 5, 6
Hypercomplex cell	Detects the orientation and length of a line	Layers 2, 3, 5, 6

shows the function of each type of cell and the layers they are typically found in [57].

In 1998, with the use of functional MRI, Tootell et al. confirmed the existance of retinal mapping in V1. The researchers also extended previous findings about retinal mapping in V1 by demonstrating that the neurons in V1 respond selectively, or are tuned to different features of the visual world, such as edges, lines, and colors [58]. The neurons in more foveal locations each respond to a smaller range of stimuli (shorter orientation bandwidth), while the neurons in the peripheral location each respond to a larger range of stimuli (longer orientation bandwidth), demonstrating that the peripheral vision is inferior to foveal vision.

Two years later, Lamme et al. found that while V1 is necessary for visual awareness (the conscious awareness), it is not required to guide visual behaviour [59]. In an experiment conducted with blindsight patients with a lesion in V1, the researchers found that these patients are unaware of visual stimuli presented to their blind hemifield, but they are still able to correctly guess attributes about the stimuli. This shows that V1 is responsible for visual awareness, but not visual behaviour.

2.3.5 Beyond the striate cortex

The brain region in charge of processing visual information is known as the occipital lobe. The occipital lobe takes up about 12% of the entire volume of the brain, and it sits in the posterior end of the brain, caudal to the parietal lobe. Besides from V1, the occipital lobe also houses many other subsections, the most well known of which are V2, V3, V4 and V5. Each of these regions process specific aspects of vision, and project onto one another to help us understand the visual stimuli presented to us.

While cells in V1 are sensitive to simple stimuli such as position and line orientation, cells in V2 are sensitive to more complex stimuli, such as angles, curves, and non-cartesian gratings, than to conventional stimuli such as bars and sinusoidal gratings. This is done by aggregating the information from V1 in V2 [60]. Huff et al. provided a great summary of the anatomy and physiology of both V1 and V2 [61]. The researchers explained that while V1 processes simple information, this information is then passed onto V2 for more complex processing. V2 then sends feedback information to V1, and sends feedforward information to V3-V5. Damage to the visual cortex can result in different forms of visual blindness, depending on the location of the lesion.

V3, a visual area with ambiguous margins, is an area with a group of cells that are sensitive to the shape of visual stimuli. Rosa and Manger published a review of the research which studied the role of V3 in visual perception [62]. Besides from processing shape, V3 is also known for processing form. This was evident from research demonstrating additional physiological response properties that align with this function, such as selectivity for slow-drifting or stationary objects in the plane of fixation and at far distances, as well as sensitivity to small objects obscured by visual noise.

In their review article about the V4 area, Ikei et al. concluded from neurophysiological studies that the V4 plays a critical role in object recognition and visual perception [63]. V4, according to them, is a higher-level visual area that is involved in the analysis of global image properties, such as colour, texture, and shape, as well as the integration of these properties into object representations. This is done by receiving feedforward input from V2, which carries information about the local image features, such as orientation, spatial frequency and edges.

V4 then uses this information to compute more complex image properties, such as the colour and texture of surfaces, the shape and size of objects, and the context and meaning of scenes.

V5, also known as MT (middle temporal), is a visual area in the human brain that is specialised in processing visual motion information. The V5 area's involvement in the perception of motion, including direction, speed, and acceleration, has been well established throughout the literature on vision research [64, 65, 66, 67]. Damage to V5 can result in motion blindness, or akinetopsia, which is the inability to perceive visual motion [67, 65]. With the use of modern brain imaging techniques, researchers can precisely pinpoint the areas of the visual cortex which are involved in vision perception. Although other areas of the visual cortex are also involved, V5 has consistently been found to be the main area involved in motion perception [68].

Another area which is in-charge of processing visual motion information is the area V6. Both V5 and V6 perform different but overlapping functions. The functional distinction between V5 and V6 becomes apparent when complex visual stimuli are presented to subjects. In a study by [69], the researchers first presented stimuli consisting of concentric rings that expanded or contracted in a radial direction, creating the perception of motion (Radial ring). Next, the stimulus consisted of a field of dots that moved coherently in a pattern that simulated the flow of visual information that would be experienced when self-motion was presented (Flow field). They found that V5 is more sensitive to the Radial ring simulation, while V6 was more sensitive to the Flow field simulations. This experiment demonstrated how V5 and V6 process different types of visual information.

Some of the stimuli processed by each subsection have been summarised in Table 2.2.

In species that rely on vision for their daily functioning, a large portion of their brain is devoted to visual processing. For example, over half of the cerebral cortex

Table 2.2: Table of functions of each subsection of the occipital lobe.

Subsection	**Function**
V1	Position and orientation of lines [58]
V2	Angles, curves and noncartesian gratings [60, 61]
V3	Shape, texture and form [62]
V4	Colour, texture and shape to object representations [63]
V5	Motion [67]
V6	Motion [69]

in both New and Old World monkeys is involved in aspects of visual processing, including visuo-motor and multi-sensory integration [62]. However, while there is much agreement on the general arrangement of the visual cortex, based on their functional contribution to visual processing, in between species, there likely exist differences between specific areas used for a specific task, especially in the extrastriate areas. Hence, it is challenging to generalise a single animal model of the visual cortex to other species, such as humans.

2.4 How different visual areas work together to help us make sense of the world around us

The neural mechanisms of visual perception can be overwhelmingly complex, and we have only scratched the surface in figuring out how we perceive visual information. However, we can develop models that help us simplify and better understand the world around us. A proposed model of visual object recognition is presented in Figure 2.11.

When the signal from the retinal cells first reach V1, the neural cells in V1 process the orientation and position of each line on the object. These lines are then put together in V2 to get contours, edges and patterns. V3 will then take these information from V2 to identify the shapes that make up the object. Finally, V4 will put these shapes together into a 3D image of the object, and recognise the object as an wooden cube.

It is important to note that this model, while there's some truth in it, is a

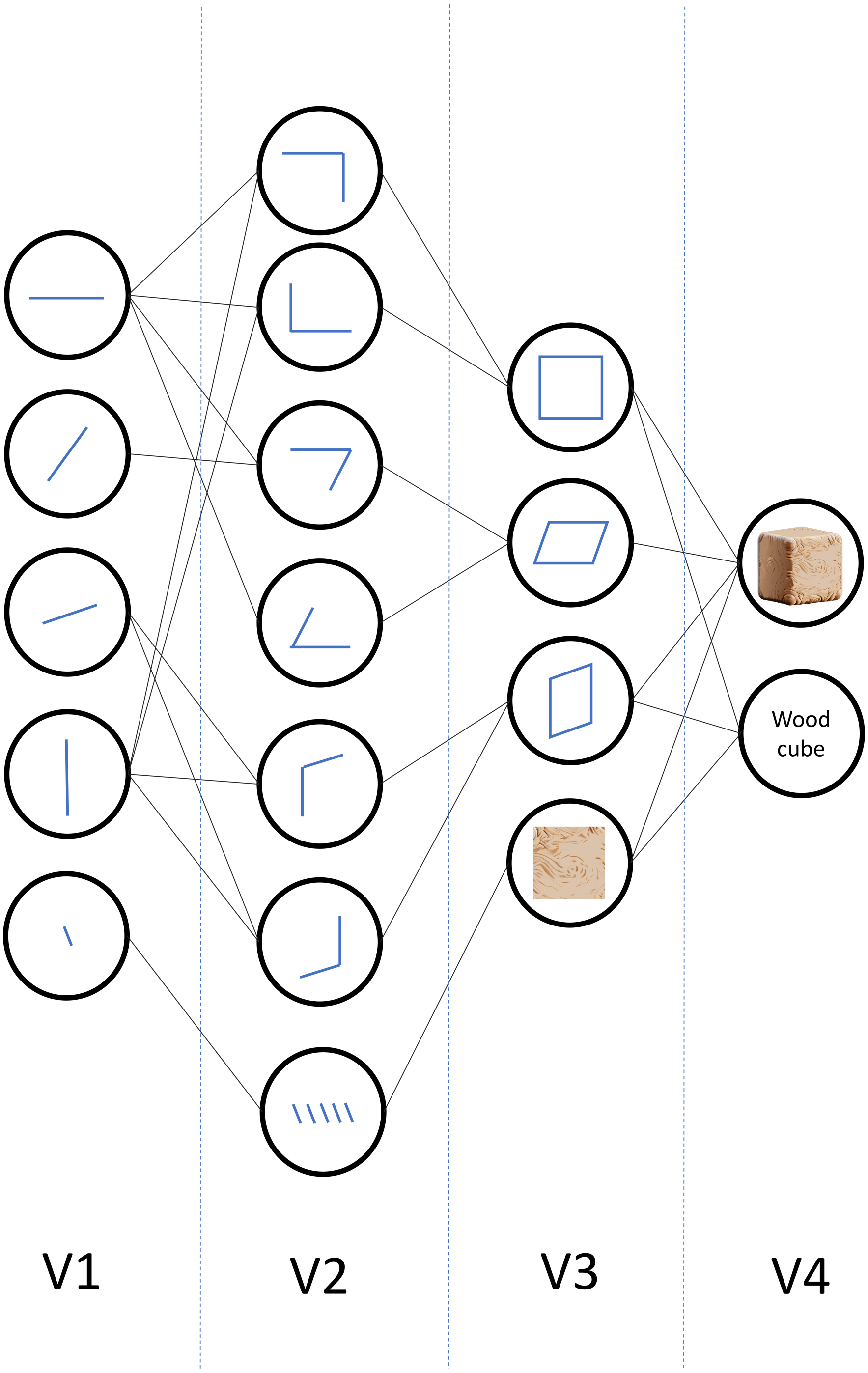


Figure 2.11: Proposed model of visual object recognition.

gross oversimplification of the mechanisms involved in our visual perception.

2.5 Reaction time

One method to study any possible differences between the visual perception of digital objects compared to non-digital objects is to study the reaction time (RT) to the presentation of each type of stimulus. RT in visual perception reflects the temporal dynamics that link sensory input to motor output. It is shaped by a cascade of biological and psychological processes that begin with the transduction of photons in the retina and culminate in the execution of an overt response.

2.5.1 Biological Determinants of Visual Reaction Time

Reaction time (RT) is fundamentally constrained by the biological architecture of the nervous system [70]. Although psychological and contextual factors influence performance, the minimum latency with which an individual can respond to a stimulus is determined by physiological processes, including neural conduction, neural pathways and motor execution.

The speed of neural conduction influences RT. Action potentials propagate more rapidly along axons that are highly myelinated and larger in diameter, due to the increased efficiency of saltatory conduction [71]. Myelination develops progressively into early adulthood and declines with age [72], contributing to well-documented age-related slowing of RT [73]. Disorders that affect myelin integrity, such as multiple sclerosis, also illustrate how disruptions to conduction velocity can significantly lengthen the response latency [74, 75]. Thus, conduction speed establishes the fundamental temporal boundaries within which higher-level processes must operate, making it a lower-bound constraint of reaction time.

Visual information travels through parallel neural pathways that differ in speed and function. The magnocellular pathway, which prioritises high temporal resolution and motion detection, supports faster RTs compared to the parvocellular pathway, which processes fine spatial detail and colour [76, 77]. The sequential relay of information from retina to LGN, V1, and higher-order regions in the dorsal and ventral streams introduces additional processing delays. The efficiency of connectivity between these regions, as well as the degree of synchronization in neural firing, influences both the speed and stability of perceptual processing, which in turn influences reaction time [78, 79].

RT can also be affected when neural signals are converted into overt movement. This is shaped by the electromechanical delay associated with muscle fibre activation [80]. The proportion of fast- versus slow-twitch fibres can also influence how fast a movement can be initiated or produced once a neuro command arrives [81]. Younger individuals and those with greater neuromuscular efficiency generally demonstrate shorter motor execution times [82, 83, 84]. Although motor delays contribute less to RT than sensory or cognitive stages, they remain an important biological component of overall response latency.

2.5.2 Psychological Determinants of Visual Reaction Time

Besides from biological factors, RT can also be influenced by psychological factors. Psychology can explain how RT can be affected by internal and external factors, such as cognition, arousal and prior knowledge.

Cognitive models conceptualise RT as the sum of multiple processing stages. Sequential stage models propose distinct encoding, decision, and response execution stages [85]. In contrast, parallel distributed processing models emphasise the overlapping interactions between perceptual and decisional processes [86]. A dominant approach in RT research is the evidence accumulation framework, such as the drift-diffusion model (DDM). Here, sensory evidence is accumulated over time until a threshold is reached, triggering a response. Faster RTs arise when evidence is strong (high signal-to-noise) or when decisional thresholds are set lower, whereas uncertainty, low contrast, or ambiguous stimuli slow the accumulation process [87].

Arousal also influences reaction time by modulating the efficiency of perceptual, attentional, and motor processes. According to the classic Yerkes–Dodson law, moderate levels of arousal generally optimize performance, facilitating faster RTs, whereas very low or very high arousal can impair processing and slow responses [88]. While low arousal limits the allocation of attentional resources and the readiness of the motor system, excessive arousal can induce distractibility, anxiety, or overactivation, which interferes with focused processing and slows decision-making. Empirical studies support these effects. For example, mild stressors, caffeine, or motivational incentives, often lead to increased RTs in simple and choice reaction tasks [89, 90, 91, 92], whereas extreme stress [93, 94] or sleep deprivation [95, 96, 97] prolongs RTs. Overall, arousal interacts with both sensory processing and decisional stages, highlighting its pervasive influence on the speed and efficiency of cognitive performance.

Expectancy and prior knowledge also exert a powerful influence on RT by shaping how sensory information is processed and how decisional mechanisms are engaged. When participants have predictive cues about when a stimulus will appear (temporal expectancy) or where it will appear (spatial expectancy), they can pre-allocate attentional resources to the expected moment or location, effectively "priming" perceptual and motor systems [98]. For example, in RT experiments, if participants were asked to react to the same stimuli over multiple iterations, their RT would improve over time. Conversely, if they were asked to react to different randomised stimuli, their RT would stay relatively consistent over time [99, 100, 101, 102]. This is because the cognitive system has to reorient attention to the new stimuli before committing to a response, which lengthens RT. This is why in RT experiments, multiple repeated stimuli are presented in random order at random intervals, rather than sequentially at fixed intervals, so that the participant do not get used to each stimuli before it changes.

2.5.3 Integrating Biological and Psychological Perspectives

Biological and psychological determinants of RT are not independent but interact dynamically. Neural architecture constrains the temporal limits of perception, while psychological factors modulate how quickly and efficiently these biological pathways operate. For example, arousal is modulated by brain regions such as the amygdala, hypothalamus, and brainstem arousal systems, which controls activity of neurotransmitters like dopamine, norepinephrine, and acetylcholine [103]. Retinal mapping also selectively enhance the retinotopic region of visual areas that corresponds to the parts of the visual field that are being focused on, increasing attention on those areas [104]. Together, these intertwined influences demonstrate that reaction time emerges from the seamless integration of biological constraints and psychological modulation within the perceptual–cognitive system.

2.6 Conclusion

In summary, this chapter discussed the journey light takes from our eye and to our visual cortex, where it is processed and interpreted. This theoretical knowledge would be the foundation for the understanding on how similar objects produced by different light sources (LED, digital screen or MR headset) can be perceived differently by the brain. The biological and psychological determinants of reaction times were also discussed. In the next chapter, this knowledge will be extended as we discuss how we perceive printed words during reading.

Chapter 3

Literature review on reading and visual perception

3.1 Introduction

The preceding chapter examined the neurobiological mechanisms underlying visual perception. Visual perception supports the recognition of diverse visual stimuli, including lines, shapes, colours, and written text. Since the broader aim of this thesis is to elucidate how visual perception may differ when individuals engage with digital compared to non-digital environments, it is necessary to devote some attention toward the processing of text as a specific stimulus type. Understanding text perception is particularly valuable because reading involves complex interactions between low-level visual analysis and higher-order linguistic interpretation, making it an informative model for investigating how digital presentation formats might alter perceptual efficiency, cognitive load, and neural activation patterns.

Reading is a skill paramount for the transfer of knowledge. This is how the reader of this thesis can read about the research that has been conducted by the author. Discerning every aspect of reading is an enormous undertaking. Topics that would have to be discussed include writing systems, linguistics, speech, phonetics, memory formation, and many more. On top of that, it is not sufficient to just examine the language of English. There are many languages in the world that have a vastly different written and spoken language system compared to English. Hence, this chapter will focus on reading as a visually driven behaviour in order to establish a foundation for later comparisons between traditional and digital media. For a more comprehensive treatment on the science of reading, readers are encouraged to read [105].

3.2 Eye movement during reading

During reading, the eyes move across the page, during which they execute a series of rapid, jerky movements, separated from the next movement by fixation pauses. These movements are also known as saccade movements. During the fixation pauses, the eyes are able to gather visual information from the page [71].

Researchers have been interested in the relationship between eye movement and reading ability since the dawn of the field of psychological research. Rayner (1978) has written a comprehensive review on research into eye movement before the 1970s. The themes explored include perceptual span, eye guidance, integration across saccades, control of fixation durations, individual differences, and eye movements as they relate to dyslexia and speed reading [106]. During this period, the focus was on the observable motion of the eye.

By the 1980s, the focus of eye movement research shifted to the cognitive and lexical processing of readers. During this period, much evidence has been collected to link eye movement data to cognitive processes occurring in a particular task [107]. Morrison (1984) proposed a model to explain eye movements during reading, focusing on the interaction between cognition and visual input [108]. The proposed model posits that, as readers process text, their gaze is guided by the need to extract and integrate information efficiently. Hence, eye movements are not random, but are closely tied to the cognitive processing of the text. Fixations occur on areas of high informational content, while saccades move the eyes to new text regions for further processing. Figure 3.1 summarises the model proposed by [108].

In this model, the word on the page first enters the reader's eyes and then the brain's visual network, where the visual information of the word is first processed. The information is then passed to the high-level cognitive processing area of the

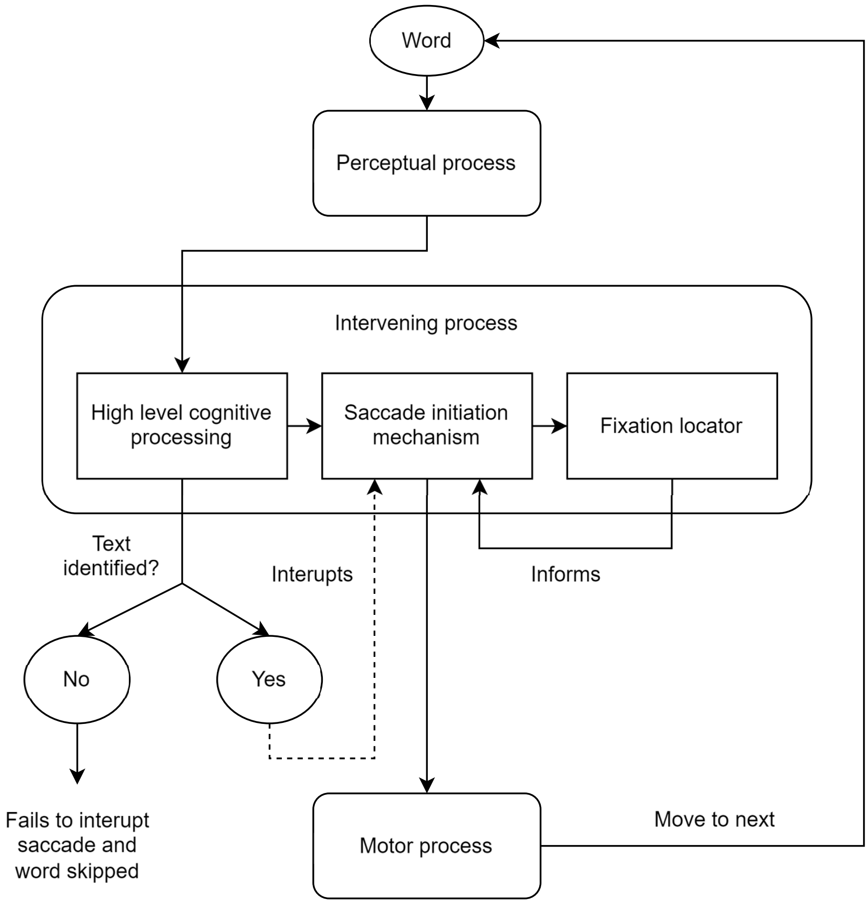


Figure 3.1: Morrison's model of eye movement during reading

brain. Here, depending on how fast the word is processed, the word may or may not be comprehended. If it has not been processed in time, the saccade initiation mechanism is activated, which uses the fixation locator to locate the next fixation point. Then a saccade is activated by motor processes to move the eye to focus on the next word. If the word is processed in time, then the saccade initiation mechanism is interrupted until the cognitive process decides on the next course of action.

While Morrison's model provides a structured explanation of how eye movements facilitate fluent reading while accommodating variability in text complexity and reader expertise, it assumes that reading is a linear process, and readers do not return to previously read text, also known as refixation. [6] proposed a model to explain eye movement during reading which accounts for refixation, known as the E-Z readers model. This model is a computational framework designed to explain eye movement control in reading by relating cognitive processing, particularly lexical access, to eye movements. It has evolved through several versions [107], and the final version (iteration 7) posits that the processes of shifting attention and programming eye movements are two separate but parallel processes.

In this model, visual information from the eye is first divided into two categories during early processing: high spatial frequency (HSF) information (words) and low spatial frequency (LSF) information (such as spaces between words). HSF information is passed to the word identification system through the attentional selection process, while LSF information is passed to the oculomotor system. In the word identification system, the HSF information goes through two stages of lexical processing. The first stage (L_1) signals to the oculomotor system to begin programming the saccade for the next word. The second stage (L_2) shifts the attention to the next word. In the oculomotor system, the saccade is programmed through two stages: labile stage (M_1) and non-labile stage (M_2). (M_1) can be cancelled and restarted by the word identification system when the word is recognised, while (M_2) cannot be interrupted. The saccade is then generated,

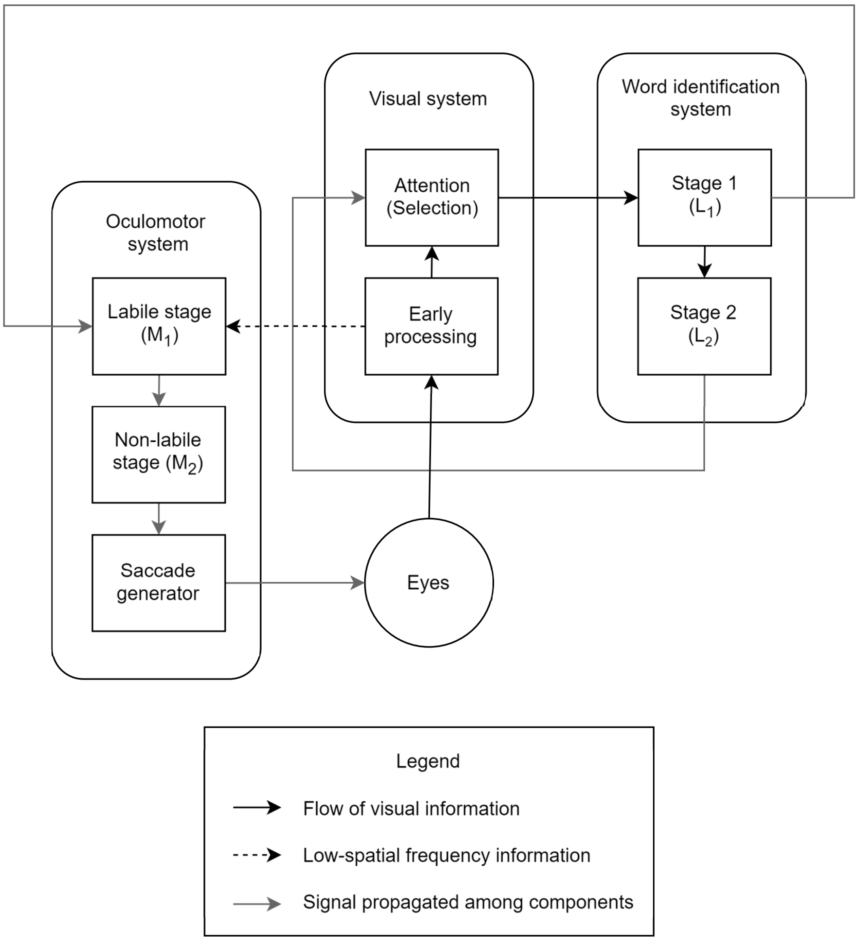


Figure 3.2: E-Z readers model, a computationalframework to explain eye movement control in reading [6].

shifting the eyes' focus to the next word, and the cycle is repeated.

In summary, Reichle's model sheds light on how saccades are programmed, and how the processes are decoupled from word comprehension, allowing for a more nuanced understanding of how readers navigate text. It also emphasizes the importance of lexical access from memory in influencing eye movement patterns [6].

3.3 Word recognition

In English, letters are the most basic unit of reading. They form words, which form phrases. The phrases then form sentences, and the sentences then form paragraphs. The following diagram illustrates how each unit of reading builds on the other. Figure 3.3 illustrate how each reading unit builds up to the other.

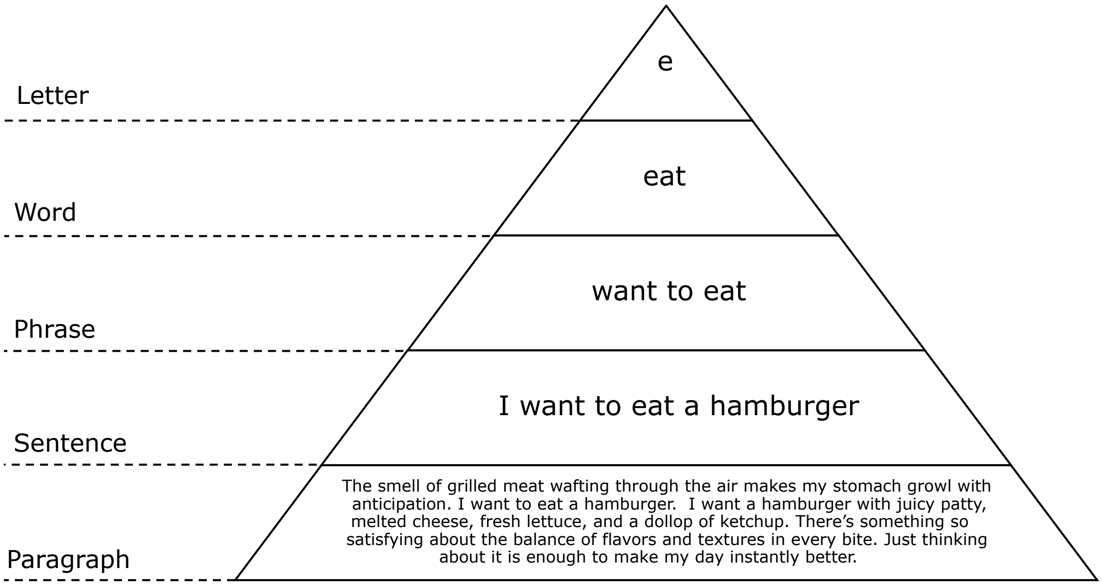


Figure 3.3: Pyramid of the reading units in English. Letters build a word, words build a phrase, phrases build a sentence and sentences build a paragraph.

An important research question in visual perception has been: how do English readers recognise each of the reading units of English? In particular, how do we perceive, identify, and understand written words? This process begins with visual word recognition, where the brain decodes a series of letters into a meaningful unit. The key to this process is the automaticity of word recognition, which allows skilled readers to identify familiar words quickly and effortlessly without needing to consciously analyse individual letters or sounds. This automaticity is developed through repeated exposure to words, enabling readers to rely on mental representations of words stored in memory.

One key concept in the study of reading is orthographic knowledge, which refers to the understanding of how spoken language is written in print [109]. Readers use this knowledge to recognize familiar letter combinations (e.g., “th“ or “ing“) and to predict plausible words even when encountering new or partially obscured text. Apel and Masterson (2019) explored orthographic knowledge as a foundational component of reading and spelling, emphasizing its two distinct levels: lexical orthographic knowledge, which involves stored representations of known words or word parts, and sublexical orthographic knowledge, which encompasses rules and patterns for representing sounds with letters. It highlights that sublexical knowledge often precedes lexical knowledge, as evidenced by early spelling attempts in children and their implicit awareness of orthographic rules [109].

There is a debate on whether orthographic knowledge should be seen as an accumulation of word-specific knowledge in a lexicon acquired through decoding or as a well-developed associative network of sublexical units. In other words, whether orthographic knowledge is accumulated in series or in parallel. In serial models of reading, such as the Dual Route Cascaded (DRC) model, word-specific knowledge in a lexicon is acquired through decoding, a process in which individual letters or letter sequences are processed sequentially to map orthography to phonology [110]. For example, say a reader wants to read the sentence “I want to eat a hamburger“. Figure 3.4 demonstrated how the sentence would have been processed by a serial reader.

In contrast, the parallel model of reading, such as the Parallel Distributed Processing (PDP) model, posits that word-specific knowledge is acquired through a well-developed associative network of sublexical units. Here, sublexical units are processed simultaneously in a distributed manner, with orthographic, phonolog-

Figure 3.4: Serial models of reading. For a serial reader, as they read through the text, they process the phonology of each letter or letter sequences before proceeding to the next.

ical, and lexical information interacting dynamically. This associative network enables the rapid recognition of words as patterns, bypassing the need for sequential decoding. Knowledge in this model is built through the strengthening of connections across these interacting sublexical units over time [110]. Figure 3.5 demonstrated how the same sentence would have been processed by a parallel reader.

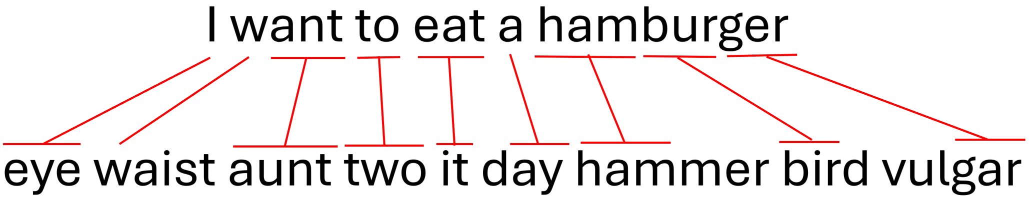


Figure 3.5: Parallel models of reading. For a parallel reader, they can relate each letter sequences with other words containing similar sounds. This dynamic processing of words allow for more efficint reading.

Van den Boer and Jong (2015) examined the conflict between serial and parallel models of reading, particularly in the context of children's development of word and non-word reading fluency [110]. The study identifies two classes of readers: serial processors and parallel processors. Importantly, the results suggest a developmental trajectory where children shift from serial decoding strategies to parallel processing as their reading proficiency improves. This developmental shift challenges both models, as neither fully accounts for the observed progression, particularly the finding that even non-words can be processed in parallel by advanced readers.

There are other factors that can influence word recognition, including word frequency and context. Becker (1979) highlights how high-frequency words are recognized faster and with fewer errors because their detectors in the mental lexicon require fewer sensory features to activate, making them more easily accessible [111]. Semantic context, such as a related word or sentence fragment, facilitates recognition by priming word detectors with relevant semantic features, reducing the sensory information required for activation. Importantly, the paper shows that the effect of context is greater for low-frequency words than for high-frequency words. This interaction suggests that semantic context compensates for the weaker lexical representation of low-frequency words, enhancing their accessibility. Overall, word frequency and semantic context work together to streamline word recognition, particularly under conditions where individual factors might otherwise hinder processing.

In recent years, computational models of visual word recognition have been developed, which provide a detailed explanation of the mechanism of how words are recognised. Norris (2013) published a comprehensive review of computational models of visual word recognition, tracing their evolution from early “box-and-arrow“ approaches to modern connectionist and mathematical models[8]. It emphasizes how these models simulate the core process of recognizing words as visual objects, enabling empirical investigations into phenomena such as lexical decision, masked priming, and word frequency effects. The paper highlights the advancements in model complexity, from basic frameworks handling only small lexicons of fixed-length words to sophisticated systems capable of simulating large-scale data sets and reaction-time distributions. Models like the Spatial Coding Model and the Bayesian Reader illustrate different theoretical approaches, such as connectionist networks and probabilistic decision-making, respectively. The review underscores the challenges of integrating insights across tasks and phenomena, particularly when testing word recognition theories against real-world data.

In summary, word recognition is a complex interplay of visual processing, orthographic and phonological knowledge, and the mental representations of words stored in memory. These processes enable rapid and efficient decoding of written language, forming the foundation for fluent reading and comprehension.

3.4 Higher level semantic processing

While the above models explain how words are perceived and recognised, they do not explain how words are comprehended. While comprehension and semantic processing is beyond the scope of this thesis, for the sake of completeness, this section will briefly discuss some of the models used to explain what happens after words are perceived and recognised.

[7] developed one of the most influential theories on the semantic processing of text, known as the construction-integration theory of text comprehension (CI-model). The CI-model posits that word recognition is achieved through a dynamic, interactive process where visual, orthographic, and lexical information is processed simultaneously. Words, letters, and features are represented as nodes in a network, and their activation levels are determined by interactions with neighboring nodes through inhibitory and excitatory connections. This model explains phenomena like word superiority (better recognition of letters in the context of a word) and masked priming, emphasizing competition among words based on orthographic similarity [8, 7]. Figure 3.6 summarises the CI model:

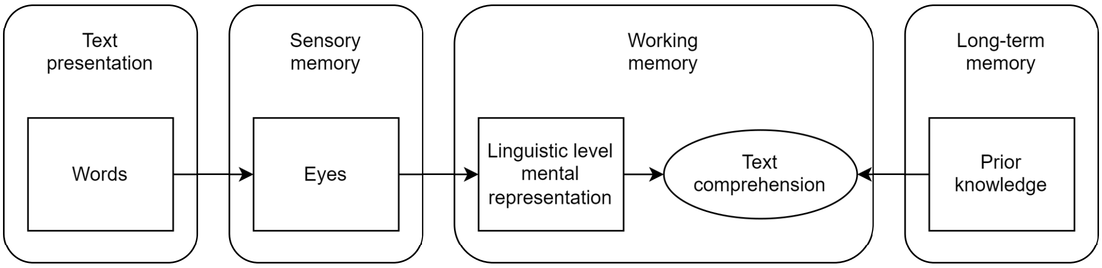


Figure 3.6: CI-model, developed by [7], figure adapted from [8]

Kintsch (1988) developed a model which builds on the CI-model by incorporating additional mechanisms to address its limitations[7]. Specifically, the model extends the CI framework to handle varying word lengths, masked priming effects, and more detailed representations of letter order, such as through spatial coding. Furthermore, Kintsch's approach integrates probabilistic decision-making processes, influenced by Bayesian principles, to model word recognition as an optimal evidence accumulation process. This adaptation allows the model to simulate a broader range of reading phenomena and refine its predictions in complex tasks like lexical decision and masked priming.

3.5 Neurobiology and pathology of language processing

Language is a construct produced by the human brain, so having discussed the cognitive aspect of reading, the neuroscience of reading will be examined next. The neuroscience of reading focuses on how the brain processes written language, converting visual symbols into meaningful representations. Reading relies on the integration of multiple neural systems, with key regions in the occipital, temporal, parietal, and frontal lobes playing distinct roles.

Many research has shown that the visual word form area (VWFA) plays a crucial role in the processing of written words. Located in the mid-fusiform gyrus of the left hemisphere of the brain, the VWFA is specifically involved in recognizing and interpreting written language, allowing individuals to identify words quickly and efficiently [112]. According to [112], the VWFA shows selectivity for visually presented words, meaning it responds more strongly to words than to other types of visual stimuli. This area is thought to be essential for the development of reading skills, as it helps in the transformation of visual input (the shapes of letters and words) into meaningful linguistic representations.

One interesting observation made by Davies-Thompson et. al. (2016) is how the VWFA interacts with the Fusiform Face Area (FFA) [112]. Both the VWFA and the FFA are located in the fusiform gyrus, with the FFA located in the right hemisphere and the VWFA in the left. Given their proximity, it is reasonable to assume that the activation of VWFA would affect the activation of FFA, and vice versa. Davies-Thompson et. al. found that these areas exhibit positive correlations in activation, suggesting that when one is engaged, the other may also show increased activity, rather than an inverse relationship as previously hypothesized. This interaction implies that while they serve distinct functions, their processing may be influenced by broader cognitive factors, such as attention, highlighting the complexity of their relationship in visual recognition tasks.

Kearns et. al. (2019) published a comprehensive tutorial on the neurobiology during reading [9]. In typical readers, the brain processes reading through two primary pathways: the dorsal pathway and the ventral pathway. The dorsal pathway, often referred to as the decoding pathway, is crucial for linking letters to their corresponding sounds, enabling the pronunciation of unfamiliar or novel words. This pathway involves regions like the temporo-parietal cortex (including the angular gyrus and supramarginal gyrus), which connect letters or letter sequences to phonemes, and the inferior frontal gyrus, which sequences and articulates these sounds. The ventral pathway, or the sight recognition pathway, processes familiar words through rapid recognition in the occipito-temporal region, including the fusiform gyrus (sometimes termed the visual word form area). This region allows for efficient word recognition by accessing stored representations of words and linking them to meaning and pronunciation. Figure 3.7 illustrates the dorsal and ventral pathway that are activated during reading:

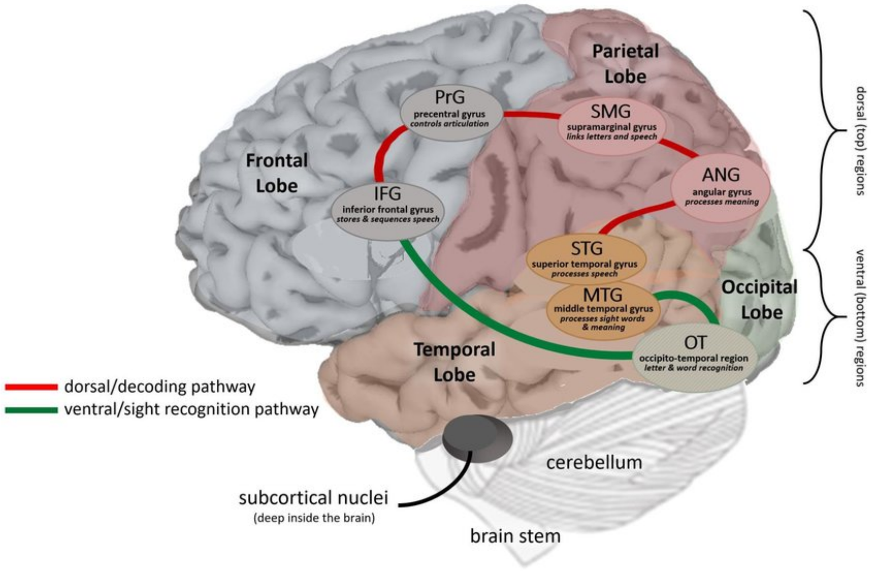


Figure 3.7: Dorsal and ventral pathway of the reading brain [9]

Kearns et. al. corroborates with Van den Boer's study, with different terms used for similar concepts. The step-by-step decoding process of the dorsal pathway aligns with the serial model of reading, while the holistic processing of words in the ventral pathway aligns with the parallel model of reading. Just as Van den Boer and Jong had concluded, Kearns et. al. also found that reading typically involves a dynamic interaction between the two pathways [110, 9]. Early readers and challenging words activate the dorsal pathway and serial processes, while skilled readers shift to ventral dominance and parallel processing for familiar words. Both pathways work together to enable fluent reading.

Another consideration in the neurobiology of reading is how it's neural pathway differs from speech. According to Buchweitz et. al. (2009), there are distinct cortical areas that are specialized for processing the respective modalities (visual or auditory) before the information converges in the amodal language network [113]. Figure 3.8 summarises Buchweitz's work.

Figure 3.8 illustrates how language information is processed in the brain. Language information can enter the brain through visual or auditory pathways. Visual language information, such as written text, is processed by the visual cortex, notably in the fusiform gyrus, while auditory language information, such as spoken words, is processed by the auditory cortex, particularly in the superior temporal gyrus. After initial processing, the information is sent for higher-level semantic analysis in an amodal network, meaning it is processed independently of the original sensory modality. This amodal processing primarily involves the left inferior frontal gyrus and middle temporal gyrus, which contribute to the integration and interpretation of language regardless of whether it was initially seen or heard [113].

It is worth mentioning that reading does not only involve orthographic, phonological and semantic processing. Reading also relies heavily on the integration of sensory, motor and memory systems [114, 115, 116]. Figure 3.9 provides a basic overview on how the word processing parts of the brain interact with other

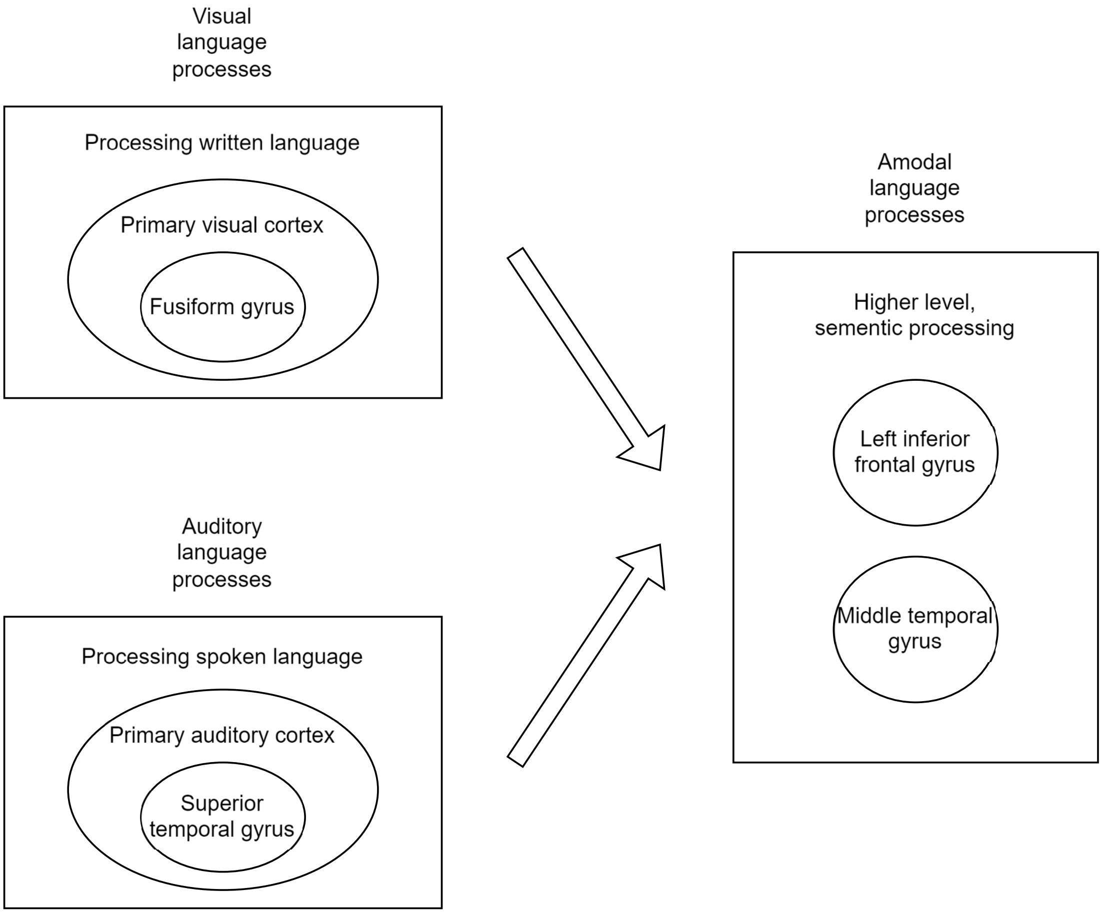


Figure 3.8: Simplified overview of how language information is processed in the brain

systems to enable reading.

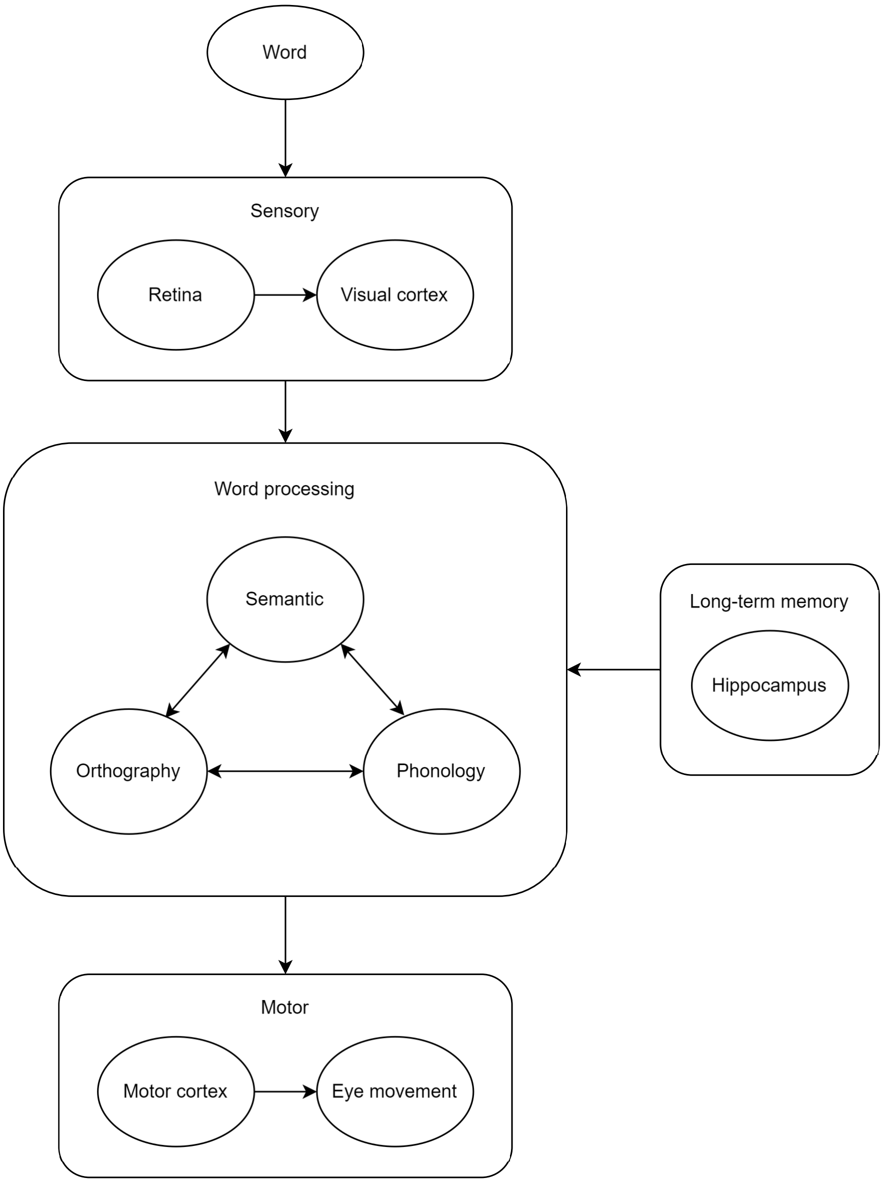


Figure 3.9: Overview of how orthographic, phonological and semantic processing integrate with the sensory, motor and memory systems

The sensory system, which is crucial for first picking up the visual information of the text, was discuss in chapter 2. The oculomotor system, which controls the eye movement to scan the text, was discussed in section 3.2. Working and long-term memory, where information about the words we learnt, are beyond the scope of this thesis. From this diagram, it can be seen that reading is not a linear process but a dynamic interaction of distributed networks, integrating visual, phonological, semantic, and cognitive resources to achieve fluent and meaningful reading.

3.6 Neurobiology of dyslexia

Understanding the various learning difficulties that can impact learner's ability to read and write can lead to improvements in the early diagnosis of these learners. This allows for early intervention, so that they may receive the education which they are entitled to. Furthermore, learners who are diagnosed young are more likely to have a positive self-image [117].

According to the British Dyslexia Association, dyslexia is defined as a specific learning difficulty which primarily affects reading and writing skills [118]. There are many theories formulated to explain the mechanism of dyslexia, including phonological theory, auditory processing deficit theory, visual processing deficit theory, attentional deficit theory, cerebellar theory [119].

In an MRI study on dyslexic patients, Paz-Alonso et. al.(2018) mapped out the language processing neural pathway of both dyslexic and control participants, and highlighted brain region hypoactivation in dyslexic participants during reading tasks [120]. 3.10 summarizes the findings of Paz-Alonso:

In the control group, language information goes from the orthographic processing, which is the conventional spelling system of a language, to the phonological processing. It then goes to sementic processing, where higher level language processing is done. In dyslexic participants, hypoactivation was found in the brain regions shown for orthographic, phonological and semantic processing. To cope with reading tasks, dyslexic people develop compensatory strategies by developing stronger functional connectivity between the thalamus and the inferior parietal cortex and ventral occipitotemporal cortex during pseudoword reading, and between the hyppocampus and the pars opercularis during word reading. These compensatory strategies may help dyslexic readers to engage more resources in their reading tasks.

Ali et. al. (2021) conducted a review on the visual and auditory percep-

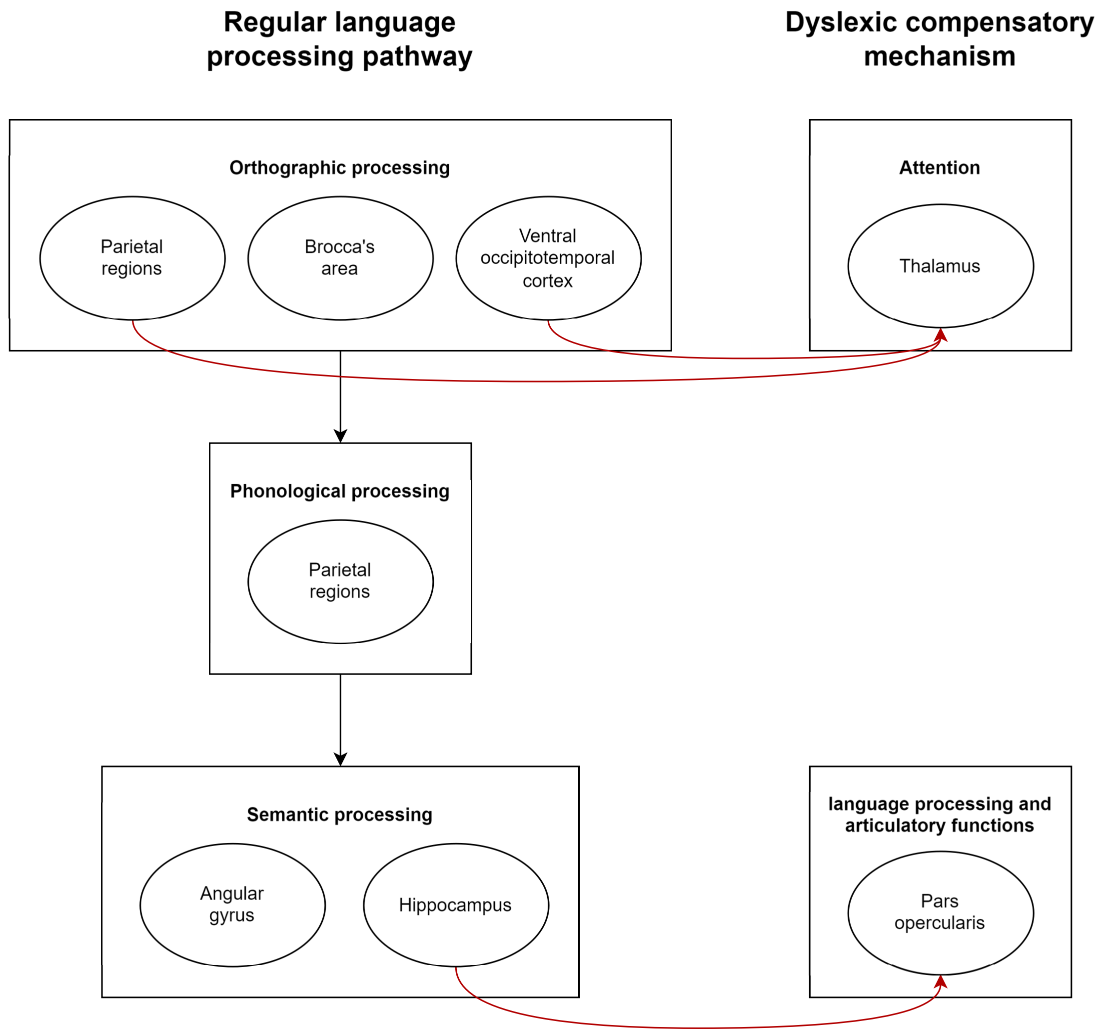


Figure 3.10: Language processing pathway and dyslexia compensatory mechanism

tion in dyslexia [121]. They found that dyslexia is closely linked to difficulties in processing phonological information within the temporal lobe, which significantly affects one's ability to read. Reading involves not only visual processing of text but also converting graphemes (letters) into phonemes (sounds). People with dyslexia often exhibit hypoactivation in the superior temporal gyrus, a key area for integrating letters and speech sounds. This deficit impairs phonological decoding, disrupting the audio-visual integration necessary for fluent reading. As a result, dyslexics may struggle to associate sounds with written words, even when their visual system functions normally, highlighting how impaired phonological processing in the temporal lobe underpins reading challenges traditionally associated with vision. This theory explains the findings in [122] and [120], and demonstrates the similarity between dyslexia and Irlen Syndrome.

In readers with dyslexia, the activation patterns in these pathways are disrupted, leading to reading difficulties. Kearns et. al. (2019) pointed out several studies showing that individuals with dyslexia exhibit reduced activity in the temporo-parietal and occipito-temporal regions during reading tasks, which correlates with difficulties in decoding and recognizing words by sight [123, 124, 9]. Similar to what Paz-Alonso et. al. (2018) found, compensatory mechanisms may emerge, such as increased reliance on the left precentral gyrus, which is involved in articulation, or subcortical regions like the thalamus and striatum, potentially reflecting alternative strategies to process written words [125, 120]. Dyslexic readers may also activate subcortical regions, such as the striatum or thalamus, more than their typical peers [126]. Due to the diverse function of each of these regions, its difficult to ascertain their role in dyslexia. These atypical patterns highlight the unique neurobiological adaptations dyslexic readers develop in response to challenges in using the typical reading pathways.

One concept which might be able to help people with dyslexia is neuroplasticity in reading acquisition. Neuroplasticity is the brain's ability to reorganize itself by forming new neural connections. Building on the concept of neuroplasticity, it follows that targeted training programs for individuals with dyslexia might be developed. Meyler et. al. (2008) have found significant changes in brain activation following remediation, such training can lead to improved reading abilities and more efficient neural processing [127]. By continuously adapting the training to the individual's progress and needs, educators can leverage neuroplasticity to foster lasting improvements in reading skills and overall academic performance.

3.7 Dyslexia and Irlen syndrome

The relationship between dyslexia and Meares–Irlen Syndrome (MIS) has been the subject of ongoing debate, particularly regarding the extent to which visual stress contributes to reading difficulties in certain individuals with dyslexia. Although dyslexia is primarily characterised as a phonological processing disorder, a subset of individuals report symptoms—such as visual distortions, glare sensitivity, and rapid fatigue when reading—that align more closely with MIS. Understanding the degree of overlap between these conditions is therefore crucial for determining whether visual stress acts as a separate but compounding factor within dyslexic populations, or whether it represents a distinct condition that requires its own screening and intervention strategies.

To investigate the potential association between the two conditions, Kriss and Evans (2005) employed the Wilkins Rate of Reading Test (WRRT), a standardised measure designed to quantify visual stress–related reading impairments [128]. The test was administered to both dyslexic participants and non-dyslexic controls under two conditions: reading with no overlay and reading with each participant's preferred coloured overlay. Their findings demonstrated that although both dyslexic and control participants showed statistically significant increases in reading speed when using overlays, the effect size was notably larger in the dyslexic group. This differential improvement was interpreted as evidence that symptoms associated with visual stress—such as perceptual distortions, glare, or instability of print—may be more prevalent or more severe among individuals with dyslexia. Kriss and Evans therefore suggested that Irlen Syndrome might occur with slightly higher frequency within dyslexic populations than in the general public, while also emphasising that MIS and dyslexia should not be considered synonymous. Instead, their results support the view that visual stress can act as a comorbid factor that exacerbates reading difficulties in some, but not all, individuals with dyslexia.

Building on the work of Kriss and Evans, a more recent study by Alanazi et. al. (2016) investigated the prevalence and symptom profile of MIS among female medical students in Saudi Arabia [47]. Using a structured questionnaire designed to quantify visual stress indicators—such as headaches, photophobia, text distortion, or difficulty sustaining reading—the researchers classified participants into four groups: those meeting criteria for MIS, those self-reporting dyslexia, those presenting both conditions, and those with neither condition. Statistical analyses were then used to compare the severity and frequency of visual stress symptoms across these groups, allowing the authors to identify patterns of discomfort and to explore potential interactions between MIS and dyslexia in a non-clinical, academically demanding population. The study found that students with MIS exhibited markedly higher levels of visual stress relative to those without the condition, supporting the validity of visual stress as a distinct construct. Moreover, individuals who reported both dyslexia and MIS tended to show the highest levels of symptom severity, suggesting an additive or compounding effect when the two conditions co-occur.

Both of these studies showed that dyslexia and MIS can be considered to be distinct but overlapping conditions, where both can be improved with the use of coloured overlays. However, there are other studies that have demonstrated that the evidence for classifying MIS as a separate disorder is weak, and visual stress should be better understood as a symptom of other conditions, such as dyslexia, rather than as a distinct diagnostic entity. Griffiths et. al. (2016) reviewed the research literature on the effect of coloured lenses or overlays on reading performance [129]. They concluded that there was no significant improvement in reading difficulties when coloured lens were used, compared to when they were not. cannot be endorsed and that any benefits reported by individuals in clinical settings are likely to be the result of placebo, practice or Hawthorne effects. Similarly, Miyasaki et. al. (2019), who reviewed decades of available studies on MIS, also concluded that the existing literature showed high heterogeneity, with many low-quality studies and insufficient high-level evidence on the existence of MIS and treatment effectiveness [130]. They stated that it remains uncertain whether MIS is a separate entity or part of the dyslexia spectrum.

All in all, whether or not MIS constitutes a distinct diagnostic entity, and regardless of ongoing debate over the efficacy of coloured lenses, the research consistently shows that a subset of individuals do experience visual stress that can hinder their reading and visual comfort. Acknowledging this phenomenon does not require full consensus on its underlying mechanisms, but it does underscore the importance of recognising and addressing these symptoms when they arise. At the same time, variability both within and between study samples highlights the need for more rigorous, standardised approaches to identifying visual stress and delineating its relationship with MIS, dyslexia and other reading-related difficulties.

3.8 Conclusion

In this chapter, the literature surrounding the visual processing of visual text was explored. The movement of the eye as a reader scanned through the text was discussed, which leads to a discussion on how the printed words are recognised by the brain. Research on the neurobiology of the brain during reading in both typical and dyslexic readers were then examined. This chapter demonstrated the complexity of the network of neurons and brain parts which work tirelessly together to enable us to read the text in this thesis.

The literature review conducted in this chapter have demonstrated the complexities between low-level analysis and higher-order processing, which, in the case of reading, would be linguistic interpretation. This will be informative in investigating how visual processing differs between digital stimuli and non-digital stimuli.

Chapter 4

Mathematics of Signal Processing

4.1 Introduction

The purpose of this research is to examine neural processes associated with processing the visual information of digital objects and non-digital objects. This can be studied by having subjects observe digital and non-digital objects while their EEG signals are being captured using an EEG system. The signals can then be analysed using EEG signal processing techniques which will be discussed in the next chapter. Techniques include frequency analysis to observe how the amplitude of the brain waves changes with frequency, statistical analysis to compare between two EEG signals, source reconstruction to localise the source of the signal and connectivity analysis to compare how different parts of the brain communicate one another.

In this chapter, we will discuss the mathematical foundations in signal processing, in particular time-frequency transform and spatial transform. Understanding the mathematics behind signal processing gives us a solid foundation on how they work, and allow us to fine tune specific variables in order to get the desired results

4.2 Time-frequency transform

Time-frequency analysis is a powerful framework for understanding signals that exhibit dynamic changes over time, making it particularly valuable in the study of complex, non-stationary signals like EEG. Methods such as the Short-Time Fourier Transform (STFT), Wavelet Transform, and the Hilbert-Huang Transform have emerged as key tools in uncovering these dynamic patterns, each offering unique strengths and limitations. The basis of these techniques is the Fourier transform, which is discussed in Appendix B. The next section will discuss some of the applications of Fourier transform, before discussing the more advance techniques listed, which are built on Fourier transform.

4.2.0.1 Applications of Fourier Transform

Fourier transform is used in applications such as the filtering of signals at high or low frequencies. These techniques are known as high-pass filtering and low-pass filtering respectively. When the two techniques are combined (both high and low frequencies are filtered out), the technique is known as band-pass filtering. Fourier transform works by converting a signal from the time-domain into the frequency-domain. In the frequency-domain, the signal can be easily manipulated to remove frequency components which are undesired. This is done by multiplying the Fourier transform $F(f)$ by a filter function $H(f)$. The equation for low-pass filter $H_{LP}(f)$ (equation 4.1), high-pass filter $H_{HP}(f)$ (equation 4.2) and band-pass filter $H_{BP}(f)$ (equation 4.3) are:

$$H_{LP}(f) = \begin{cases} 1, & \text{if } |f| \leq f_c \\ 0, & \text{if } |f| > f_c \end{cases} \tag{4.1}$$

$$H_{HP}(f) = \begin{cases} 0, & \text{if } |f| \leq f_c \\ 1, & \text{if } |f| > f_c \end{cases} \tag{4.2}$$

$$H_{BP}(f) = \begin{cases} 1, & \text{if } f_L \leq |f| \leq f_H \\ 0, & \text{otherwise} \end{cases} \tag{4.3}$$

Where f_c is the predetermined cut-off frequency. The filtered signal in the frequency domain can then be calculated using the equation 4.4:

$$F_{filtered}(f) = F(f) \cdot H(f) \tag{4.4}$$

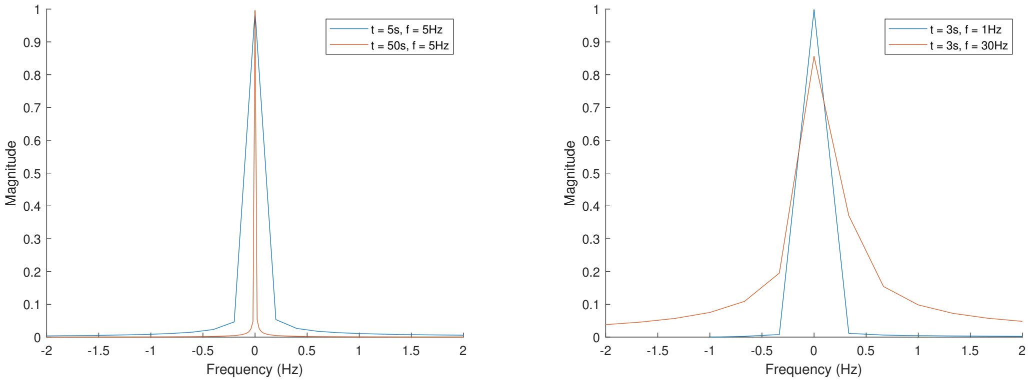
After the filter function has been applied, the resulting frequency-domain signal can then be converted back to time-domain using the inverse Fourier transform equation, equation 4.5:

$$H(t) = \int_{-\infty}^{+\infty} F(f) \mathrm{e}^{i2\pi ft} df \tag{4.5}$$

This produces a time-domain signal that contains only the frequencies within the desired band. Equations B.1 and 4.5 form a reciprocal pair, which is the basis for Fourier's reciprocal integral relation [131]. This relation is the basis of the Uncertainty principle discussed in the next section.

4.3 Uncertainty principle

One characteristic of the Fourier transform is that there is a trade-off between a signal's time resolution and frequency resolution. This characteristic is known as the uncertainty principle. This principle is mathematically derived from properties of the Fourier transform [131], as will be shown below.



(a) Change in time (b) Change in frequency

Figure 4.1: Change in standard deviation of Fourier transform when the (a) duration and (b) frequency is adjusted. Frequency has been normalised to zero.

The uncertainty principle is derived by analyzing the spread of the signal $f(t)$ in the time domain and the spread of its Fourier transform $F(\omega)$ in the frequency domain. Summarizing the work of [132], the time density can be defined as $|s(t)|$, which leads to the following definition for the mean time:

$$\langle t \rangle = \int t \, |s(t)|^2 \, dt \tag{4.6}$$

Hence, the standard deviation of the signal in time Δt is defined as:

$$\Delta t = \sqrt{\int (t - \langle t \rangle)^2 \, |f(t)|^2 \, dt} \tag{4.7}$$

Similarly, we can derive the following definition for the mean frequency from the frequency density $|s(\omega)|$:

$$\langle \omega \rangle = \int t \, |s(\omega)|^2 \, dt \tag{4.8}$$

Hence, the standard deviation in the frequency domain ω_0 is:

$$\Delta \omega = \sqrt{\int (\omega - \langle \omega \rangle)^2 \, |F(\omega)|^2 \, d\omega} \tag{4.9}$$

These measures quantify the "spread" of the signal in time and frequency. Using Cauchy–Schwarz inequality [133], it can be shown that the product of these spreads satisfies the inequality:

$$\Delta t \cdot \Delta \omega \geq \frac{1}{2} \tag{4.10}$$

This inequality expresses the uncertainty principle, illustrating that the narrower a signal is in the time domain (smaller Δt), the broader it must be in the frequency domain (larger $\Delta \omega$), and vice versa.

4.3.1 Short-term fourier transform

While the Fourier Transform allows us to analyzes signals in the frequency domain by breaking them into their constituent frequencies, it loses all time-related information, meaning it cannot indicate when specific frequency components occur in the signal. This limitation is particularly problematic for non-stationary signals, like EEG, where frequency content changes over time. Hence, we need methods which can transform a time-dependent signal into a two-dimensional time-frequency product space. Mathematically, this can be represented using equation 4.11:

$$H(t) \rightarrow F(t, f) \tag{4.11}$$

where $H(t)$ is the time-dependent signal and $F(t, f)$ is the time-frequency transform. One of the methods which can achieve this is the Short-Time Fourier Transform (STFT). The STFT expands on the Fourier transform by dividing the signal into small time segments using a sliding window and performing the Fourier transform on each segment. This allows the STFT to capture both time and frequency information, providing a time-frequency representation that shows how the spectral content of the signal evolves over time.

Here, a non-stationary surrogate signal will be generated to demonstrate how STFT works. Suppose we have the following surrogate signal $F(t)$, defined by equation 4.12:

$$F(t) = \begin{cases} sin(8\pi t) + sin(20\pi t), & \text{if } 0 \leq |t| \leq 3 \\ sin(12\pi t) + sin(24\pi t), & \text{if } 3 < |t| \leq 6 \\ sin(16\pi t) + sin(28\pi t), & \text{if } 6 < |t| \leq 9 \end{cases} \tag{4.12}$$

On a Cartisian plane, $F(t)$ can be plotted as shown in figure 4.2:

We can see that in $F(t)$, there exists different frequencies at different time points. Performing a Fourier transform, we get the following magnitude plot as shown in figure 4.3:

As can be seen from Figure 4.3, by applying a Fourier transform to our signal, we only see that the signal contains the frequencies 4 Hz, 6 Hz, 8 Hz, 10 Hz, 12 Hz and 14 Hz. However, we do not know in which time point do these frequencies

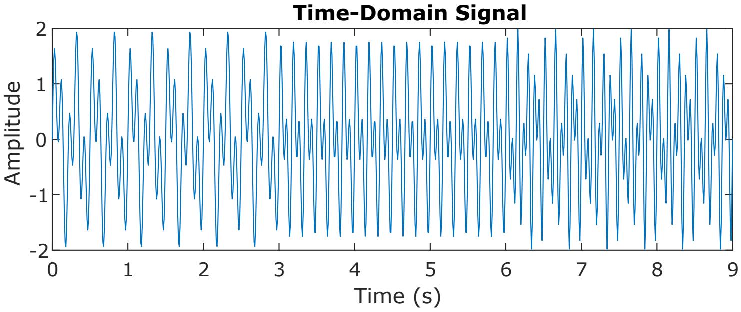


Figure 4.2: Raw surrogate signal

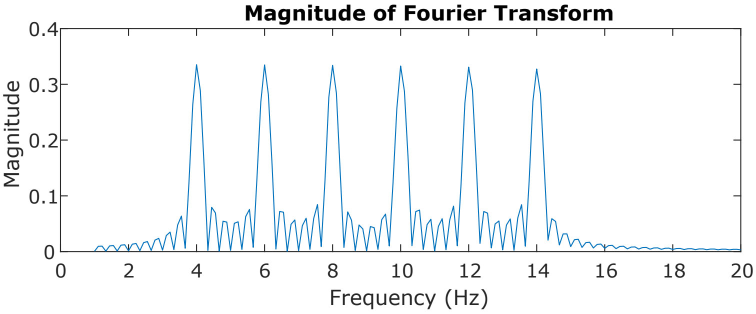


Figure 4.3: Magnitude plot of FFT of signal

exist. Hence, if we want to be able to examine the different frequencies that exist at different time point, we need a different method of analysis.

One way to resolve the issues with Fourier transform is through the use of Short-Time Fourier transform (STFT). STFT first divides the signal into smaller, overlapping time segments, and then performs the Fourier Transform on each segment individually. Mathematically, this is done using equation 4.13:

$$S^{(F)}(f,t) = \mathbf{F}_{\tau}\left\{H(\tau)\cdot w^{(F)}(t-\tau)\right\} \tag{4.13}$$

Given a signal $H(\tau)$, the signal is first multiplied by the window function $w^{(F)}(t-\tau)$, which is a time window centered around t. The result is then subjected to Fourier transform $\mathbf{F}$. The result is a complex-valued function of time t and frequency f [10].

By doing this, the STFT provides a time-frequency representation of the signal, where each segment's Fourier Transform reveals the frequency content within each short time segment. This allows the STFT to capture how the frequency content of the signal evolves over time. The results of an STFT transform can be represented on a spectrogram, which is a visual representation of the frequency content of a signal as it changes over time [134]. It displays three key dimensions of information:

1. Time (typically on the horizontal axis)
2. Frequency (typically on the vertical axis)
3. Amplitude or intensity (represented by color or brightness)

Performing a STFT on our signal using a segment length of 0.5s and no overlaps, we get the spectrogram shown in figure 4.4.

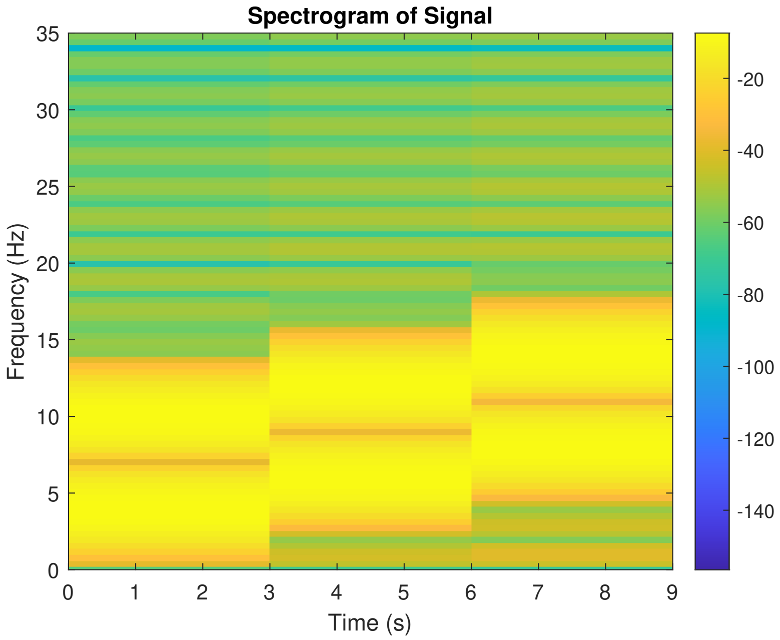


Figure 4.4: Spectrogram of STFT of signal

As can be seen from figure 4.4, the frequency is plotted against time, and the magnitude is represented using a colour scale, as shown on the right of the spectrogram. This spectrogram shows how the signal's frequencies vary over time, enabling the analysis of non-stationary signals where frequency components change over the course of the signal. There are three segments shown, corresponding to the frequencies in each of the three segments in the surrogate signal. We see that the signal contained frequencies at 4Hz and 10Hz from 0s to 3s, 6Hz and 12Hz from 3s to 6s and 8Hz and 14Hz from 6s to 9s.

While the STFT is useful in extracting information on how the frequencies in a signal change over time, it is still limited due to its fixed window size. This means that it has a constant resolution for both time and frequency throughout the signal. This results in a trade-off: a narrow window provides good time resolution but poor frequency resolution, while a wide window does the opposite.

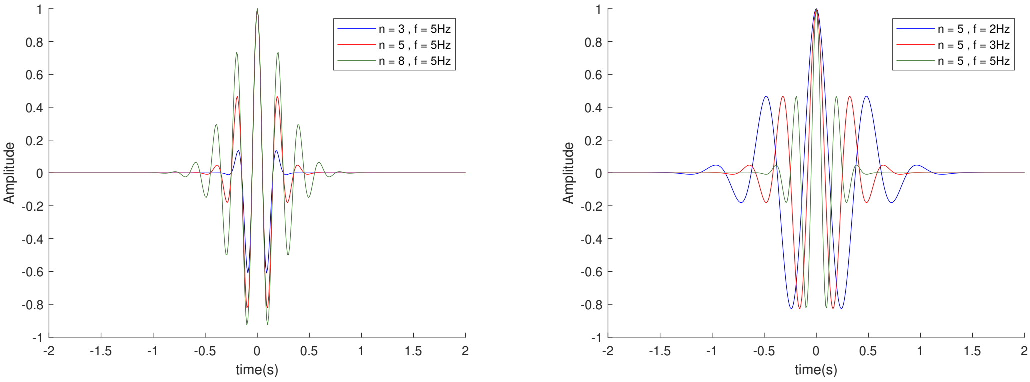
4.3.2 Wavelet transform

The wavelet transform breaks down signal components across both time and frequency domains using scalable wavelets, localized wave-like functions that can stretch and compress, to capture both the frequency content and the time at which specific frequencies occur. This allows the resolution to be adjusted on the basis of the frequency content of the signal. It uses shorter windows for higher frequencies (providing better time resolution) and longer windows for lower frequencies (providing better frequency resolution), making it ideal for analyzing non-stationary signals.

Wavelets are constructed by scaling and translating a "mother wavelet", a base function that serves as the prototype for all other wavelets. To qualify as a wavelet, the function $\psi(t)$ must satisfy two main conditions:

1. It must have a zero mean ($\int_{-\infty}^{\infty} \psi(t)dt = 0$)
2. It must be square-integrable (finite total energy) ($\int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty$)

One of the basic wavelets that has been used to process EEG signals is the Morlet wavelet. Morlet wavelet, as defined by [135], is a sine wave tapered by a Gaussian [136, 135]. While many improvements on the Morlet wavelet have been proposed [136], and there are many different types of mother wavelets used for signal processing [137], this chapter will focus on the original Morlet wavelet proposed by [138].



(a) Change in number of cycles

(b) Change in frequency

Figure 4.5: Morlet wavelets in the time domain being transformed when the (a) number of cycles and (b) frequency is varied.

Mathematically, the Morlet wavelet can be represented using equation 4.14:

$$w = e^{2i\pi ft} e^{\frac{-t^2}{2\sigma^2}} \tag{4.14}$$

with a range of t centered at $t = 0$ to avoid phase shift. σ is the width of the Gaussian, which is defined mathematically using equation 4.15

$$\sigma = \frac{n}{2\pi f} \tag{4.15}$$

where n defines the time-frequency precision trade-off, and is often referred to as the “number of cycles.” The wavelet can be scaled in the x-axis by modifying f, and scaled in the y-axis by modifying n. This is demonstrated in Figure 4.6:

As with STFT, the aim of wavelets is to generate a time-frequency spectrogram from the raw signal. Once the mother wavelet has been chosen, a signal can be converted from time-domain to its time-frequency representation by means of convolution between the signal and the wavelet. According to [139], a convolution is an integral that expresses the amount of overlap of one function as it is shifted over another function. In other words, two functions are placed one over the other and the area under the graph, which is covered by both functions, is calculated. Mathematically, this can be represented using the equation 4.16.

$$f * g = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \tag{4.16}$$

where $f * g$ denotes the convolution between f and g. An animation on the calculated area can be found in [139]. This is particularly useful in the signal processing of EEG signals as the signals can be split into segments, which are then compared to a wavelet of fixed frequency. The amplitude of each frequency at a given time frame in the signal can then be determined.

Mathematically, the convolution between the signal $F(t)$ and the mother wavelet $\Psi(t)$ can be expressed using equation 4.17:

$$H(t) = F(t) * \Psi(t) \equiv \int_{-\infty}^{+\infty} F(\tau)\Psi(t-\tau, f_0)d\tau \tag{4.17}$$

Scaling the wavelet by a factor of a, we get the equation 4.18:

$$H(a,t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} F(\tau)\Psi(\frac{t-\tau}{a}, f_0)d\tau \tag{4.18}$$

Performing a wavelet transform using the Morlet wavelet on our signal, we get the scalogram shown in figure 4.4.

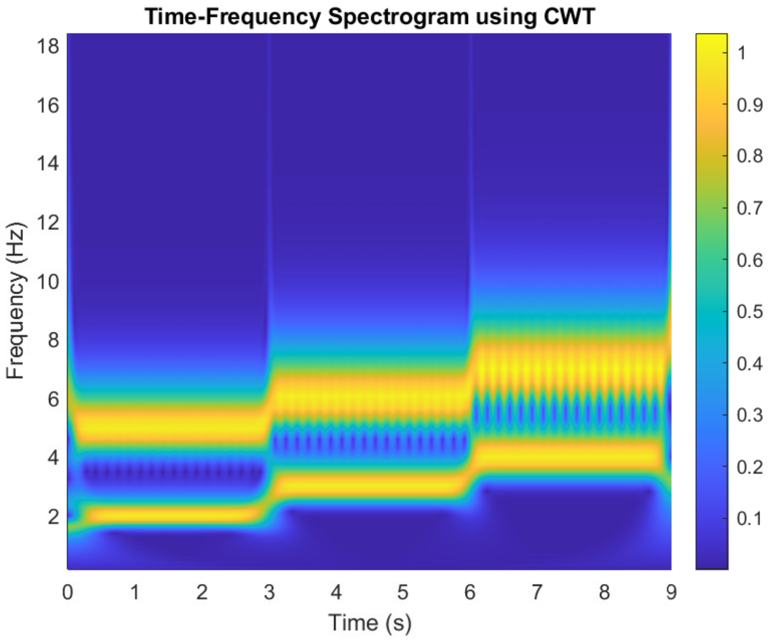


Figure 4.6: Scalogram of wavelet transform of signal with Morlet wavelet

Looking back at figure 4.6, we observe 3 segments similar to what we observed in the STFT spectrogram (figure 4.4). However, we also observed that there is a trade-off between time and frequency resolution. At higher frequencies, the band are wider due to the higher range of possible frequencies and lower frequency resolution, and the band shift as time progresses are sharper due to the higher time resolution, demonstrating the uncertainty principle in wavelet transform. The ability for wavelet transform to allow for adaptive resolution—using narrow windows for high-frequency components (which often vary quickly and need good time resolution) and wide windows for low-frequency components (which are more stable and need better frequency resolution) improves on this trade-off compared to the Short-Time Fourier Transform (STFT).

In particular, the bandwidths of neural rhythms are organized in a logarithmic progression to account for the uncertainty principle. As [140] has discussed, the uncertainty principle, when applied to the log scaling of bandwidths in physiological signals such as EEG, highlights the fundamental trade-off between time and frequency resolution. Logarithmic scaling is crucial because neural rhythms span a wide range of frequencies, from slow delta waves to fast gamma oscillations, each associated with different cognitive and physiological states. As will be discussed in the next chapter, this allows for the dynamic adjustment of the time-frequency resolution, so that finer frequency details are captured during low brain activity, and changes in frequency can be captured more readily during high brain activity. Logarithmic approaches, such as wavelet transforms, make this trade-off explicit and efficient for analyzing the complex, non-stationary properties of EEG signals.

4.3.3 Hilbert transform

Another technique which is used in the processing of EEG data is called Hilbert transform. As explained by [10], Hilbert transform generates an analytic signal from the raw signal. This is achieved by applying a frequency-dependent phase shift of 90 degrees to the signal's Fourier components and suppressing negative frequencies. The resulting spectral signal is then converted back to the time-domain using inverse transform, generating a complex analytic signal, where the real part is the original signal, and the imaginary part is its Hilbert transform. This allows for the calculation of instantaneous amplitude and phase, which are useful for extracting time-varying features from band-limited signals.

While the Hilbert transform is useful for generating complex analytic signal from real signal, when applied to broadband and multi-component signal such as EEG, the instantaneous frequency derived from the Hilbert transform becomes ambiguous and difficult to interpret. The Hilbert-Huang Transform (HHT) builds on the Hilbert transform by combining it with Empirical Mode Decomposition (EMD). EMD adaptively decomposes a signal into a set of intrinsic mode functions (IMFs), each representing an oscillatory mode whose frequency and amplitude may vary over time. Once the signal is decomposed, the Hilbert transform is applied to these IMFs to extract instantaneous frequencies and amplitudes. The result is a Hilbert spectrum, a time-frequency-energy plot that reveals how signal components vary over time. The Hilbert spectrum obtained from the surrogate signal is shown in Figure 4.7:

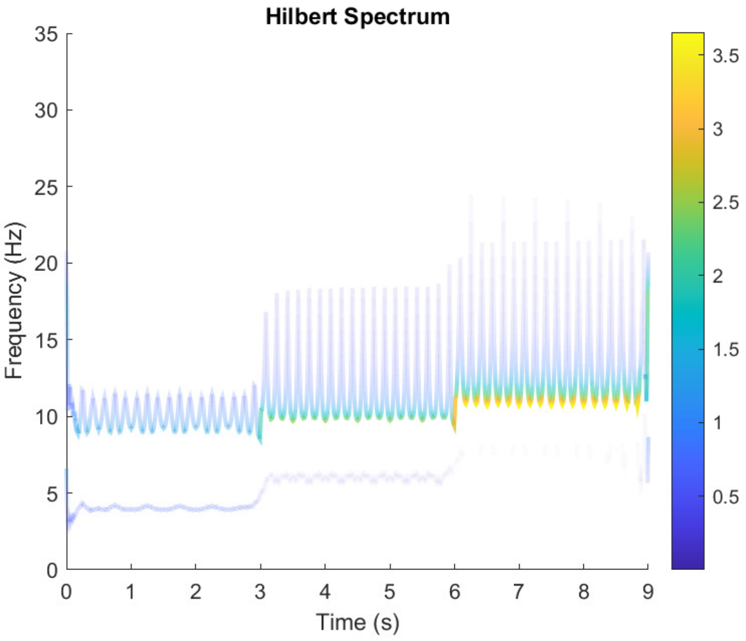


Figure 4.7: Spectrogram of HH of signal

Similar to the spectrogram from the STFT and the scalogram from the wavelet transform, multiple temporal segments and dominant oscillatory components can be observed. However, the instantaneous frequency estimates exhibit oscillatory fluctuations around the expected frequencies, reflecting the sensitivity of EMD to signal discretization, noise, and end effects. These fluctuations tend to increase at higher frequencies due to limitations in phase differentiation and sampling resolution.

4.3.4 Comparing the methods of transform

While each of the methods of transform presented above has its distinct advantages, all of them ultimately resulted in similar transform signals. As highlighted by [10], all three methods give time-frequency spectral results which are essentially the same, as long as the relevant analysis parameters are matched with each other. To demonstrate this point, [10] presented a graphical comparison of the Fourier, Hilbert and wavelet convolution kernels, with center frequencies and bandwidths being matched with each other, as shown in Figure 4.8.

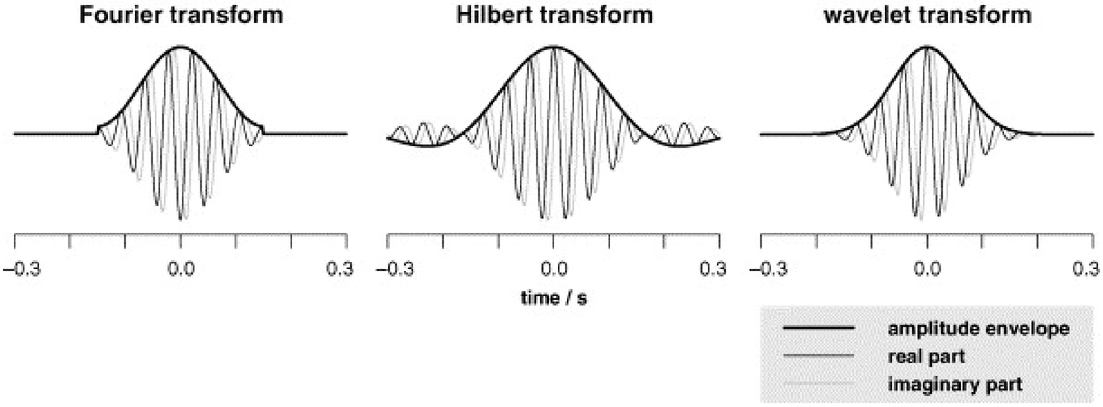


Figure 4.8: Figure to demonstrate the mathematical equivalence of Fourier, Hilbert and Wavelet transforms [10].

In Figure 4.8, Bruns (2004) demonstrated that despite choosing different filter forms for each transform, results are still very similar [10]. Had all three transforms used Gaussian filters, the results from these three approaches could have been even more perfectly identical.

4.3.5 Applying methods of transform to EEG signal

While the above analysis did corroborate with Bruns's results, it is important to consider the characteristics of the signal to be analyzed in choosing a method of transform. Figure 4.11 a shows a sample of the target of analysis: EEG signal:

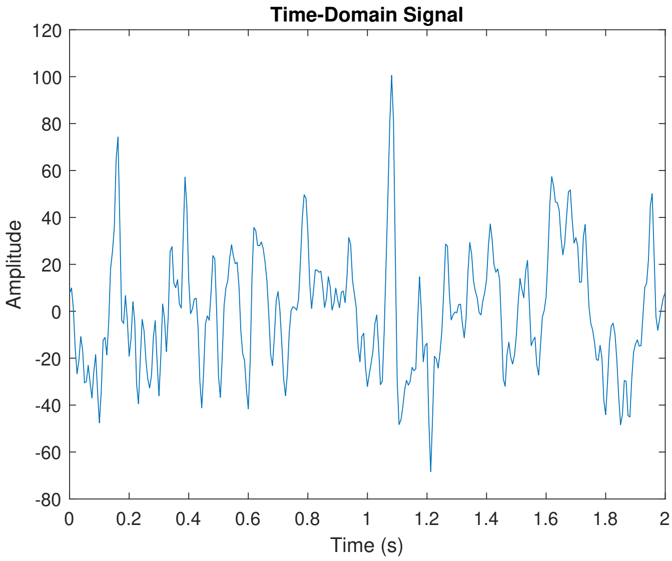
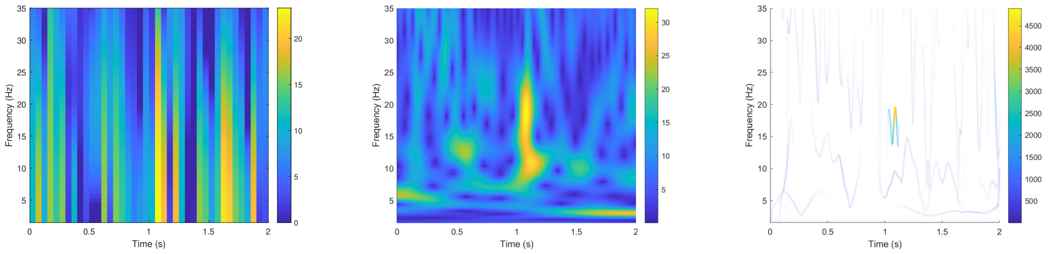


Figure 4.9: Sample of Raw EEG signal

As can be seen from Figure 4.11, EEG signals are complex, dynamic, and non-stationary. They exhibit fluctuations in amplitude and frequency over time. The non-stationary nature of EEG signals means that their statistical properties, such as mean and variance, change over time rather than remaining constant. This is due to the brain's varying states, such as wakefulness, attention, and sleep, which modulate neural activity across different regions. This non-stationarity requires advanced analytical tools which can accurately capture the dynamic patterns of the signal, so that the underlying physiological or pathological processes may be interpreted.

Performing STFT, Morlet wavelet transform and Hilbert-Huang transform, we get the results shown in Figure 4.10:

As can be seen from Figure 4.10, the Morlet wavelet transform resulted in the time-frequency representation which captures the most accurate information. While the STFT is conceptually and computationally simple, it is more suitable



(a) Short-term Fourier transform

(b) Morlet wavelet transform

(c) Hilbert-Huang transform

Figure 4.10: STFT, Morlet wavelet and Hilbert-Huang transform of the EEG signal

for stationary signals, where the statistical properties of the signal do not change over time. Since STFT uses a fixed window throughout the signal, assuming stationary within each window, it cannot adapt to the dynamic nature of EEG signals, potentially missing important events. From Figure 4.10a, while the brain activation at time 1.1 s was captured, it is not clear which frequency band the activation was in. Furthermore, there are artifacts at time 1.6 s, while many other weaker activations were not captured.

The Hilbert–Huang transform, although adaptive and well suited to non-stationary data in principle, relies on Empirical Mode Decomposition (EMD). This process can suffer from mode mixing, where different frequency components are combined into a single IMF or split across multiple IMFs. Additionally, the HHT produces instantaneous frequency estimates that are only reliable if the IMFs are approximately narrowband and well-behaved, which is not always the case for EEG signals that exhibit irregular, broadband, and highly dynamic characteristics. As can be seen in Figure 4.10c, while the transform did capture the brain activation at time 1.1 s, and other weaker activations, their frequency characteristics are less clearly defined compared to the wavelet representation.

While the STFT and Hilbert-Huang transform are powerful tools for analyzing signals, the wavelet transform is better suited to provide a time-frequency representation for the non-stationary, broadband and dynamic nature of EEG data. Unlike the STFT and the Hilbert-Huang transform, the wavelet transform offers multi-scale analysis, enabling simultaneous decomposition of EEG signals into distinct rhythms. Its dynamic resolution makes it ideal for capturing transient events like task-related bursts or evoked potentials. Additionally, the wavelet transform's robustness to noise and flexibility in choosing mother wavelets tailored to EEG patterns further enhance its utility. Thus, the wavelet transform is the most versatile and effective for EEG signal processing, and will be used for this thesis.

4.3.6 Limitations of single-sample evaluation

While the above comparison demonstrates the differences in the time-frequency representations produced by the STFT, Morlet wavelet transform, and Hilbert-Huang transform, it is important to acknowledge that this evaluation is based on a single EEG recording. As such, the results should not be interpreted as a universal performance ranking applicable to all EEG signals. EEG recordings exhibit substantial variability depending on subject, task, recording conditions, and the presence of artifacts. Consequently, the observed behaviour of each transform must be interpreted in relation to the specific signal characteristics present in this example.

EEG signals are subject to a number of well-established signal characteristics that place distinct demands on time-frequency analysis methods. These include transient artifacts such as eye blinks and muscle activity [141], broadband noise resulting in low signal-to-noise ratios [142], non-stationarity arising from time-varying neural dynamics [143], periodic interference such as mains line noise [144], and low-frequency baseline drift [145]. Each of these characteristics violates, to varying degrees, the assumptions of stationarity or narrowband behaviour that underpin many classical signal processing techniques.

Transient events require high temporal resolution to be accurately localised, while non-stationary oscillations necessitate adaptive frequency resolution. Periodic interference can dominate spectral representations if not appropriately handled, and broadband noise can obscure weak but physiologically meaningful neural activity. The suitability of a given transform therefore depends not only on computational considerations, but also on how its inherent assumptions align with the dominant characteristics of the signal under analysis.

To complement the real EEG example and to isolate the influence of individual signal characteristics, a set of synthetic EEG-like signals was generated. Each synthetic signal was designed to emphasise a single common EEG challenge, including transient activity, broadband noise, non-stationarity, periodic interference, and baseline drift, while maintaining a consistent underlying oscillatory structure. This approach enables controlled illustration of how different transforms respond to specific signal features.

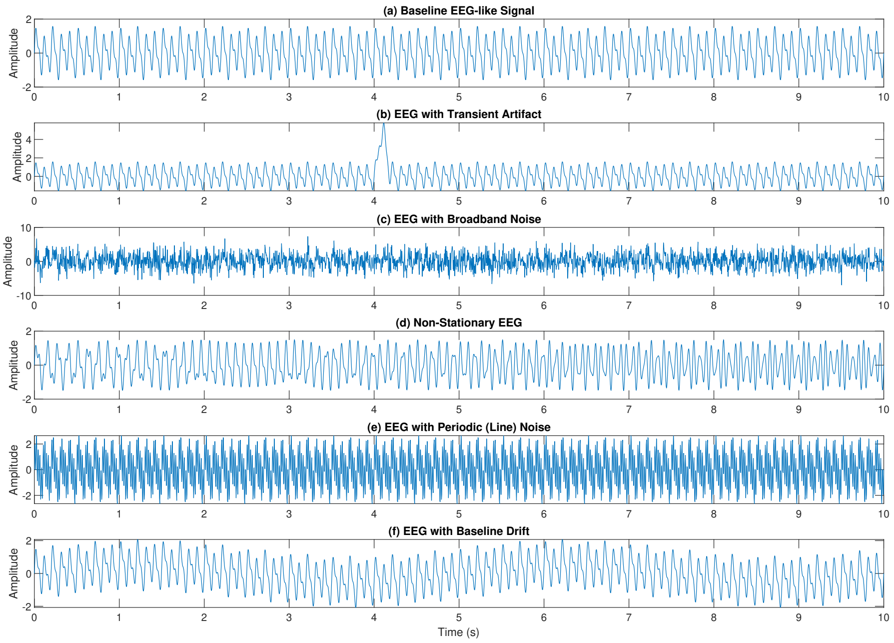
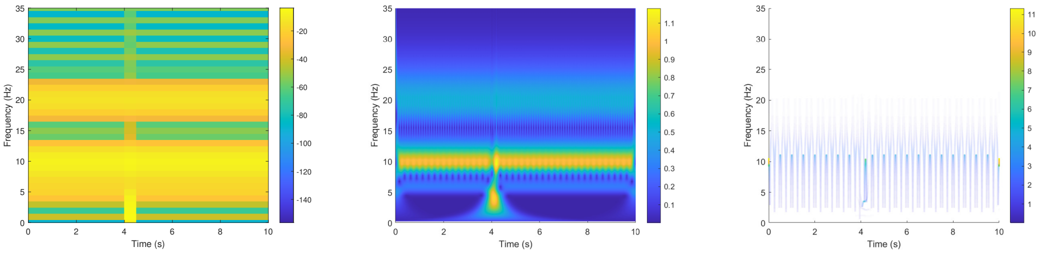


Figure 4.11: Synthetic EEG-like signals emphasising one common EEG challenge each

The synthetic signals are not intended to replicate the full physiological complexity of real EEG data, but rather to provide targeted stress tests that highlight the interaction between signal properties and analytical assumptions.

The baseline synthetic EEG signal (Figure 4.11a) was constructed as a lin-



(a) Short-term Fourier transform

(b) Morlet wavelet transform

(c) Hilbert-Huang transform

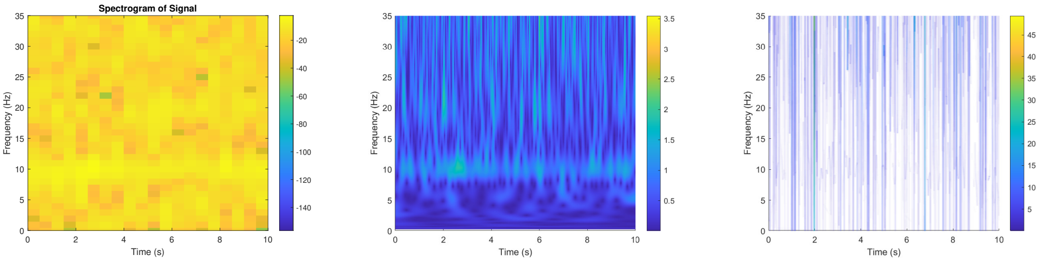
Figure 4.12: STFT, Morlet wavelet and Hilbert-Huang transform of the synthetic EEG signal with transient artifact

ear superposition of three sinusoidal components representing canonical EEG rhythms: a dominant alpha oscillation at 10 Hz, a beta component at 20 Hz with half the amplitude of the alpha rhythm, and a lower-amplitude theta component at 6 Hz. This combination produces a stationary, noise-free signal with a stable oscillatory structure and well-defined spectral content, serving as a reference signal against which the effects of specific signal distortions can be assessed.

In Figure 4.11b, a transient artifact was localized to a narrow temporal window and shaped using a Hann window to produce a smooth onset and offset, minimizing spectral leakage. STFT, Morlet wavelet and Hilbert-Huang transform was applied to this signal in Figure 4.12:

For the signal containing a transient artifact, the STFT (Figure 4.12a) detects a short-duration broadband energy increase, but the fixed analysis window results in temporal smearing and reduced localization of the artifact onset. The Morlet wavelet transform (Figure 4.12b) provides the clearest representation of the transient, capturing its abrupt temporal structure across a wide range of scales while preserving the surrounding oscillatory activity. In contrast, the Hilbert–Huang transform (Figure 4.12c) decomposes the transient into multiple intrinsic mode functions, which, while temporally precise, can complicate interpretation due to mode mixing and reduced consistency across realizations.

In Figure 4.11c, broadband noise was introduced by adding zero-mean Gaussian white noise to the baseline EEG signal, resulting in a low signal-to-noise



(a) Short-term Fourier transform (b) Morlet wavelet transform (c) Hilbert-Huang transform

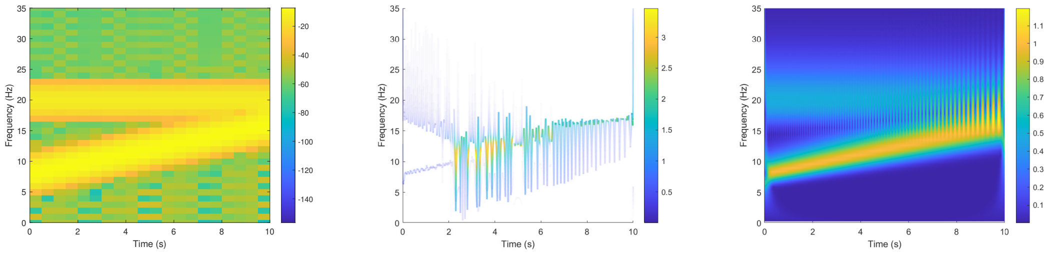
Figure 4.13: STFT, Morlet wavelet and Hilbert-Huang transform of the synthetic EEG signal with broadband noise

ratio condition. STFT, Morlet wavelet and Hilbert-Huang transform was applied to this signal in Figure 4.13:

In the broadband noise condition, the STFT (Figure 4.13a) produces a diffuse time–frequency representation with poor contrast between noise and oscillatory components. The Morlet wavelet transform (Figure 4.13b) offers superior interpretability by concentrating rhythmic activity into well-defined time–scale regions, allowing dominant EEG rhythms to remain identifiable despite elevated noise levels. Although the Hilbert–Huang transform Figure 4.13c adaptively separates components, its instantaneous frequency estimates are sensitive to noise, leading to fragmented and less stable representations compared to the smoother and more robust wavelet-based results.

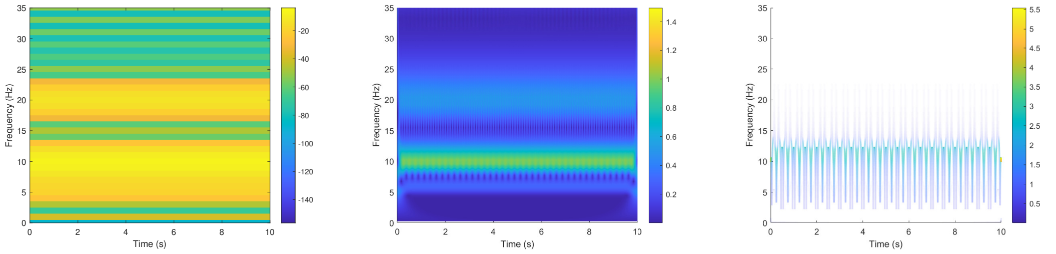
In Figure 4.11d, non-stationarity was introduced by replacing the stationary alpha component with a sinusoid whose instantaneous frequency increased linearly over time from 8 Hz to 12 Hz across the recording duration. This frequency-modulated oscillation was combined with a stationary beta rhythm, producing a signal in which spectral characteristics evolve continuously over time. STFT, Morlet wavelet and Hilbert-Huang transform was applied to this signal in Figure 4.14:

For the non-stationary signal with frequency drift, the STFT (Figure 4.14a) captures the overall trend but lacks sufficient resolution to track gradual frequency changes accurately. The Morlet wavelet transform (Figure 4.14b) performs best



(a) Short-term Fourier transform (b) Morlet wavelet transform (c) Hilbert-Huang transform

Figure 4.14: STFT, Morlet wavelet and Hilbert-Huang transform of the synthetic EEG signal with non-stationarity



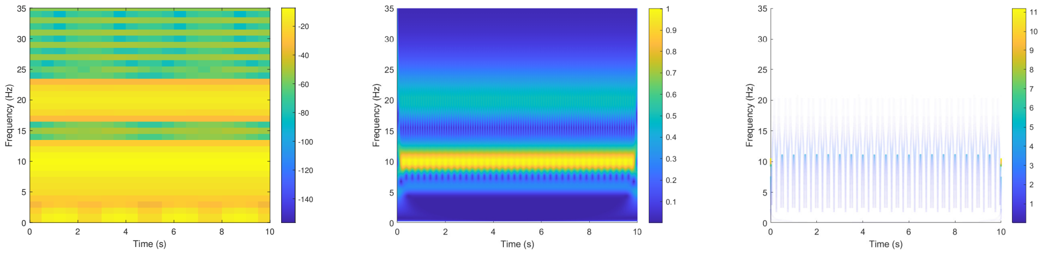
(a) Short-term Fourier transform (b) Morlet wavelet transform (c) Hilbert-Huang transform

Figure 4.15: STFT, Morlet wavelet and Hilbert-Huang transform of the synthetic EEG signal with periodic interference

in this scenario, clearly revealing a smooth, continuous ridge corresponding to the evolving dominant frequency over time. While the Hilbert–Huang transform (Figure 4.14c) provides high temporal resolution, its decomposition can introduce variability in frequency estimates, making the wavelet transform the most reliable and interpretable method for characterizing gradual non-stationary dynamics.

In Figure 4.11e, a sinusoidal component at 50 Hz was added to the baseline EEG signal, representing power-line contamination commonly encountered in EEG recordings. STFT, Morlet wavelet and Hilbert-Huang transform was applied to this signal in Figure 4.15:

In the presence of periodic line noise, the STFT (Figure 4.15a) identifies the interference as a persistent narrowband component, though its representation is influenced by windowing effects. The Morlet wavelet transform (Figure 4.15b) distinctly isolates the periodic interference across time with minimal distortion, while simultaneously maintaining separation from lower-frequency EEG rhythms.



(a) Short-term Fourier transform (b) Morlet wavelet transform (c) Hilbert-Huang transform

Figure 4.16: STFT, Morlet wavelet and Hilbert-Huang transform of the synthetic EEG signal with baseline drift

Although the Hilbert–Huang transform (Figure 4.15c) separates the interference into a distinct intrinsic mode function, its utility is limited by sensitivity to decomposition parameters, whereas the wavelet transform provides a more stable and transparent representation.

In Figure 4.11f, baseline drift was simulated by adding a low-frequency sinusoidal component at 0.2 Hz to the baseline EEG signal. STFT, Morlet wavelet and Hilbert-Huang transform was applied to this signal in Figure 4.16:

For the signal containing baseline drift, the STFT (Figure 4.16a) shows increased low-frequency power but offers limited insight into the temporal evolution of the drift. The Morlet wavelet transform (Figure 4.16b) most effectively captures the slow-varying baseline component as a large-scale feature while preserving higher-frequency oscillations, enabling clear multi-scale interpretation. The Hilbert–Huang transform (Figure 4.16c) isolates the drift into a low-frequency mode; however, the wavelet transform yields a more consistent and readily interpretable separation of slow trends from neural activity.

Taken together, both the real EEG example and the synthetic signal illustrations indicate that, for non-stationary, broadband EEG signals containing transient activity, the wavelet transform provides the most informative and interpretable time-frequency representation among the methods considered. Accordingly, the wavelet transform is adopted as the primary analytical tool for subsequent analyses in this thesis.

4.3.7 Section summary

In summary, this section discussed the mathematics of some of the most commonly used time-frequency transform, namely STFT, wavelet transform and Hilbert-Huang transform. These methods are tested on a sample EEG signal, and the wavelet transform was found to be the most appropriate for the analysis of EEG signal. The next section will move on from a discussion of time-frequency transform to discuss a method of spatial transform, common spatial pattern.

4.4 Spatial transform

While time-frequency transforms are powerful tools for analyzing the temporal and spectral dynamics of EEG signals, they primarily focus on how signal frequency content changes over time. However, EEG signals are not only defined by their temporal and spectral properties but also by their spatial distribution across multiple electrode channels. This is where spatial filtering techniques can come in. Spatial filtering of EEG signals involves enhancing or isolating specific neural activity by focusing on the spatial distribution of the signals across different electrodes. Spatial transforms complement time-frequency analyses by focusing on the spatial characteristics of EEG data.

4.4.1 Common spatial transform

One spatial transform technique used for EEG is the common spatial pattern (CSP) algorithm [146, 147, 148, 149]. The CSP algorithm is a complex mathematical algorithm which utilizes eigendecomposition, a mathematical technique to separate a matrix into its eigenvector and its corresponding eigenvalue, to amplify the variance between multi-channel EEG datasets. This makes CSP particularly valuable for identifying patterns of signal variances between different conditions or tasks, and discriminating between them [150].

In this subsection, the geometric interpretation of the CSP algorithm will be explored, which allows us to visualize how the CSP algorithm differentiates two different sets of EEG data.

4.4.1.1 Covariance matrix

The first step in the CSP algorithm is to generate the covariance matrix. The covariance matrix is a fundamental mathematical concept used in statistics, signal processing, and machine learning to describe the variability and relationships between multiple random variables. It expands on the concept of variance, which measures how much a single variable deviates from its mean, to a multivariate setting. Specifically, the covariance matrix captures the variance of each variable and the pairwise covariances between them, providing insight into how these variables interact.

Mathematically, let X be a dataset with n observations of p variables, where X is represented as an $n \times p$ matrix. Each column of X corresponds to a variable, and each row represents an observation. To compute the covariance matrix Σ, we first center the data by subtracting the mean of each variable from its respective column. This yields a mean-centered matrix X_c. The covariance matrix is then computed using equation 4.19:

$$\Sigma = \frac{1}{n-1} X_c^T X_c = \begin{bmatrix} \Sigma_{1,1} & \Sigma_{1,2} & ... & \Sigma_{1,j} \\ \Sigma_{2,1} & \Sigma_{2,2} & ... & \Sigma_{2,j} \\ ... & ... & ... & ... \\ \Sigma_{i,1} & \Sigma_{i,2} & ... & \Sigma_{i,j} \end{bmatrix} \tag{4.19}$$

where X_c^T is the transpose of the centered data matrix X_c, and the factor $\frac{1}{n-1}$ accounts for the degrees of freedom in the sample covariance estimation.

The resulting covariance matrix Σ is a $p \times p$ symmetric matrix, where the diagonal elements represent the variances of the individual variables, and the off-diagonal elements represent the covariances between pairs of variables. A positive covariance indicates that two variables tend to increase or decrease together, while a negative covariance suggests an inverse relationship. If the covariance is zero, the two variables are uncorrelated.

One important property of the covariance matrix is that it provides a compact representation of the data's spread and relationships in multiple dimensions. In the Common Spatial Pattern (CSP), the covariance matrix identify spatial patterns that maximize the variance between conditions.

4.4.1.2 Eigenvector and eigenvalues

The next step in the CSP algorithm is the eigendecomposition of the covariance matrix. To begin a discussion on how eigendecomposition works, we first need to discuss the concept of eigenvector and eigenvalues.

When we transform a vector into a different vector shape, the magnitude and direction of the vector changes. However, there are special vectors, which would not change direction regardless of the transformation. These vectors are known as eigenvectors. Eigenvectors can still change in magnitude after a transformation. The change in magnitude is known as eigenvalues. Mathematically, this can be expressed using equation 4.20:

$$A\overrightarrow{v} = \lambda\overrightarrow{v} \tag{4.20}$$

Where A is the transformation matrix, $\overrightarrow{v}$ is the eigenvector, and λ is the scalar eigenvalue. This equation can also be understood as: the linear transformation A of eigenvector $\overrightarrow{v}$ is equivalent to scaling $\overrightarrow{v}$ by a magnitude of λ.

4.4.1.3 Eigendecomposition

Eigendecomposition is the mathematical process used to decompose a square matrix, such as the covariance matrix, into its corresponding eigenvector and eigenvalue. The eigenvector set generated from a covariance matrix can be used as a spatial filter to project the signal onto a different space in order to maximise or minimise its variance.

Consider 2 eigenvectors, $\overrightarrow{v}_1$ and $\overrightarrow{v}_2$, which are transformed by the same linear transformation A. By taking $\overrightarrow{v} = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 \end{bmatrix}$ and substituting it into equation 4.20, we get equation 4.21:

$$A \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 \end{bmatrix} = \begin{bmatrix} \overrightarrow{v}_1 & \overrightarrow{v}_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{4.21}$$

Taking $\lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, and rearranging the above equation, we get equation 4.22:

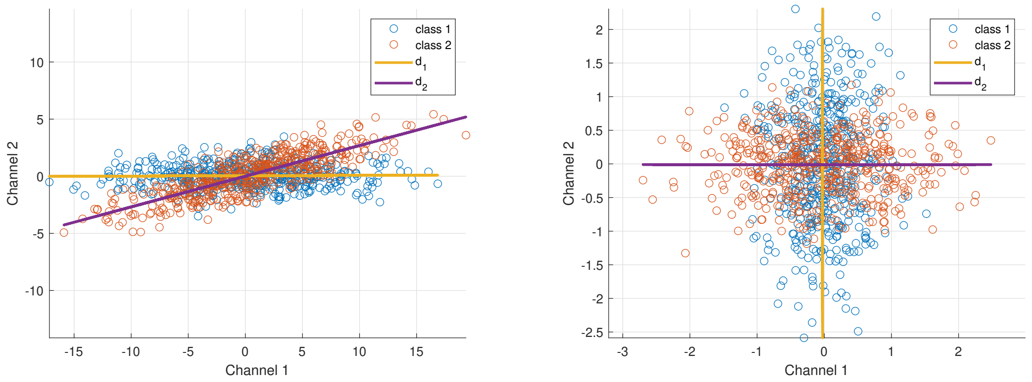
$$A = \overrightarrow{v} \lambda \overrightarrow{v}^{-1} \tag{4.22}$$

This is the formula for eigendecomposition, where a square matrix, a matrix with the same number of rows as columns, is decomposed into its eigenvectors and eigenvalues. The resulting eigenvector is then used as the CSP filter.

4.4.1.4 Generalized eigenvalue problem

Eigendecomposition only decomposed one signal at a time. However, the CSP filter works with two signals from two different channels, such that it maximizes the variance of one signal and simultaneously minimizes the variance of the other signal. This is achieved through solving a generalized eigenvalue problem, which ensures that the filters project the signals into a new space where the variances of the two conditions are maximally separable. Consider the signal from 2 EEG channels, each plotted on one axis. Figure 4.17 shows a visual intuition of how the CSP filter projects the signal to maximise variance:

As can be seen from Figure 4.17, the CSP filters was applied on 2 surrogate EEG channel signals. The signal from the 2 channels are plotted as coordinates on a 2-D graph: one channel on the x-axis and the other on the y-axis. Each moment in time becomes a point in this space, forming a cloud of points that reflects how the two channels vary together. Eigendecomposition finds new axes



(a) Before CSP filtering

(b) After CSP filtering

Figure 4.17: Visualisation of how CSP maximises the variance of the signal. D_1 and D_2 are the linear regression line of channel 1 and channel 2 respectively.

that best describe the structure of this cloud. In the context of CSP, these new axes are chosen so that the variance of the signal is maximised along one direction and minimised along the other. Graphically, it is like rotating the coordinate system so that one new axis aligns with the direction where the EEG points spread out the most, capturing the channel whose activity is strongest for that condition, while the second axis aligns with the direction where the spread is smallest, capturing the channel that is suppressed for that condition. The result is a pair of transformed "virtual channels" where one output emphasises the signal of interest and the other deemphasises it.

Mathematically, the generalised eigenvalue problem can be solved using using equation 4.23:

$$A\overrightarrow{v} = \lambda B\overrightarrow{v} \tag{4.23}$$

Where A and B are the matrixes to be decomposed, $\overrightarrow{v}$ is the common eigenvector set, or spatial filter, and λ is the corresponding eigenvalues. Here, the eigenvalues represent the variance ratio between the variances of the projected A and projected B. Mathematically, this can be expressed using the equation 4.24:

$$\frac{\text{Variance of projection of A}}{\text{Variance of projection of B}} = \lambda_i \tag{4.24}$$

This means that the variance ratio λ_i can be used to determine the effectiveness of the spatial filter $\overrightarrow{v}$. When $\lambda_i >> 1$, variance of A is maximised while variance of B is maximised. When $\lambda_i << 1$, variance of B is maximised while variance of A is maximised. When $\lambda_i \approx 1$, variance of both A and B are similar, so these projections are typically not as useful for discriminating between the classes and are often ignored in classification tasks. This means that the further λ_i is from 1, the more effective the filter $\overrightarrow{v}$ is in discriminating between A and B.

Once an appropriate spatial filter $\overrightarrow{v}$ have been calculated, the signals A and B can be multiplied by $\overrightarrow{v}$ to project A and B onto a space which maximises the differences between their variances.

4.5 Other spatial transform methods

There are other spatial transform techniques like connectivity analysis and beamforming technique that are commonly used on EEG signal. However, they are out of the scope of this thesis due to the fact that they require an fMRI machine to capture the structure of the participants' brains, which were not accessible throughout the duration of this research. While it is possible to use a generic brain model for every single participant, the results would not be as accurate. Hence, these techniques have been left out of the thesis. For interested readers, [151] and [152] have done fantastic tutorials on connectivity analysis and beamforming respectively.

4.6 Conclusion

In this chapter, we looked at some of the mathematical foundation upon which signal processing of EEG signals rely on. We compared some commonly used time-frequency techniques in EEG signal processing, namely STFT, Wavelet transform and Hilbert-Hwang transform. We concluded that Wavelet transform would work the best for EEG applications, due to the non-stationary, broadband and dynamic nature of EEG signal. We then examined CSP filters, a common spatial transform technique used in EEG signal processing to amplify the differences between the signals from 2 EEG channels. In the next chapter, we will examine some of the signal processing techniques which have been applied to EEG signals, and build a signal processing pipeline, which uses both wavelet transform and CSP filters, and will be used in the EEG experiment on visual processing of digital and non-digital stimuli.

Chapter 5

EEG signal processing techniques and pipeline

5.1 Introduction

In the previous chapter, the mathematics of signal processing techniques used for processing EEG signals were discussed. In this chapter, the signal processing techniques will be put together and applied to a public dataset to assess its effectiveness.

For many centuries, what goes on in our brain had been a mystery. In the past, the only way to study the brain is to obtain the brain slices from a deceased patient. In recent years, new neuroimaging techniques and technologies have emerged, which allows clinicians and researchers to examine patients' brains while they are still alive. The ability to observe brain activity provides researchers with valuable insights for developing tailored interventions for individuals. Research into brain function and structure can reveal how various factors, such as lesion location, functional requirements, and medical history, affect treatment approaches.

There are many different techniques which can be used to detect the signals from the brain. These techniques can be classified as invasive, semi-invasive and non-invasive. Rao (2013) and Eliassen et. al. (2008) have already written comprehensive details about how each of these techniques work [153, 154]. In this section, we will focus on the non-invasive techniques, in particular, electorencephalogram (EEG), magnetoencephalogram (MEG), functional near-infrared spectroscopy (fNIRs) and functional magnetic resonance imaging (fMRI). The works of Rao and Eliassen will be summarised, with examples on how each of these techniques are applied.

Electroencephalogram (EEG) measures the electrical signals from the brain through electrodes placed on the scalp. In particular, it measures the sum of postsynaptic potentials from the neurons oriented radially to the scalp [154]. The advantages of using EEG include good temporal resolution, portability and cheap implementation. However, EEG typically have poor spatial resolution, easily corruptible signals and may require intensive post-processing. EEG have been used to gather information about patient's recovery during rehabilitation [153]. Finnigan and Van Putten (2013) conducted a literature review on the different features and quantitative measures of EEG which have shown to provide value to medical diagnosis [155]. EEG have also been used to allow patients to control a robotic limb or exoskeleton. He et. al. (2018) conducted a literature review on robotic systems controlled by EEG and found that while EEG has some potential to be used as a robotic control system, more research would need to be conducted, with particular focus on EEG denoising techniques and clinical trials [156].

Magnetoencephalogram (MEG) works similarly to the EEG, except that instead of capturing electrical signals from the brain, it captures the magnetic signals. This is possible due to the concept of Maxwell's displacement current, which postulates that a changing electric field can induce a magnetic field. As the neurons fire electrical signals, these electrical signals can generate small magnetic field oscillations at right angles to the current flux, which can be picked up by an MEG system [153]. MEG are at times preferred over EEG due to its higher spatialtemporal resolution as compared to an EEG [154, 157]. Mellingeer (2007) developed an MEG-based brain computer interface (BCI) system and compared the training output with a state-of-the-art EEG-based BCI system [157]. They found that both MEG-based and EEG-based BCI systems have comparable performance, thus demonstrating the feasibility of using MEG for BCI applications. Another drawback to MEG is its large immobile hardware. However, in recent years, researchers have developed the optically pumped magnetometer (OPM) based MEG, which is capable of capturing brain signals even with large movements [158, 159].

Functional near-infrared spectroscopy (fNIRs) measures changes in blood oxygen level caused by increased neural activity in the brain, by detecting near-infrared light absorbance of hemoglobin in the blood, which could be oxygen-rich or oxygen-poor [154]. The advantages of using fNIRs include portability, no movement restriction, less signal corruption, and cheap implementation. However, fNIRs can only measure signal close to the skull, unlike fMRI. Rea et. al. (2014) first presented the evidence that fNIRs can detect the intention to move a lower limb in the brain signals of stroke patients. fNIRs have been used to investigate different brain activities in stroke patients [160]. Lin et. al. (2013) have used fNIRs to investigate cortical control while stroke patients are cycling, and Arun et. al. (2020) have used fNIRs to investigate the resting state pattern for the motor cortical network in stroke patients [161, 162]. While advanced signal processing techniques enabled fNIRs to be used for the diagnosis of certain neurological conditions, the temporal and spatial resolution of fNIRS will need to be improved for a more robust clinical usage within subjects [163].

Functional magnetic resonance imaging (fMRI) detects changes in blood flow due to increased activations of neurons in specific regions of the brain while the subject is performing certain tasks. As neurons become active, they would require more oxygen from the blood. The local capillaries will dilate in response to this increased need for blood, resulting in an increase in blood flow, which can be detected by fMRI [154]. fMRI has a much higher spatial resolution compared to other non-invasive techniques. However, it has poor temporal resolution, not portable, expensive, and subjects will not be able to move while being scanned. fMRI have been used to map the brain region responsible for various neurological functions, including those involved in brain plasticity and stroke recovery [164, 165, 166]. fMRI can also be used to predict each patient's treatment outcome by examining the lesion in their brain [167].

It is worth mentioning that these technologies can be combined to leverage on each of their advantages. [168] developed a multimodal neuroimaging system integrating OPM-based MEG, EEG, fNIRS. This system measures both brain electrophysiological activity and hemodynamic responses simultaneously. The OPM sensors provide flexible placement and enhanced signal power due to proximity to the scalp. The EEG cap is compatible with MEG and includes electrodes with minimal magnetic interference. The fNIRS component employs light sources and detectors for measuring cortical hemodynamics. These modalities are combined in a magnetically shielded environment, enabling high-resolution, simultaneous acquisition of neural and vascular signals, facilitating advanced studies of neurovascular coupling and potential applications like BCI and wearable neuroimaging.

Each of the techniques discussed above have their benefits and drawback, differing in factors such as spatial resolution, temporal resolution, portability and financial cost. For example, fMRI have excellent spatial resolution while lacking in temporal resolution, while EEG is the opposite. The key to choosing a system is to understand the needs and aims of the research project [169]. In this research, the system needs to be able to accommodate an HMD for virtual reality (VR) and augmented reality (AR) projection. This limits our choice to EEG, which is portable, cost effective and doesn't take up too much space on the patients' head. It is possible to design VR and AR systems that do not require an HMD, such as the Cave Automatic Virtual Environment (CAVE) system, which is a large, room-scale, multi-display immersive VR setup that creates visual immersion by surrounding the user with synchronized stereo displays [170]. However, those system would not be able to provide as much immersion as using an HMD. Hence, the rest of this chapter will focus on how EEG signals are processed.

5.2 Illustrations of EEG systems

Figure 5.1 shows a person wearing an EEG cap, which is designed to hold electrodes in predefined locations on the scalp. The cap facilitates consistent electrode placement across participants by conforming to standardized electrode positioning systems.

Figure 5.1: Person wearing an EEG cap

EEG electrodes measure electrical potentials generated by synchronized neural activity at specific scalp locations. EEG electrodes can broadly be classified into wet and dry types. Wet electrodes require the application of a conductive gel to reduce electrode-skin impedance and generally provide higher signal quality, while dry electrodes do not require gel and offer faster setup and improved usability at the expense of increased susceptibility to noise [171]. Electrodes may also be categorized as active or passive: active electrodes incorporate a built-in preamplifier at the electrode site to improve signal quality, whereas passive electrodes transmit the raw signal to an external amplifier [172].

Figure 5.2 shows a schematic of a typical EEG acquisition system, adapted from Velarde et. al. [11].

EEG electrodes are connected to amplification circuitry, either integrated within the electrodes themselves or located in a central acquisition unit. The amplified analog signals are then digitized via an analog-to-digital converter (ADC). In many systems, analog filters are applied prior to digitization, such

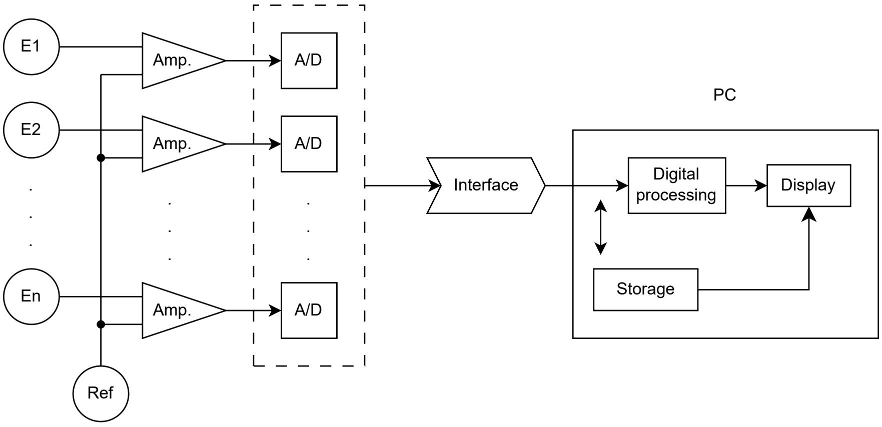


Figure 5.2: Typical EEG acquisition system schematic [11]

as high-pass filters to remove slow drifts and low-pass or notch filters to suppress high-frequency noise and power-line interference. The digitized signals are subsequently transmitted to a computer via a communication interface, where they can be processed, visualized, and stored for offline analysis [11].

Figure 5.3 shows an example of raw EEG signals visualized using the FieldTrip toolbox [12].

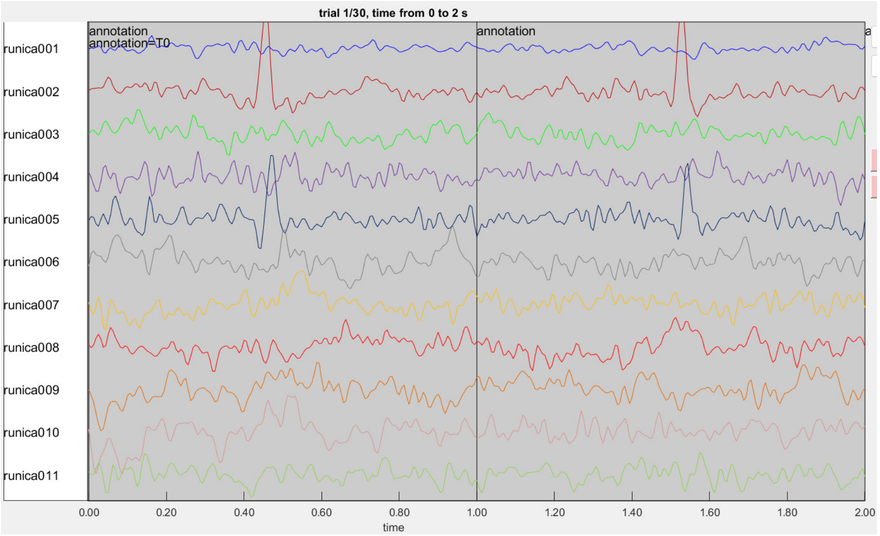


Figure 5.3: Example of raw EEG signals visualized using FieldTrip [12]

Visual inspection of raw EEG signals allows for the identification of artifacts such as eye blinks, muscle activity, or excessive noise. Trials containing severe artifacts may be corrected using artifact removal techniques or excluded from further analysis.

Figure 5.4 shows an example of a scalp topographical map generated using

FieldTrip [12].

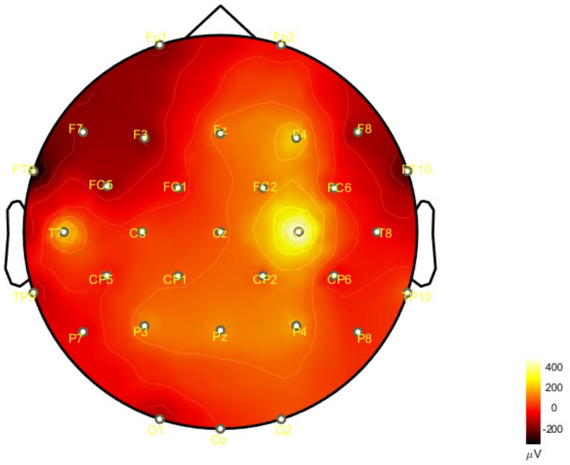


Figure 5.4: EEG scalp topographical map generated using FieldTrip [12]

Scalp topographical maps provide a spatial representation of EEG activity across the scalp, enabling researchers to identify regions associated with task-related neural activity. These spatial patterns can then be interpreted in the context of known neurophysiological mechanisms and cortical functional organization.

5.3 Brain-computer interface

The aim of this thesis is to study the visual perception (VP) of various subjects. One way of doing that is with the use of Brain-Computer Interface (BCI) systems. BCI systems are systems which allows for information to be directly transmitted from the brain to a computer system, without relying on the brain's regular sensorimotor neural pathways [173, 174, 175, 176, 177]. BCI systems work by first capturing electrical signals generated from neuron clusters within the brain, using neuroimaging techniques such as EEG. These signals are then converted to digital output, which can then be analysed or even used to control external devices or computer system [178].

One of the most studied controlled signal types is Motor imagery (MI). MI may be defined as a conscious mental state during an internally rehearsed motor act with any motor output [179, 180, 181]. In the context of BCI, MI involves getting participants to mentally rehearse a predefined set of actions, while their brain signals are being captured by EEG electrodes. These brain signals can then be processed and analysed to study the brain processes involved in initiating and generating motion [182, 183, 176]. While MI and VP are fundamentally different neural activities, the techniques used to analysed the EEG signals generated from both neural activities can be similar. Hence, in this chapter, techniques used to analyse MI EEG signals will be reviewed, and tested on a public MI dataset.

The EEG analysis of MI signals aims to discriminate between the left MI from the right MI by using the special characteristics of event-related desynchronization and synchronization (ERD and ERS, respectively) in mu and beta rhythms over the sensorimotor cortices during MI tasks [184]. Mu and beta rhythms are sensorimotor EEG oscillations (mu $\approx$ 8–13 Hz, beta $\approx$ 13–30 Hz) that are strongest over central scalp regions and typically decrease in power during movement execution, motor imagery, or action observation, reflecting sensorimotor cortical engagement [185]. On the other hand, the aim of this research is to discriminate between EEG signals when different types of visual stimuli are being perceived, namely 2D and 3D stimuli, or real and virtual stimuli. In processing both MI and VP signals, the accuracy of classification depends heavily on the quality of the noise removal, feature extraction and feature selection algorithms. The methods for each of these techniques will be discussed later in this chapter.

Hamedi et al. (2016) conducted a review on current research into EEG-based MI brain connectivity analysis [183]. Past research on brain regions involved in MI were detailed, and challenges in using BCI systems for motor imagery were identified, including motor variability and damage to parts of the brain involved in motor imagery. Saha and Baumert (2020) conducted a review on inter-session and inter-subject performance predictors could potentially improve BCI performance [186]. In their review, they discussed the factors which affects motor variability, which includes individual differences in motor learning rates, structural and functional differences in brain networks, and the degree of motor impairment in patients. Saha found that the EEG patterns associated with motor variability partly explain intra-individual variability in sensorimotor rhythm-based BCIs. Higher motor variability may allow some individuals to learn BCI skills faster. Additionally, alterations in EEG signatures due to motor training are dependent on intra- and inter-subject variability.

5.4 EEG signal processing pipeline

BCI techniques do not work with just the raw EEG signal. Instead, the signal need to be processed before they can be interpreted or used. Various signal processing and machine learning techniques have been used to process brain signals. Figure 5.5 shows a typical pipeline used to proccess EEG signals.

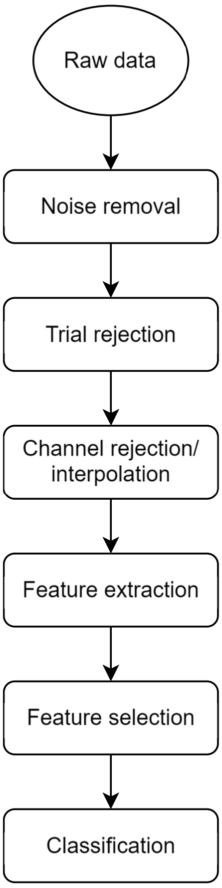


Figure 5.5: Typical EEG processing pipeline

In the next few subsections, each of these steps in the EEG processing pipeline will be discussed in detail.

5.5 Noise removal

EEG signals are inherently noisy due to spontaneous synaptic activity, which introduces fluctuations in membrane potentials. These fluctuations cause random noise that, while part of normal neural activity, can mask or interfere with detecting specific brain signals [187]. Such noises are known as intrinsic noise. On the other hand, noise can also come from external sources, such as muscle activity (EMG), eye movements (EOG), power line interference, and other artifacts. Such noise are known as contamination noise.

Noise in EEG requires specific filtering techniques to separate genuine neural activity from these external interferences for more accurate analysis. As discussed in Chapter 4, specific frequency filters can be used to filter out signal components with frquencies which are either too high or too low. A low-pass filter can remove high-frequency noise, such as muscle artifacts, while preserving the slower brainwave frequencies. A high-pass filter can eliminate low-frequency drifts and artifacts, like those caused by sweat or slow eye movements, which can obscure the relevant brainwave activity. Band-pass filters are particularly useful for isolating specific frequency bands of interest, allowing the extraction of meaningful neural activity while filtering out both lower and higher frequency noise [188, 189].

Lai et. al. (2018) conducted a comprehensive review of various EEG noise removal techniques, including Independent Component Analysis (ICA), statistical methods, wavelet-based approaches, adaptive filtering, and machine learning. ICA, a widely used technique, decomposes EEG signals into independent components, enabling the identification and removal of artifacts. Statistical methods, such as Multivariate Empirical Mode Decomposition (MEMD), are employed to isolate artifact components. Wavelet-based techniques are often combined with ICA or statistical approaches to further decompose signals and identify artifacts. Adaptive filtering has demonstrated effectiveness in removing specific artifacts, such as eye blinks [190, 189]. More recently, machine learning methods have been introduced, utilizing classifiers trained on labeled artifact data to automate the detection and removal of noise. The review concluded that hybrid approaches, which integrate multiple techniques, achieve the best performance by effectively eliminating artifacts while preserving critical signal information [190, 154].

One of the techniques mentioned, ICA, is particularly effective for separating mixed signals into their original sources. In the context of EEG, each electrode picks up signal from different part of the brain. They can also pick up artifacts, such as eye movements, muscle activity, or electrical interference. With multiple electrodes places on different positions of the scalp, the individual source signal needs to be extract from the mixed signals captured by the electrodes, without prior knowledge of the sources or the mixing process. This is known as the cocktail party problem. ICA assumes that the mixed signals are linear combinations of statistically independent source signals and that at least one source has a non-Gaussian distribution. By estimating a demixing matrix, ICA separates the mixed signals into independent components that correspond to the original sound sources. An example of such a matrix would be the JADE (Joint Approximate Diagonalization of Eigen-matrices) algorithm, which uses higher-order statistical properties to ensure effective separation of sources [191, 192]. As Iriarte et. al. (2003) have found, ICA can be a useful tool to clean artifacts in short EEG samples, without having the disadvantages associated with digital filters [193]. There are many demixing matrices available, and readers who are interested in the mathematical treatment of ICA demixing matrices may refer to Hyvärinen et.al. (2001) [194].

Although ICA is a powerful tools to separate and remove noisy signals from the desired signal, they require the researcher to 1) have knowledge of how the noisy component looks like, and 2) visually inspect each component and manually remove the noisy component. This might not be feasible especially if the dataset is a public set, and the researcher analysing the dataset might not have complete information on the noise sources that may be present during the experiment. Also, visual inspection of EEG signals is labour intensive. Hence, this might be feasible if the dataset is small, but not if the dataset is big.

5.5.1 Trial rejection

Noise removal techniques are not magic erasers. They are not perfect in removing all noises while at the same time leaving the desired signal untouched. There are instances when a trial is just too noisy, perhaps due to excessive movement during the trial, and the only option left is to remove them from processing. This is typically done through a visual inspection of the EEG signal from each trial. During this step, a researcher would typically look through the signal from each trial and check if there are any unusual peaks, troughs, or entropy.

While visual inspection might be sufficient for a small dataset, when the dataset gets larger, visually inspecting every trial will become tedious. As with the decomposition methods of noise removal discussed above, this might be feasible for smaller datasets but not for larger ones. A researcher might then consider automation to process larger datasets. One effective way to automate this is by setting predefined thresholds for key signal metrics, such as amplitude, variance, or power, to flag and remove trials with excessive artifacts. For example, trials exceeding a certain amplitude range or exhibiting abnormally high variance across channels can be automatically marked for rejection.

One issue with defining a global threshold is that it is not optimised to account for different sensor noise level and experimental conditions, creating a risk of overrejection and losing relevant EEG signals. To address this issue, Jas et. al. (2016) introduced “Autoreject“, an innovative algorithm designed for the automatic detection and correction of artifacts in single-trial M/EEG data [13]. Instead of simply defining the trial, Jas identifies the bad channels, then use that information to repair the channels if the number of bad channels identified does not exceed a predefined limit, and reject the trial if the limit is exceeded.

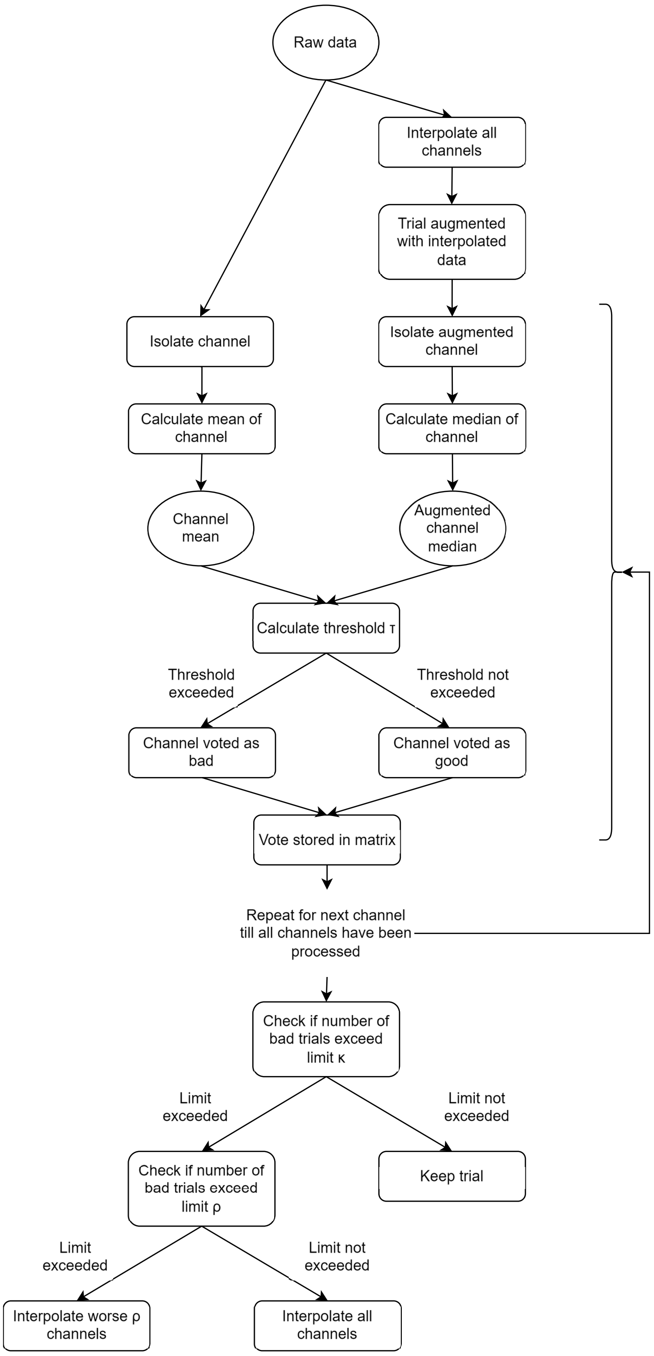


Figure 5.6: [13]'s autoreject (local) pipeline

The parameters τ, κ and ρ are determined through a grid search method aimed at minimizing a specific error metric. This is done using the following calculations. Say we have a set of trials X, and thresholds L. For each threshold l, where $l \in L$, a trial x is considered a good trial when the PTP of x is less then l. Mathematically, this can be expressed as:

$$G_l = x \in X | ptp(x) < l \tag{5.1}$$

Next, the error matrix e_{xl} for the cross validation of one trial x and one threshold l is calculated using equation 5.2

$$e_{xl} = \left\| \bar{X}_{G_l} - \tilde{X}_{aug} \right\|_{Fro} \tag{5.2}$$

where $\|\cdot\|_{Fro}$ is the Frobenius norm. The root mean square error (RMSE) is computed between the mean of the good trials G_l with the median of the augmented trials X_{aug}. Here, X_{aug} is used as the validation set. This process of cross validation will find an optimal value which rejects as many bad trials as possible, while keeping as many good trials as possible. Autoreject consistently matches or surpasses the performance of current alternatives, highlighting the importance of sensor- and subject-specific parameter optimization.

5.5.2 Channel rejection/interpolation

Channel rejection is similar to trial rejection, in that channels that are deemed to be too noisy are rejected. While this might work if the EEG was recorded with a sufficiently large number of electrodes, the loss of information from one electrode might significantly skew the results, such as in the case of EEG studies focused on a specific region of the brain. Furthermore, in BCI applications, the continuity of the EEG data might be critical. Hence, many EEG researchers opt to interpolate the bad EEG channel instead.

Channel interpolation in EEG is a technique used to repair or reconstruct data from faulty or noisy electrodes that produce unreliable signals, ensuring the integrity of the overall dataset. When a channel is identified as problematic due to noise, poor connectivity, or other issues, it is temporarily excluded, and its signal is estimated using information from the surrounding channels. This is typically achieved by applying spatial interpolation methods, such as the spherical, barycentric or spline interpolation [195], which consider the geometric arrangement of electrodes on the scalp. These methods assume that EEG signals vary smoothly across neighbouring electrodes, allowing the missing data to be approximated by weighting the contributions of nearby channels based on their distances and relationships.

Perrin et. al. (1989) proposed the spherical spline method, which assumes that the head can be modelled as a sphere [196]. From there, this technique calculates interpolated potentials using a combination of weights (coefficients) that are determined by solving a system of equations derived from the known potential values at electrode locations. These weights are applied to a function of the cosine of the angular distance between points on the sphere, ensuring a smooth and continuous map. Spherical splines are particularly advantageous compared to other methods like thin plate splines because they are computationally efficient and maintain accuracy in areas with fewer electrodes. Additionally, the spherical spline framework simplifies the computation of scalp current density by directly using the spherical Laplacian of the potential. This method improves the precision of EEG mapping and current density estimation, even in undersampled regions, while also being faster to compute. Lastly, this method is also integrated within the Fieldtrip toolbox set, making it's implementation seemlessly easy [12].

Nouira et. al. (2014) propose a novel framework for EEG channel interpolation using deep encoder-decoder networks, which autonomously identifies the spatio-temporal properties of EEG data to predict the values of missing channels [195]. Their method is designed to effectively interpolate EEG data for any missing channel at any time, leveraging transfer learning to fine-tune the model using clean data from new subjects or tasks. This approach not only enhances the accuracy of EEG data reconstruction but also allows for practical application across various datasets without the need for extensive electrode localization. However, due to the need to train a deep learning model, this method can be more computationally intensive than the spherical spline method, especially during the training phase.

5.5.3 Feature extraction

After the noisy trials and channels have been dealt with, the next step in the EEG signal processing pipeline is to then extract meaningful features that can represent the underlying neural activity. Feature extraction involves applying mathematical transformations such as time-frequency transform and spatial transform approaches. As Chapter 4 has already discussed extensively about the mathematics behind these transforms, this subsection will present examples of these transforms applied on EEG signals.

Yamawaki et. al (2006) proposed two novel methods on time-frequency feature extraction, expression and classification for high-density EEG recordings [197]. The first method was a single trial motor imagery (MI) classification strategy for the brain–computer interface (BCI) applications by using time–frequency synthesis approach to accommodate the individual difference, and using the spatial patterns derived from electroencephalogram (EEG) rhythmic components as the feature description [198]. The second method was a new algorithm by means of frequency decomposition and weighting synthesis strategy for recognizing imagined right- and left-hand movements [199]. They then applied it to a one-dimensional “cursor control” BCI experiment with online feedback [197].

Ferrante et. al. (2015) developed a classifier that combines Morlet wavelets and Common Spatial Pattern (CSP) algorithms to address the noisy and nonstationary nature of EEG signals [200]. The method achieves an average accuracy of 88% across subjects using only about 10-15 training examples per class, making it significantly more data-efficient than existing approaches that typically require hundreds or thousands of trials. This algorithm exploits both the oscillation frequency range (focusing on the μ-rhythm EEG frequency band at 7.5-12 Hz) and scalp location information relevant to motor imagery. The researchers tested the method on 8 healthy subjects performing left/right hand motor imagery tasks, using single-trial analysis rather than averaging over multiple trials. The data-efficiency of this approach could potentially reduce BCI training times by up to 100-fold, which may lead to improved adoption rates and make BCIs more feasible for larger patient groups.

Murugappan (2011) developed a pipeline following a similar principle of combining time-frequency filter and spatial filter [50]. They started the preprocessing with the Surface Laplacian (SL) filter to remove artifacts and noise, improving spatial resolution. Following this, the signals are decomposed into five frequency bands (δ, θ, α, β, and γ) using a Discrete Wavelet Transform (DWT) with three wavelet functions, namely db8, sym8 and coif5. Both linear (e.g., power, variance) and non-linear (e.g., entropy) statistical features are extracted from each band. The extracted features are then fed into two classification models: K-Nearest Neighbor (KNN) and Linear Discriminant Analysis (LDA), with the classification process validated using 5-fold cross-validation [50]. Murugappan found that KNN performed better than LDA, particularly when using entropy as a feature. The study demonstrates the effectiveness of combining spatial filtering, wavelet-based feature extraction, and simple classifiers for emotion recognition through EEG signals.

In summary, combining a time-frequency transform and a spatial transform in EEG analysis provides a powerful approach for capturing both the temporal and spatial dynamics of brain activity. Together, these methods complement each other: time-frequency transforms capture the temporal and spectral characteristics of neural processes, while spatial transforms identify their anatomical or functional origins. This combination improves the resolution and interpretability of EEG data, and increases the discriminability between signals so as to enhance the classification of signals.

5.5.4 Feature selection

After the EEG signals have been transformed into meaningful representations (features), the next step in the pipeline is feature selection. This step narrows down these representations to retain only the most task-relevant ones.

Feature selection can be done using an understanding of the brain physiology. Firstly, a knowledge on the function of each part of the brain allows researchers to focus on the electrodes placed on the relevant parts of the brain. Figure 5.7 illustrates how electrodes can be isolated based on the type of experiment conducted:

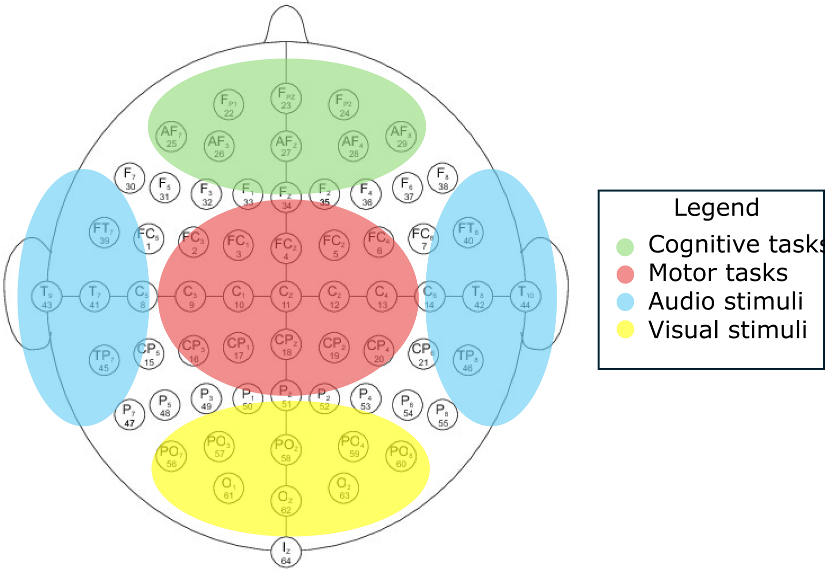


Figure 5.7: General areas where activations can be expected in specific tasks and when specific stimuli are presented. Please note that this is not meant to be a comprehensive diagram on the areas of brain activation, but rather a basic illustration to show that electrodes can be isolated based on the type of experiment conducted.

Secondly, different brain activities and cognitive state emit electrical signals in different frequency bands, and the relevant frequency bands can be selected depending on the type of experiment conducted. For instance, delta waves (0.5-4

Hz) are linked to deep sleep, theta waves (4-8 Hz) are associated with drowsiness and light sleep, alpha waves (8-13 Hz) are connected to resting activity of the occipital cortex and normal adult wakefulness, and beta waves (13-30 Hz) are related to resting activity of the precentral cortex, speech and intellectual effort [201]. Once the desired frequency band has been identified, a band-pass filter can be used to isolate the desired frequency band.

Feature selection techniques in EEG signal processing can also be automated. However, this thesis focuses more on knowledge on brain physiology for feature selection rather than computational methods, so automated feature selection techniques are beyond the scope of this thesis.

5.5.5 Classification

Once the features have been extracted and selected, the next step would be to classify the features. Traditionally, classification of EEG data is done using statistical methods. Recently, however, machine learning techniques have been used to classify EEG data with reasonable accuracy.

Statistical methods are analytical techniques used to summarize, interpret, and draw conclusions from data by identifying patterns, relationships, and significant differences through mathematical calculations and probabilistic reasoning. Two commonly used statistical methods in EEG data analysis are Analysis of Variance (ANOVA) and the independent t-test [202, 203]. These methods are designed to test differences between groups or conditions. The independent t-test compares the means of a single variable between two independent groups, such as patients and controls, providing a simple measure of whether the difference is statistically significant. However, the t-test is limited to two groups or conditions. In contrast, ANOVA can compare the means of a dependent variable across multiple groups or conditions, making it more suitable for experiments with more than two levels. While the t-test evaluates one variable at a time, ANOVA extends this by considering variability within and between groups, producing an overall

F-statistic to indicate whether significant differences exist among groups. Both methods require assumptions like normality and homogeneity of variance, and in EEG studies with multiple comparisons, corrections like Bonferroni or false discovery rate (FDR) are often applied to reduce the likelihood of false positives [204].

On the other hand, machine learning methods are computational algorithms that enable systems to automatically learn patterns and make predictions or decisions from data without being explicitly programmed. Machine learning algorithms such as Support Vector Machines (SVM), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks are widely used algorithms for the classification of EEG data, each leveraging different strengths. SVM is a powerful supervised learning algorithm that works well with high-dimensional EEG data by finding the optimal hyperplane to separate classes, making it particularly effective for linear and non-linear EEG feature classification, such as mental state or motor imagery detection[205]. CNN, on the other hand, excels at capturing spatial and temporal patterns in raw EEG data by learning hierarchical features directly from input matrices, such as spectrograms or raw signals, without the need for manual feature extraction. They are especially effective in identifying spatial dependencies across electrodes [206, 207]. LSTM networks, a type of recurrent neural network (RNN), are designed to handle sequential data, making them ideal for classifying EEG signals by modeling temporal dependencies, such as changes in brain activity over time [208]. While SVMs are more interpretable and computationally efficient, CNNs and LSTMs often achieve higher accuracy in complex EEG tasks by learning both spatial and temporal patterns, though they require larger datasets and greater computational resources for training. Together, these algorithms provide versatile tools for analyzing and classifying EEG signals in diverse applications like brain-computer interfaces, emotion recognition, and seizure detection.

Machine learning methods and statistical methods each offer distinct advantages in EEG data analysis. Statistical methods, such as ANOVA and independent t-tests, are interpretable, computationally efficient, and well-suited for hypothesis-driven studies where specific relationships between variables need to be quantified. They provide robust tools for identifying significant effects and trends in small to moderate datasets. In contrast, machine learning methods, such as SVM, CNN and LSTM, excel at handling the high dimensionality, complexity, and non-linear nature of EEG data. Machine learning can automatically detect patterns, optimize features, and classify multi-dimensional data, making it ideal for larger, more complex datasets. With the power that machine learning methods offer, they often require extensive computational resources and larger training datasets. With the strengths and weaknesses of both methods, it is important to understand the requirements of a research project before deciding on the method to be used.

While these machine learning methods have demonstrated promising performance in large-scale EEG datasets by learning discriminative features directly from raw or minimally processed signals, their effectiveness is strongly dependent on the availability of substantial training data [209] and careful model regularization [210]. In this thesis, preliminary experiments were conducted using a support vector machine classifier. However, classification accuracy remained close to chance level, likely due to the limited number of subjects and trials. More complex deep learning architectures, such as LSTM-based models, were therefore not pursued, as the available dataset was insufficient to support reliable training and validation. Consequently, this work focuses on traditional signal processing and statistical analysis methods, which are more suitable for small-sample EEG studies, such as the visual processing study conducted within this thesis, and offer greater interpretability.

5.6 Conclusion

In this chapter, we discussed a general EEG processing pipeline, and some examples on each step of the pipeline. In the next chapter, we will apply this pipeline to a public dataset.

Chapter 6

Applying processing pipeline to motor imagery dataset

6.1 Introduction

In the previous chapter, a general EEG processing pipeline was discussed. In this chapter, a proposed pipeline, based on the previously discussed techniques, will be presented and applied to a motor imagery (MI) dataset. The purpose of applying the pipeline to the dataset will be to validate the pipeline for the experiment conducted for this thesis.

6.2 Methods

6.2.1 Dataset

A dataset containing a set of 64-channel EEGs from subjects who performed a series of motor/imagery tasks [14] was obtained from PhysioNet [211] by the developers of the BCI2000 instrumentation system for brain-computer interface research [212].

In this dataset, participants were asked to either perform or imagine a set of motor movements in response to a visual stimulus presented on screen. The visual stimulus is a target which would appear on either the left or the right side of the screen. The participant will then perform or imagine moving the limb, either their hand or leg, on the same side as the target on screen. While this is being done, the participant's EEG signal will be recorded by the BCI2000 system. The EEG signal were recorded from 64 electrodes as per the international 10-10 system (excluding electrodes Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9, and P10) [14]. Figure 6.1 illustrates the placement of the electrodes on the participants' scalp:

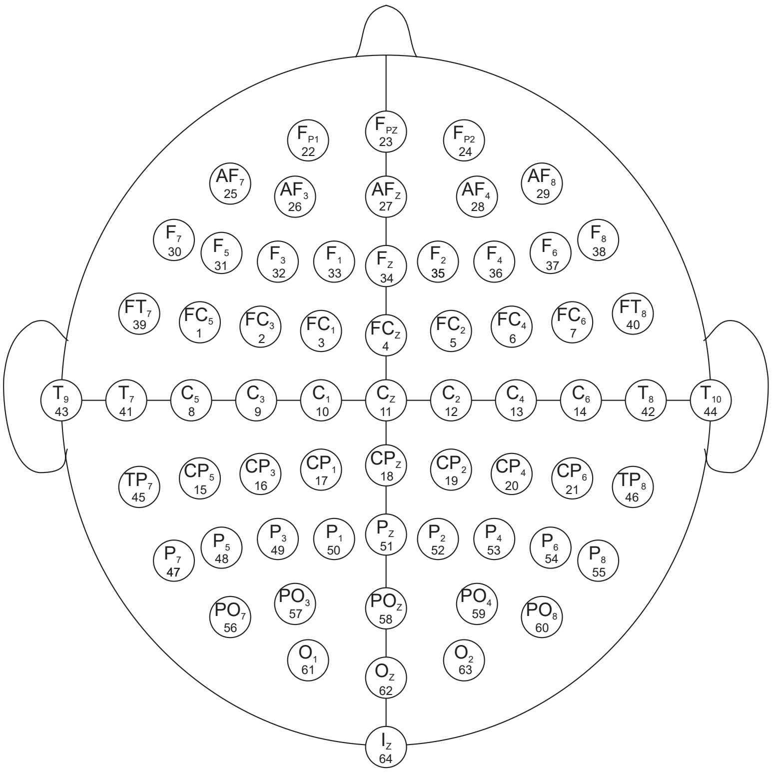


Figure 6.1: 10-10 electrode placement system used in motor imagery dataset[14].

While there are four different protocols used in this dataset: real hand motion, imagined hand motion, real leg motion, imagined leg motion. While it is possible to analyse the whole dataset, it should be considered that hand MI usually produces stronger, more localized EEG signals compared to leg MI [213, 150]. Furthermore, since the EEG visual perception experiment involves responding to stimuli with a button press, which involves finger movement, and it would more closely resemble hand MI. Hence, to simplify the analysis process, the analysis performed in this chapter will focus on the protocol where the participants imagine their hand motion.

6.3 Procedure

Figure 6.2 summarises the pipeline used for the motor imagery dataset.

The analysis was done with Matlab and Fieldtrip [12]. The data is first

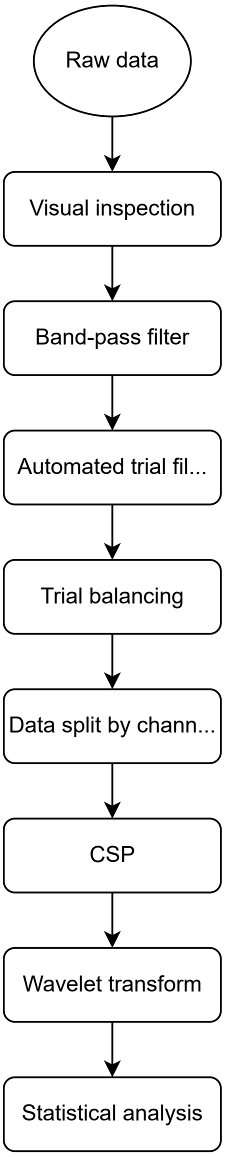


Figure 6.2: Proposed EEG processing pipeline

imported into Matlab and converted into Fieldtrip readable format. The imported data is then spliced into epoches. Each epoch is 2s long, both from the time the stimuli is shown and the time the stimuli disappears (baseline). Each epoch is then inspected through a visual inspection. The signal was then demeaned, before a 4th-order, butterworth, infinite impulse response (IIR) band-pass filter between 1 Hz and 45 Hz was applied. The power spectral density of the signals before and after filtering were inspected to ensure that frequency components outside the target range were effectively attenuated while preserving relevant neural activity. Next, an automated trial rejection algorithm, modified based on the trial rejection algorithm proposed by Jas et. al. (2016) [13], was applied.

Most of the steps outlined by Jas in their algorithm were followed, except the last step. In Jas' algorithm, the trials that goes over the threshold have their bad channels interpolated. The number of bad channels interpolated was based on another threshold, as outlined in Chapter 5. In the proposed algorithm, to reduce the computational load and speed up each iteration of the algorithm, the last step was modified to just reject the trial instead.

Following the trial rejection, each trial is then split by channel, based on which side of the motor cortex the channel is on. C5, C3 and C1 correspond to the left motor cortex and C2, C4 and C6 correspond to the right motor cortex.

The resulting data were spatially filtered using the Common Spatial Patterns (CSP) algorithm. CSP was applied to pairs of conditions (motor imagery vs. baseline) and computed at the class level. For each condition, trial-wise covariance matrices were first estimated and normalized by their trace to ensure equal contribution of each trial regardless of signal amplitude. These normalized covariance matrices were then averaged within each condition, allowing unequal numbers of trials across conditions without bias. To improve numerical stability and reduce overfitting, particularly given the large number of baseline trials, covariance matrices were regularized using shrinkage toward the identity matrix. Spatial filters were obtained by solving the generalized eigenvalue problem of the

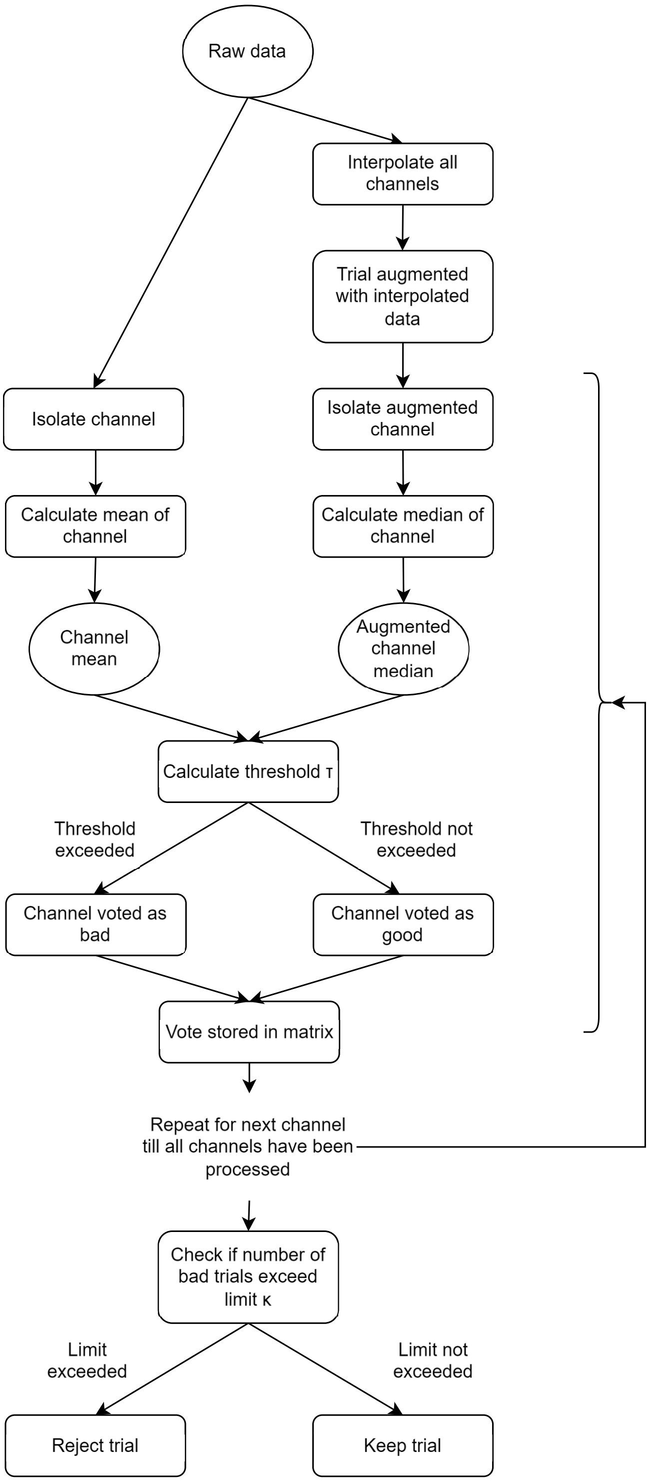


Figure 6.3: Proposed autoreject (local) pipeline

two regularized class covariance matrices, yielding components that maximized variance for one condition while minimizing it for the other. The resulting CSP filters were subsequently applied to all trials, producing a reduced set of spatially filtered time series for further time–frequency analysis.

After applying the CSP filters, time–frequency representations of the resulting signals were computed using a continuous wavelet transform with complex Morlet wavelets. The wavelet transform was applied separately to each CSP component and each trial. Complex wavelet coefficients were converted to power by taking the squared magnitude, and no averaging was performed in the complex domain to avoid phase cancellation effects. The resulting wavelet power was then averaged across trials within each participant and condition, yielding participant-level time–frequency representations. The mathematical details of each of the transforms can be found in Chapter 4.

Statistical analysis was performed using IBM SPSS Statistics for Windows, version 31 (Armonk, NY: IBM Corp). Repeated-measures t-test is applied to compare the participant-level time–frequency representations during baseline with each side of the motor imagery. To compare both baseline/ motor imagery and left MI/ right MI, a 2x2 repeated measures ANOVA was performed.

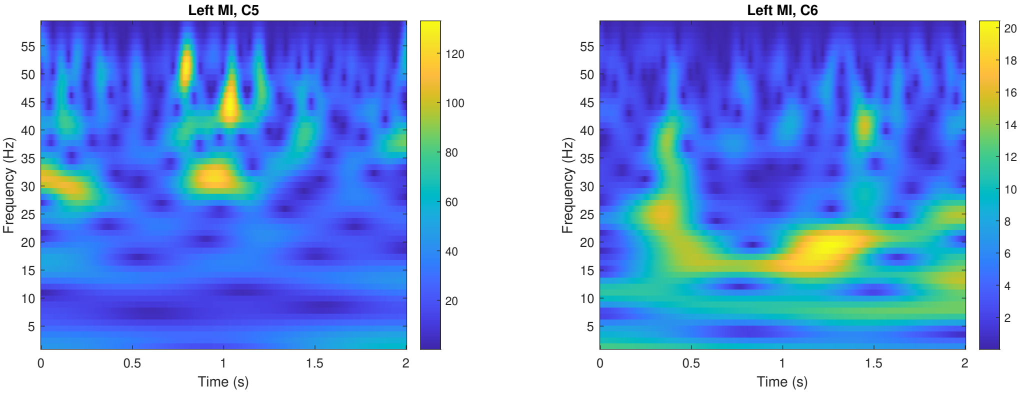
6.4 Results

6.4.1 Visualisation of wavelet transform

Figure 6.4 shows the wavelet transform of a sample trial on channel C5 and C6 during left hand motor imagery.

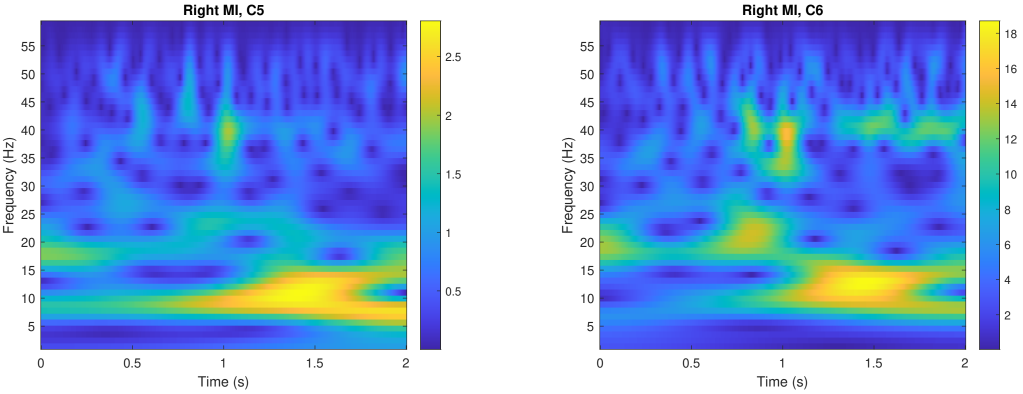
Figure 6.5 shows the wavelet transform of a sample trial on channel C5 and C6 during right hand motor imagery.

Figure 6.4a shows the time–frequency representation of a representative trial during left motor imagery at electrode C5. In this trial, an increase in power was observed in the μ band (7–13 Hz) approximately 0–1 s following stimulus onset. This increase was not observed at the electrode C6 during left motor imagery



(a) Left motor imagery (b) Left motor imagery

Figure 6.4: Wavelet of sample trial of channel C5 and C6 during left motor imagery



(a) Right motor imagery (b) Right motor imagery

Figure 6.5: Wavelet of sample trial of channel C5 and C6 during right motor imagery

(Figure 6.4b). Furthermore, no comparable μ-band power increase was evident at either C5 or C6 during right motor imagery (Figures 6.5a and 6.5b).

As shown in Figure 6.1, electrode C5 overlays the left motor cortex, whereas C6 overlays the right motor cortex. Due to the contralateral organization of the motor system, activation of the left motor cortex is associated with motor processes involving the right side of the body. The μ-band power increase observed at C5 during left motor imagery is therefore consistent with established neurophysiological models of motor imagery–related cortical activation. The absence of a similarly clear pattern during right motor imagery may be related to participant characteristics, such as right-handedness, although this interpretation remains speculative.

Based on this observation, a hypothesis was formed that isolating μ-band activity in electrodes overlying the left motor cortex (C5, C3, and C1) would enhance the discriminability between left and right motor imagery EEG signals. This hypothesis was subsequently tested using the full dataset and participant-level statistical analyses. To provide a more comprehensive characterization of motor cortical activity, electrodes overlying the right motor cortex (C2, C4, and C6) were also included in the analysis.

Following the hypothesis, a 4th-order, butterworth, infinite impulse response (IIR) band-pass filter between 7 Hz and 13 Hz was applied on the demeaned signal, and the analysis pipeline was rerun with the updated band-pass filter.

6.4.2 Statistical analysis comparing left hand motor imagery with baseline

Table 6.1, 6.2 and 6.3 show the descriptive statistics the average power of each participant, comparing left MI and baseline, after CSP and Morlet wavelet transform.

Table 6.4 shows the result of the repeated-measures t-test comparing the mean power of each participant.

Table 6.1: Table for descriptive statistics for left motor imagery and baseline

Average power	Mean	Sample size	Std. deviation
Left	119.018	522	162.539
Baseline	125.995	522	157.911

Table 6.2: Table for descriptive statistics for each side of the motor cortex during left motor imagery and baseline

Average power	Mean	Sample size	Std. deviation
Left C5,C3,C1	138.781	261	190.557
Left C2,C4,C6	99.255	261	125.869
Baseline C5,C3,C1	141.341	261	173.325
Baseline C2,C4,C6	110.649	261	139.476

Table 6.3: Table for descriptive statistics for each electrode during left motor imagery and baseline

Average power	Mean	Sample size	Std. deviation
Left C5	121.539	87	108.317
Left C3	162.184	87	168.647
Left C1	132.621	87	262.132
Left C2	116.347	87	115.110
Left C4	31.181	87	29.889
Left C6	150.238	87	161.682
Baseline C5	131.792	87	112.373
Baseline C3	167.090	87	156.929
Baseline C1	125.141	87	229.217
Baseline C2	141.612	87	155.685
Baseline C4	32.000	87	30.442
Baseline C6	158.339	87	155.306

Table 6.4: Repeated-measures t-test comparing left motor imagery and baseline EEG

	t	df	2-tail p	Mean diff	SD diff	Lower	Upper
Left,Baseline	2.910	521	0.004	6.977	54.773	2.267	11.686

It was found that the average power during left motor imagery, at 7 Hz to 13 Hz, from C1 to C6, was statistically significantly lower (119.018 ± 162.539) compared to baseline EEG (125.995 ± 157.991), $t(521) = 2.910$, $p = 0.004$.

Table 6.5 shows the result of the repeated-measures t-test comparing the mean power of each participant on each side of the motor cortex, assuming equal variances.

Table 6.5: Repeated-measures t-test comparing left motor imagery with baseline EEG for each side of the motor cortex

	t	df	2-tail p	Mean diff	SD diff	Lower	Upper
Left,Baseline C5, C3, C1	0.787	260	0.432	2.560	52.545	−3.845	8.964
Left,Baseline C2, C4, C6	3.248	260	0.001	11.394	56.671	4.486	18.301

In the left motor cortex (C5, C3, C1), it was found that the average power during left motor imagery, at 7 Hz to 13 Hz, was not statistically significantly different (138.781 ± 190.557) compared to baseline EEG (141.341 ± 173.325), $t(260) = 0.787$, $p = 0.432$.

However, in the right motor cortex (C2, C4, C6), it was found that the average power during left motor imagery, at 7Hz to 13Hz, was statistically significantly lower (99.255 ± 125.869) compared to baseline EEG (110.649 ± 139.476), $t(260) = 3.248$, $p = 0.001$.

Table 6.6 shows the result of the repeated-measures t-test comparing the mean power of each participant for each individual electrode, assuming equal variances.

From Table 6.6, it was found that the average power of 2 electrodes during left motor imagery, at 7 Hz to 13 Hz, was statistically significantly different compared to baseline EEG. The average power of C5 was statistically significantly lower (121.539 ± 108.317) compared to baseline EEG (131.792 ± 112.373), $t(86) = 2.550$, $p = 0.013$. The average power of C2 was statistically significantly lower (116.347 ± 115.110) compared to baseline EEG (141.612 ± 155.685), $t(86) = 3.522$, $p < 0.001$. The other 4 electrodes (C3, C1, C4, or C6) were not statistically significantly different compared to baseline EEG (all $p > 0.05$).

However, following Bonferroni correction for multiple comparisons ($\alpha = 0.0083$),

Table 6.6: Repeated-measures t-test comparing left motor imagery with baseline EEG for each electrode

	t	df	2-tail p	Mean diff	SD diff	Lower	Upper
Left,Baseline C5	2.550	86	0.013	10.253	37.502	2.260	18.245
Left,Baseline C3	0.913	86	0.364	4.906	50.137	−5.779	15.592
Left,Baseline C1	−1.069	86	0.288	−7.480	65.265	−21.390	6.430
Left,Baseline C2	3.522	86	<0.001	25.264	66.909	11.004	39.525
Left,Baseline C4	1.856	86	0.067	0.816	4.101	−0.0582	1.690
Left,Baseline C6	1.080	86	0.283	8.101	69.979	−6.814	23.016

only electrode C2 remained statistically significant. At this electrode, average power during left motor imagery was significantly lower (116.347 ± 115.110) compared to baseline EEG (141.612 ± 155.685), $t(86) = 3.522$, $p < 0.001$. No statistically significant differences were observed at the remaining electrodes (C5, C3, C1, C2, C4, or C6; all corrected $p > 0.05$).

6.4.3 Statistical analysis comparing right hand motor imagery with baseline

Table 6.7, 6.8 and 6.9 show the descriptive statistics the average power of each participant, comparing left MI and baseline, after CSP and Morlet wavelet transform.

Table 6.7: Table for descriptive statistics for right motor imagery and baseline

Average power	Mean	Sample size	Std. deviation
Right	119.978	522	137.199
Baseline	125.984	522	143.568

Table 6.10 shows the result of the repeated-measures t-test comparing the

Table 6.8: Table for descriptive statistics for each side of the motor cortex during right motor imagery and baseline

Average power	Mean	Sample size	Std. deviation
Right C5,C3,C1	135.598	261	165.628
Right C2,C4,C6	104.356	261	98.979
Baseline C5,C3,C1	141.209	261	164.614
Baseline C2,C4,C6	110.760	261	117.217

Table 6.9: Table for descriptive statistics for each electrode during right motor imagery and baseline

Average power	Mean	Sample size	Std. deviation
Right C5	154.194	87	135.302
Right C3	134.402	87	115.870
Right C1	118.198	87	224.824
Right C2	103.428	87	96.541
Right C4	132.882	87	118.404
Right C6	76.760	87	68.440
Baseline C5	164.830	87	138.762
Baseline C3	137.127	87	118.518
Baseline C1	121.668	87	218.299
Baseline C2	115.899	87	133.033
Baseline C4	135.270	87	124.053
Baseline C6	81.110	87	83.238

mean power of each participant, assuming equal variances.

Table 6.10: Repeated-measures t-test comparing right motor imagery and baseline EEG

	t	df	2-tail p	Mean diff	SD diff	Lower	Upper
Right,Baseline	1.881	521	0.061	6.007	72.969	−0.267	12.281

It was found that the average power during right motor imagery, at 7 Hz to 13 Hz, from C1 to C6, was not statistically significantly different (119.978 ± 137.199) compared to baseline EEG (125.984 ± 143.568), $t(521) = 1.881$, $p = 0.061$.

Table 6.11 shows the result of the repeated-measures t-test comparing the mean power of each participant on each side of the motor cortex, assuming equal variances.

Table 6.11: Repeated-measures t-test comparing right motor imagery with baseline EEG for each side of the motor cortex

	t	df	2-tail p	Mean diff	SD diff	Lower	Upper
Right,Baseline C5, C3, C1	1.136	260	0.257	5.610	79.793	−4.115	15.336
Right,Baseline C2, C4, C6	1.577	260	0.116	6.403	65.591	−1.591	14.398

In the left motor cortex (C5, C3, C1), it was found that the average power during right motor imagery, at 8 Hz to 30 Hz, was not statistically significantly different (135.598 ± 165.628) compared to baseline EEG (141.209 ± 164.614), $t(260) = 1.136$, $p = 0.257$.

Similarly, in the right motor cortex (C2, C4, C6), it was found that the average power during right motor imagery, at 8 Hz to 30 Hz, was not statistically significantly different (104.356 ± 98.979) compared to baseline EEG (110.760 ± 117.217), $t(260) = 1.577$, $p = 0.116$.

Table 6.12 shows the result of the repeated-measures t-test comparing the mean power of each participant for each individual electrode, assuming equal variances.

Table 6.12: Repeated-measures t-test comparing right motor imagery with baseline EEG for each electrode

	t	df	2-tail p	Mean diff	SD diff	Lower	Upper
Right,Baseline C5	1.354	86	0.179	10.636	73.276	−4.981	26.253
Right,Baseline C3	0.382	86	0.703	2.725	66.551	−11.459	16.909
Right,Baseline C1	0.334	86	0.739	3.470	97.014	−17.206	24.147
Right,Baseline C2	1.414	86	0.161	12.471	82.248	−5.058	30.001
Right,Baseline C4	0.299	86	0.766	2.389	74.589	−13.509	18.286
Right,Baseline C6	1.629	86	0.107	4.350	24.913	−0.960	9.660

From Table 6.12, it was found that the average power of none of the electrodes during right motor imagery, at 7 Hz to 13 Hz, was statistically significantly different compared to baseline EEG. This pattern was consistent across all electrodes, suggesting an absence of robust electrode-level power changes associated with right motor imagery in this frequency band.

6.4.4 ANOVA comparing left MI and right MI

A 2x2 repeated measures ANOVA was performed with the factors Condition and Laterality. The factor Condition has levels baseline and motor imagery, while the factor Laterality has levels left and right. As there are only 2 levels for each of the 2 factors, no tests of sphericity or post-hoc comparisons are required.

Table 6.13 shows the result of the repeated-measures ANOVA comparing the mean power of each participant.

Table 6.13: Repeated-measures ANOVA comparing left and right motor imagery EEG

		df	df error	F	Sig.	n_p^2
Left,Right	Condition	1	521	7.826	0.005	0.015
	Laterality	1	521	0.015	0.902	0.000
	Cond × Lat	1	521	0.091	0.763	0.000

From Table 6.13, a significant main effect of Condition was observed (F(1, 521) = 7.826, p = 0.005, n_p^2= 0.015), indicating that EEG power differed between motor imagery and baseline conditions when averaged across hemispheres. The main effect of Laterality was not statistically significant (F(1, 521) = 0.015, p = 0.902), nor was the Condition × Laterality interaction (F(1, 521) = 0.091, p = 0.763), indicating that the magnitude of the motor imagery–related change did not differ between the left and right hemispheres.

Table 6.14 shows the result of the repeated-measures ANOVA comparing the mean power of each participant on each side of the motor cortex.

From Table 6.14, for the left motor cortex (C5, C3, C1), the main effect of

Table 6.14: Repeated-measures ANOVA comparing left and right motor imagery EEG

		df	df error	F	Sig.	n_p^2
Left,Right	Condition	1	260	1.478	0.225	0.006
C5,C3,C1	Laterality	1	260	0.196	0.658	0.001
	Cond × Lat	1	260	0.376	0.540	0.001
Left,Right	Condition	1	260	7.730	0.006	0.029
C2,C4,C6	Laterality	1	260	0.151	0.698	0.001
	Cond × Lat	1	260	1.500	0.222	0.006

Condition was not statistically significant (F(1, 260) = 1.478, p = 0.225, n_p^2 = 0.006). Similarly, no significant main effect of Laterality was observed (F(1, 260) = 0.196, p = 0.658, n_p^2 = 0.001) and the Condition × Laterality interaction was also not significant (F(1, 260) = 0.376, p = 0.540, n_p^2 = 0.001). These results indicate that motor imagery did not elicit differential changes in EEG power relative to baseline within the left motor cortex cluster.

For the right motor cortex (C2, C4, C6), a significant main effect of Condition was observed, F(1, 260) = 7.730, p = 0.006, n_p^2 = 0.029, indicating a reliable difference in EEG power between motor imagery and baseline conditions when averaged across hemispheres. However, the main effect of Laterality was not significant (F(1, 260) = 0.151, p = 0.698, n_p^2 = 0.001), nor was the Condition × Laterality interaction (F(1, 260) = 1.500, p = 0.222, n_p^2 = 0.006). Thus, although motor imagery produced significant power changes over the right motor cortex, these effects were not lateralised.

Table 6.15 shows the result of the repeated-measures ANOVA comparing the mean power of each participant for each electrode.

From Table 6.15, at C5, a significant main effect of Condition was observed (F(1, 86) = 4.218, p = 0.043, n_p^2 = 0.047), indicating greater EEG power during motor imagery relative to baseline when averaged across hemispheres. A strong main effect of Laterality was also present (F(1, 86) = 50.709, $p < 0.001$, n_p^2 = 0.371), reflecting a robust hemispheric difference in power at this electrode.

Table 6.15: Repeated-measures ANOVA comparing left and right motor imagery EEG

		df	df error	F	Sig.	n_p^2
Left,Right	Condition	1	86	4.218	0.043	0.047
C5	Laterality	1	86	50.709	<0.001	0.371
	Cond × Lat	1	86	0.003	0.958	0.000
Left,Right	Condition	1	86	0.669	0.416	0.008
C3	Laterality	1	86	14.583	<0.001	0.145
	Cond × Lat	1	86	0.066	0.798	0.001
Left,Right	Condition	1	86	0.075	0.785	0.001
C1	Laterality	1	86	3.129	0.080	0.035
	Cond × Lat	1	86	1.215	0.273	0.014
Left,Right	Condition	1	86	6.367	0.013	0.069
C2	Laterality	1	86	14.433	<0.001	0.144
	Cond × Lat	1	86	4.707	0.033	0.052
Left,Right	Condition	1	86	0.155	0.694	0.002
C4	Laterality	1	86	99.139	<0.001	0.535
	Cond × Lat	1	86	0.040	0.842	0.000
Left,Right	Condition	1	86	2.082	0.153	0.024
C6	Laterality	1	86	68.227	<0.001	0.442
	Cond × Lat	1	86	0.268	0.606	0.003

However, the Condition × Laterality interaction was not significant (F(1, 86) = 0.003, p = 0.958, n_p^2 = 0.000), indicating that the motor imagery–related change did not differ between hemispheres.

At C3, no significant main effect of Condition was found (F(1, 86) = 0.669, p = 0.416, n_p^2 = 0.008). In contrast, a significant main effect of Laterality was observed (F(1, 86) = 14.583, p ¡ 0.001, n_p^2 = 0.145). The Condition × Laterality interaction was not significant (F(1, 86) = 0.066, p = 0.798, n_p^2 = 0.001).

At C1, neither the main effect of Condition (F(1, 86) = 0.075, p = 0.785, n_p^2 = 0.001), nor the main effect of Laterality (F(1, 86) = 3.129, p = 0.080, n_p^2 = 0.035), reached statistical significance. The Condition × Laterality interaction was also non-significant (F(1, 86) = 1.215, p = 0.273, n_p^2 = 0.014).

At C2, a significant main effect of Condition was observed (F(1, 86) = 6.367, p = 0.013, n_p^2 = 0.069). A significant main effect of Laterality was also present (F(1, 86) = 14.433, p ¡ 0.001, n_p^2 = 0.144). Importantly, a significant Condition

× Laterality interaction was detected (F(1, 86) = 4.707, p = 0.033, $n_p^2 = 0.052$), indicating that the motor imagery–related change in EEG power differed between hemispheres at this electrode.

At C4, no significant main effect of Condition was found (F(1, 86) = 0.155, p = 0.694, $n_p^2 = 0.002$). However, a strong main effect of Laterality was observed (F(1, 86) = 99.139, $p < 0.001$, $n_p^2 = 0.535$). The Condition × Laterality interaction was not significant (F(1, 86) = 0.040, p = 0.842, $n_p^2 = 0.000$).

Finally, at C6, neither the main effect of Condition (F(1, 86) = 2.082, p = 0.153, $n_p^2 = 0.024$) nor the Condition × Laterality interaction (F(1, 86) = 0.268, p = 0.606, $n_p^2 = 0.003$), reached significance. A significant main effect of Laterality was observed (F(1, 86) = 68.227, $p < 0.001$, $n_p^2 = 0.442$).

Across electrodes, strong and consistent main effects of Laterality were observed, indicating systematic hemispheric differences in EEG power that were largely independent of motor imagery. In contrast, significant Condition × Laterality interactions were rare, emerging only at electrode C2, suggesting that lateralised motor imagery effects were spatially limited and not consistently expressed across the motor cortex. This pattern reinforces the conclusion that motor imagery modulates EEG power, but that robust hemispheric asymmetries are weak and electrode-specific at the scalp level.

6.5 Discussion

This chapter examined whether motor imagery (MI) elicits lateralised changes in EEG power over the motor cortex, with particular emphasis on differences between left (C5, C3 and C1) and right (C2, C4 and C6) hemispheric activity. Initial repeated-measures t-tests indicated a significant decrease in EEG power at 7 to 13 Hz, during left motor imagery relative to baseline in the right motor cortex, whereas no significant difference was observed in the left motor cortex. This observed reduction in EEG power reflects event-related desynchronisation of the sensorimotor mu rhythm. This decrease in oscillatory power is a well-established marker of motor system activation and is commonly observed during both executed and imagined movements. Higher power during baseline reflects greater neural synchrony and cortical idling, whereas the reduction in power during motor imagery indicates increased engagement of motor cortical networks. Thus, the lower power observed during left motor imagery represents functional activation rather than reduced neural activity.

In contrast, comparisons between right motor imagery and baseline did not reveal significant differences in either hemisphere. One possible explanation for the observed asymmetry in detectability is participant handedness. While the public dataset did not indicate the handedness of the participants, based on the percentage of left-handed people in the general population, it can be assumed that most participants were right-handed. Hence, imagery of the dominant right hand may rely on more efficient or automatised neural representations, leading to weaker measurable effects. Conversely, imagery of the left hand may have required greater neural recruitment, potentially resulting in larger or more detectable EEG power changes. However, this interpretation remains speculative, as the present analysis did not include handedness as an explicit factor, nor did it directly compare dominant versus non-dominant imagery effects within a single statistical model.

The observation of reduction in EEG power in the right motor cortex during left motor imagery also appears consistent with the contralateral organisation of the motor system, whereby imagery of one limb preferentially engages neural populations in the opposite hemisphere. However, the repeated-measures ANOVA provides a more appropriate framework for evaluating hemispheric asymmetry, as it formally tests whether the magnitude of the MI-related change differs between hemispheres. In this analysis, no significant Condition × Laterality interaction was observed for either motor cortex electrode cluster. This indicates that although motor imagery produced detectable changes in EEG power, these changes were not reliably larger in one hemisphere than the other. Importantly, the absence of an interaction demonstrates that differences in statistical significance observed in the t-tests cannot be interpreted as evidence of a true hemispheric difference in effect magnitude.

The main effects observed in the ANOVA further contextualise the t-test findings. Over the left motor cortex (C5, C3, C1), no significant main effects of Condition or Laterality were detected, indicating that motor imagery did not reliably modulate EEG power in this region. In contrast, over the right motor cortex (C2, C4, C6), a significant main effect of Condition was observed, reflecting greater EEG power during motor imagery compared to baseline when averaged across hemispheres. While this suggests that motor imagery effects were more readily detectable over the right motor cortex, the lack of a significant Laterality effect or interaction indicates that this activity cannot be attributed to hemispheric specialisation per se.

Overall, the findings suggest that while motor imagery reliably modulates EEG power relative to baseline, lateralisation effects at the scalp level are weak and variable. The small effect sizes associated with Laterality and Condition × Laterality terms further support this conclusion. These results underscore the importance of using interaction-based analyses when investigating hemispheric asymmetries, as reliance on separate within-hemisphere tests may lead to over-interpretation. Future work incorporating higher spatial resolution measures or more sensitive modelling approaches may be required to robustly capture contralateral motor imagery effects.

No strong or consistent spatial pattern emerged when comparing individual electrodes beyond those already discussed. While isolated electrodes such as C2 showed statistically significant differences in some conditions, these effects were not systematic across conditions or hemispheres. This suggests that motor imagery-related activity was distributed rather than focal, and that single-electrode analyses may lack sufficient sensitivity or robustness in this dataset. The absence of a clear electrode-level pattern is therefore not unexpected, particularly given inter-individual variability and the smoothing effects of spatial filtering and trial averaging.

A fundamental limitation of this study is the poor spatial resolution of EEG, which restricts precise localisation of neural generators. Although EEG offers excellent temporal resolution, its ability to resolve fine-grained cortical sources is limited by volume conduction and electrode density. For this reason, EEG is often paired with fMRI to combine temporal and spatial precision. The absence of fMRI data in this dataset prevents confirmation of whether the observed spectral changes originated strictly from primary motor cortex or from associated premotor and parietal regions. This limitation should be considered when interpreting hemispheric and electrode-level findings.

An important overall observation is that statistically significant differences were more readily detected when comparing motor imagery with baseline EEG than when comparing left versus right motor imagery directly. This suggests that while the pipeline effectively captures task-related motor imagery activity, discriminating between imagery classes remains challenging with power-based features alone. This finding aligns with existing literature indicating that motor imagery classification often benefits from combining spectral, spatial, and temporal features or from using advanced machine learning approaches. Future work could explore subject-specific models, alternative frequency bands such as mu and beta rhythms, or multimodal data to improve discrimination performance.

6.6 Conclusion

This chapter demonstrated that the proposed EEG processing pipeline is capable of capturing task-related neural activity associated with hand motor imagery, as evidenced by consistent differences between motor imagery and baseline EEG across sensorimotor regions and relevant frequency bands. Time–frequency analyses revealed physiologically plausible patterns, including reductions in 7–13 Hz power consistent with event-related desynchronisation, contralateral engagement during left motor imagery, and asymmetries plausibly related to participant handedness. Together, these observations support the validity of the analysis approach and its sensitivity to established markers of motor system engagement.

However, while reductions in sensorimotor power were most evident during left motor imagery, particularly over electrodes associated with the right motor cortex, interaction-based analyses did not provide evidence for hemispheric lateralisation of these effects. Instead, motor imagery–related changes were modest, distributed, and variable across participants and electrodes. Overall, the results suggest that although motor imagery reliably engages motor-related neural processes, its lateralised expression at the scalp EEG level may be subtle and influenced by methodological factors such as spatial resolution, inter-individual variability, and task design.

Chapter 7

Design of a 3D virtual and augmented reality platform

7.1 Introduction

In the previous chapter, we have explored technologies and techniques for neuroimaging. In this chapter, we will explore the technologies which generate the simulations for the brain to respond to, in particular, extended reality (XR). These XR technologies have opened new avenues of exploration, offering innovative ways to study brain function, cognition, and perception. These systems have the potential to push the boundaries of neuroscience research by creating controlled and customizable environments for experiments, enabling researchers to probe the complexities of human perception, behavior, and brain dynamics in novel ways.

XR is an umbrella term that encompasses all immersive technologies, including Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR). These technologies blend real and virtual environments to varying degrees, offering users a spectrum of experiences that range from augmenting the real world with digital overlays to fully immersing users in a virtual environment. AR is a system that allows users to interact with their real environment while simultaneously overlaying digital content onto it. This creates a composite view, where digital objects coexist with the physical world. VR, in contrast, immerses the user in a fully digital, computer-generated environment, blocking out their real surroundings entirely. [214].

In between AR and VR, we have MR, a technology that combines elements of both the physical and digital worlds, creating a hybrid environment where virtual and real objects coexist and can interact in real time. Unlike AR or VR, MR enables seamless integration between the two. In mixed reality, users can manipulate and interact with virtual objects as though they were part of their physical environment, often using advanced tools such as motion tracking, spatial mapping, and artificial intelligence. Table 7.1 summarises the characteristics of the different XR technologies.

Table 7.1: Summary of AR,VR,MR and XR

Technology	Stimuli	Real-world visibility
AR	Digital content overlayed on the real world	Visible
VR	Fully immersive virtual environment	Not visible
MR	Anchored digital content that can interact with the real world	Visible
XR	Umbrella term for AR, VR and MR	

From the table, both AR and MR technologies allow users to see their environment. This can be achieved with either a see-through lens or a front-facing camera. While VR technologies don't require the user to be able to see the real world to be immersed in the virtual environment, many of them also have AR and MR capabilities.

Many modern smartphones also have the capability for AR and VR. Modern smartphones are equipped with advanced sensors, processing power, and portable display. In VR, smartphones are often used in conjunction with headsets like Google Cardboard [215] or Samsung Gear VR [216], providing immersive 360-degree experiences by utilizing their gyroscopes, accelerometers, and high-resolution screens. For AR, smartphones leverage their cameras, GPS, and accelerometers to overlay digital content onto the physical world, enabling AR applications like Pokémon GO [217]. AR development platforms such as ARKit (iOS) and ARCore (Android) [218] enhance smartphone functionality, enabling accurate spatial mapping and real-time interaction with virtual objects. The affordability and widespread availability of smartphones granted more people access to VR and AR experiences.

Ko et. al. (2021) summarised the current technological advancements in functional materials and devices for XR technologies [219]. It highlights recent progress in developing haptic feedback systems, such as vibro-haptic and thermo-haptic technologies, which enhance user immersion by simulating realistic sensations of touch and temperature in XR environments. The paper also emphasizes the importance of wearable, skin-like sensors and actuators, which integrate seamlessly with human sensory systems to improve comfort and interactivity. Furthermore, it explores innovations in optical components for head-mounted displays, such as transparent and deformable displays, as well as soft robotics for applications like surgical training and teleoperation. While these advancements offer significant opportunities for immersive experiences and practical applications, the paper also raises ethical concerns, including mental health risks, privacy issues, and the potential for manipulation of users' beliefs and emotions.

7.1.1 XR in neuoscience

AR, VR and MR systems have already been widely adopted in industries such as gaming [220], retail [221], education [222], and therapy[223] for their ability to enhance user engagement and simulate real-world scenarios. In neuroscience research, these XR systems push the boundaries of the phenomena that can be studied. They overcome the limitations of traditional 2D screen-based setups by providing dynamic, immersive, and interactive environments that better mimic real-world or controlled scenarios. Hence, more naturalistic studies of brain processes such as spatial navigation, attention, and sensory integration, bridging the gap between artificial experimental setups and real-world behaviour can be designed using XR technologies. This increased ecological validity offers deeper insights into how the brain functions in dynamic, real-life contexts.

Kober, Kurzmann and Neuper designed and ran a spatial navigation experiment involving the use of VR [224]. They noted the concept of spatial presence, which was defined as “the sense of being in an environment“ [225], was connected to the use of VR applications due to their high level of immersion. They also noted that an increase in spatial presence correlates with an increase in the transfer of knowledge. In the study, participants were asked to walk through a virtual maze in both 2D and 3D virtual environments, and both EEG signals and subjective presence ratings were obtained. It was found that there were differences in brain activation between the 2D and 3D environment. The authors hypothesised that the differences were caused by difference in presence experience. This was based on an increased parietal brain activation, which was associated with an increased presence experience, in the 3D environment. While this study gave us a better understanding of our brain processes in the virtual environment, the 3D environment was presented on a large projection screen by a VR projector. Today, with the use of VR headsets, such as HTC Vive or Oculus Rift, an even greater level of immersion can be provided, which could further push the boundaries of what we know about our brain processes in 3D environments.

Krugliak and Clarke sought to understand our neural processes in real-world and AR environments [226]. In this study, the authors used the face inversion effect to study the cognitive effects which correlates to the different environments. The participants were asked to perform face inversioon tasks while placed in three different environments: AR, VR and 2D. At the same time, a mobile EEG was used to capture the participants’ brain signal. This study demonstrated the feasibility of combining AR and EEG in neuroscience studies. It was found that there were significant differences in the face inversion effect in when the participants were placed in different environments. The author raised an interesting question, whether there exist a difference in neural responses when virtual objects, as opposed to real objects are perceived. To build on this study, the use of a VR system might provide interesting insights into our neural processes.

In this chapter, we explore the role of AR and VR technologies in neuroscience research, with a particular focus on their contributions to the study of visual perception. We will discuss the technical principles of AR and VR, their unique advantages for neuroscience experiments, and the insights they have provided into how the brain interprets and interacts with complex sensory environments. By bridging the gap between real and virtual worlds, AR and VR systems are not only expanding the methodological toolkit of neuroscientists but also reshaping how we study and understand the brain.

7.1.2 How XR systems work

7.1.2.1 Stereoscope

If we look back at the history of XR systems, XR technology was born from stereoscopes back in 1838 by Sir Charles Wheatstone, where he demonstrated that a sense of immersion and depth can be given to an image when two photographs taken from different sides of the same object are combined by the brain [227]. The stereoscope operates by leveraging the concept of stereopsis, the brain's ability to perceive depth from binocular disparity. Stereopsis arises because each eye views the same scene from slightly different angles, resulting in images on the two retinas that differ in the position of objects. These positional differences, termed binocular disparities, provide crucial information about an object's depth [15]. Figure 7.1 explains how stereopsis works:

Suppose a person is looking at a bar placed on F. Light would be reflected off bar F, and enter the fovea in each eye. Each eye will hence capture an image where bar F is in a slightly different position. The right eye would capture bar F at N, and the left eye would capture bar F at P. Using the differences between the angular distance from the fovea (α_L and α_R), also known as binocular disparity, the brain is able to match the two images to determine the true position of F. While there is a false match formed by N and P, marked by the open circle labelled "false match“, the brain is able to ignore this and match the correctly

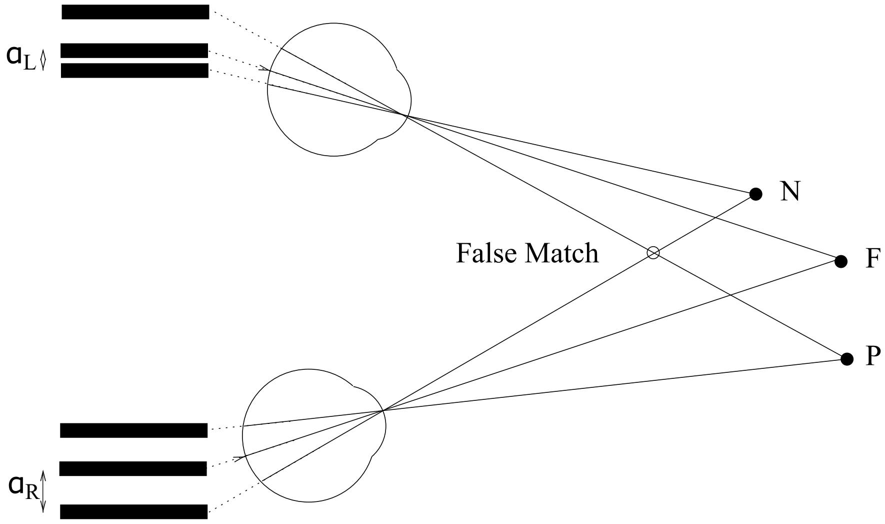


Figure 7.1: Figure of how light is captured by both eyes, resulting in two images that differ slightly in the position of the object [15].

corresponding features [15].

In a stereoscope, two slightly different images, mimicking the perspectives of the left and right eyes, are presented separately to each eye. The lenses on the device are placed strategically so that the images are aligned to correspond with how the eyes naturally perceive depth. The brain fuses these two images into a single 3D perception, resolving the "stereo correspondence problem", which involves matching features between the two retinal images. The stereoscope is a simple yet effective tool that exploits these fundamental principles of binocular vision to create a compelling illusion of depth, making it instrumental in early stereopsis research and 3D imaging studies. The complete history of how the stereoscope evolved to become the XR systems we see on the market today is out of scope of this thesis. Interested reader may turn to chapter 1.2 of [227].

7.1.2.2 Virtual Reality

VR systems are built based on the stereoscope. Just like how the stereoscope present two slightly different image separately to each eye, a VR headset has two tiny stereoscopic screens which were positioned a few inches right in front of both eyes. VR systems can range from non-immersive to fully immersive. Non-immersive VR systems allow users to interact with virtual environments through standard interfaces like monitors and keyboards. Semi-immersive VR systems enhance engagement by using larger displays or projection systems, such as the Audio-Visual Experience Automatic Virtual Environment (CAVE) system [170] or other large screen setups, creating a more enveloping experience without full immersion. Fully immersive VR systems provide a complete virtual experience by surrounding users with the environment through head-mounted displays (HMDs) and motion tracking, such as the Oculus Rift and HTC Vive, which enable users to move and interact within a 3D space [228]. Figure 7.2 shows how different parts work together within a VR system to make it work.

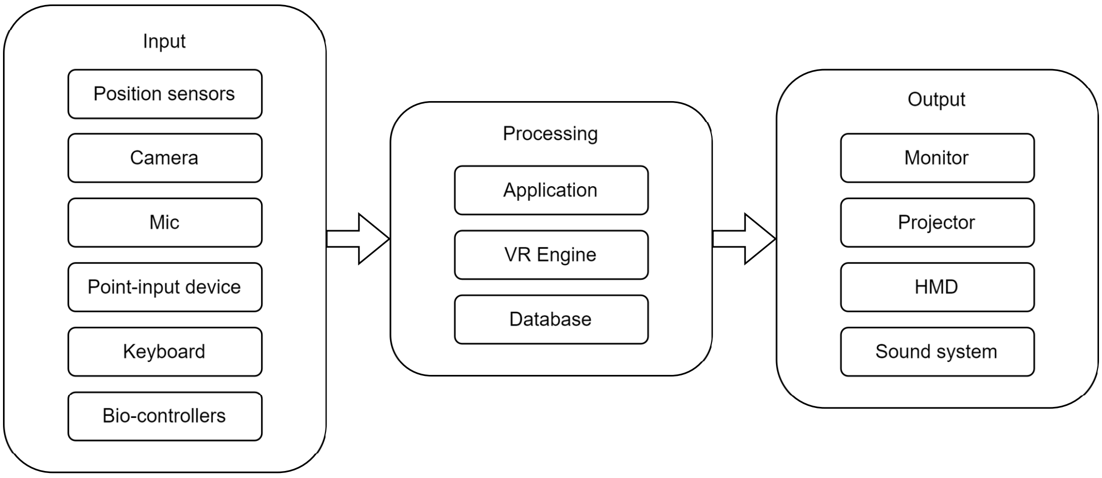


Figure 7.2: Language processing pathway and dyslexia compensatory mechanism

In a VR system, it first needs an input, which can be received from various sensors which collect information such as the user's movement, environment layout and sound. This information is then sent to the VR application, which compares with the existing database to make a decision. The decision is then sent to the VR Engine to generate an output, which is sent to a module which is able to project the output, such as a monitor or a HMD. At the same time, the output sound is played through the sound system.

Virtual Reality (VR) environments are rendered by a VR engine, such as Unity or Unreal engine [229]. A VR engine integrates various components such as 3D models, 2D graphics, audio, and scripting to compile and render a VR application. Figure 7.3 shows a summary of the software used to build a VR application.

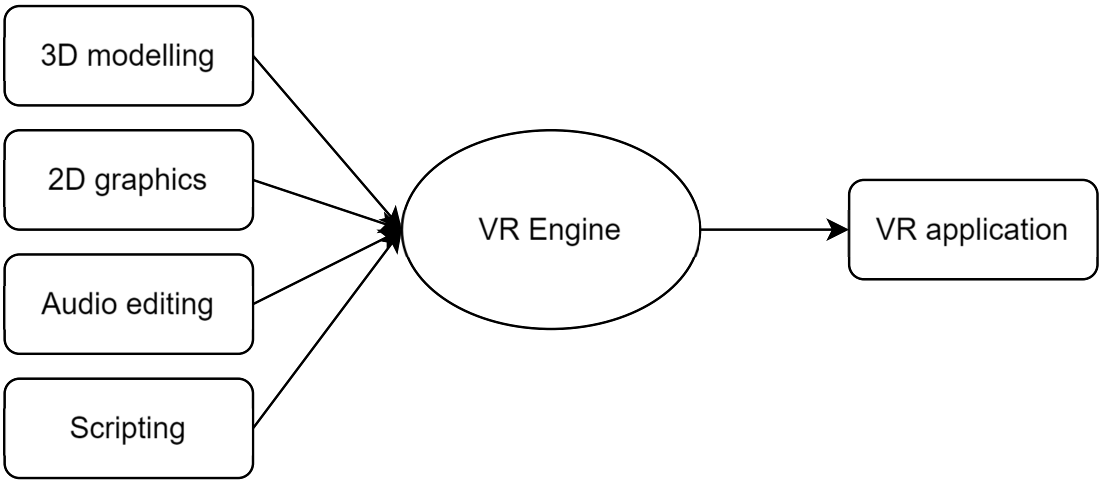


Figure 7.3: Language processing pathway and dyslexia compensatory mechanism

Firstly, 3D Models are the visual representations of objects within the virtual environment, allowing users to interact with and explore the space. Secondly, 2D Graphics can include user interface elements, menus, and overlays that enhance the user experience. Thirdly, sound effects and background music are integrated to create an immersive atmosphere, contributing to the realism of the VR experience. Lastly, scripting involves programming logic and behaviours that dictate how users interact with the environment, enabling dynamic responses and interactivity. The VR engine brings all these elements together to compile and render the VR application for users to explore and interact with [227].

7.1.2.3 Augmented Reality

AR system uses similar technology as compared to VR systems, in that similar hardware and software are used to receive input, generate graphics and project display for both AR and VR systems. In fact, many of the well-known VR HMD have AR capablilities as well, all they needed was a camera on the front of the headset to capture the environment. Chapter 4 of [227] goes into in-depth detail about the mechanism behind AR technology, which will be summarised in this subsection.

AR elements can be rendered by two different types of devices: holographic devices and immersive devices [230]. Holographic devices have a see-through lens and projects the digital stimuli on the see-through lens. This technology typically employs holographic displays, which manipulate light to project images that can be viewed from different angles without the need for special glasses. Holographic AR systems often integrate sensors and cameras to track the user's position and movements, enabling the virtual content to adjust dynamically to the user's perspective [227]. Examples of holographic devices include Magicleap and Microsoft Hololens [231].

On the other hand, immersive devices have an opaque display and a camera on the front of the headset. The opaque display blocks the user's view of the environment, allowing for a fully immersive experience similar to VR. The front camera captures the user's surroundings, enabling the system to recognize and track real-world objects and surfaces. This information is processed in real-time to accurately position and render virtual elements within the user's field of view. The display will then show the output from the camera, with the virtual elements overlayed on top [227]. Most VR headsets with a front camera, such as HTC vive and Oculus rift, have such a capability.

One notable system used by AR systems is the marker-based system. In a marker-based AR system, the process begins with a camera capturing the real-world environment, where specific visual markers, such as QR codes or images, are identified by the software. Once a marker is detected, the AR system uses algorithms to determine its position and orientation, allowing it to accurately overlay virtual objects onto the marker in real-time. This creates an interactive experience where users can see and engage with digital content seamlessly integrated into their physical surroundings [227].

7.1.2.4 Mixed Reality

MR technologies are any technology that exist somewhere between the real world and VR in the virtuality continuum. Proposed by [16], the virtuality continuum describes a mixture of classes of objects presented in any particular display situation. Figure 7.4 illustrates the virtuality continuum:

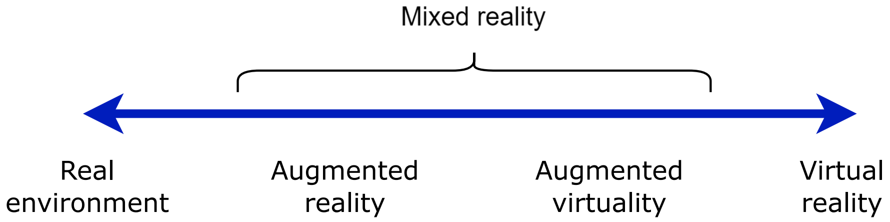


Figure 7.4: Simplified illustration of the virtuality continuum. Adapted from [16]
.

The primary aim of MR technology, as detailed by [232], is to seamlessly blend real and virtual environments to create immersive, interactive experiences that enhance human interaction with digital content. MR seeks to bridge the gap between AR and VR by allowing users to interact with both physical and virtual objects in real-time while maintaining spatial awareness. The hardware and software used by MR systems are similar to those used by VR and AR systems. However, more sensors would be required to allow the user to seamlessly interact with virtual objects. Due to the complexity of MR systems, the details on the mechanism behind MR technology is out of the scope of this thesis. Interested readers may turn to [232] for a detailed review of the current state-of-the-art MR technology.

7.2 System setup

A simple platform that can be used to build VR simulations for neuroscience experiments is proposed. Here the experiment in chapter 8 would be an example usage of this plaform.

For an EEG experiment which involves the use of VR and AR equipment, four crucial design requirements that need to be fulfilled were identified:

- Hardware with sufficient processing power to run both AR/VR systems and EEG system without compromising processing speed, which refers to how quickly a computer's Central Processing Unit (CPU) can execute instructions and process data [233].

- Software which is capable of connecting the AR/VR system with the EEG system so that both systems can communicate with each other.
- Able to run reaction-time simulations with low latency, where latency is the time used to process a user input, or the time between client request and server response.
- Input hardware is constant.

With these design requirements in mind, two separate computer systems were set up, one which is connected to the EEG system and records the brain signal, called the signal recording system, and the other which is connected to the headset or display which displays the stimuli to the participant, called the stimuli display system. The purpose of using two separate computer systems was to ensure that there was an acceptable level of processing speed. The two systems are then connected via ethernet TCP/IP protocol, which minimises the speed of data transfer.

For this experiment participants were placed in four different environments: 2D screen, real world, virtual reality (VR) and augmented reality (AR). The hardware and software used for each environment is listed in Table 7.2

Table 7.2: Table for hardware and software used for each environment

Environment	Hardware	Software
2D	TV screen	Psychopy
Real world	Arduino, RGB LED	Psychopy
AR	Magicleap	Unity
VR	HTC Vive	Unity

Subsection 7.2.1 will explain the hardware setup for the experiment, and Subsection 7.2.2 will explain the software used for the simulation and EEG processing.

7.2.1 Hardware

Figures 7.5 to 7.8 shows the hardware set up of each of the system used in the various environments.

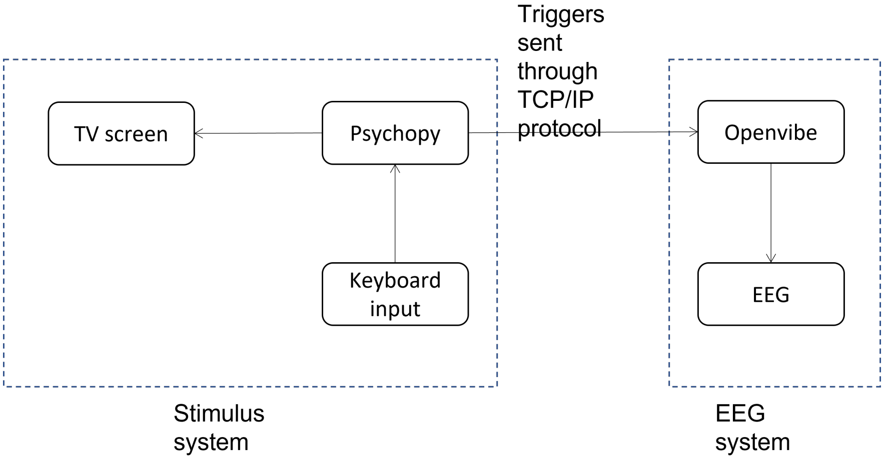


Figure 7.5: Flowchart of 2D environment setup

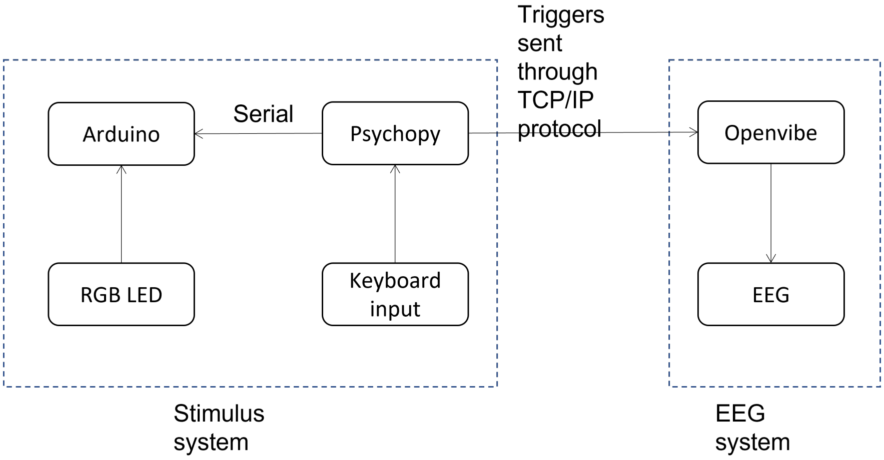


Figure 7.6: Flowchart of real world environment setup

The stimuli of both the 2D and the real world environment setup are controlled by Psychopy, a python open source software package, including both a Python library and a graphical interface, primarily to build simulations in neuroscience and experimental psychology research [234]. In the 2D environment setup, Psychopy directly shows the stimuli on a TV screen, in particular, the Samsung PS50C450B1W 50”, 720p plasma display TV. However, in the real world environment setup, Psychopy will send a serial signal to Arduino, which would then

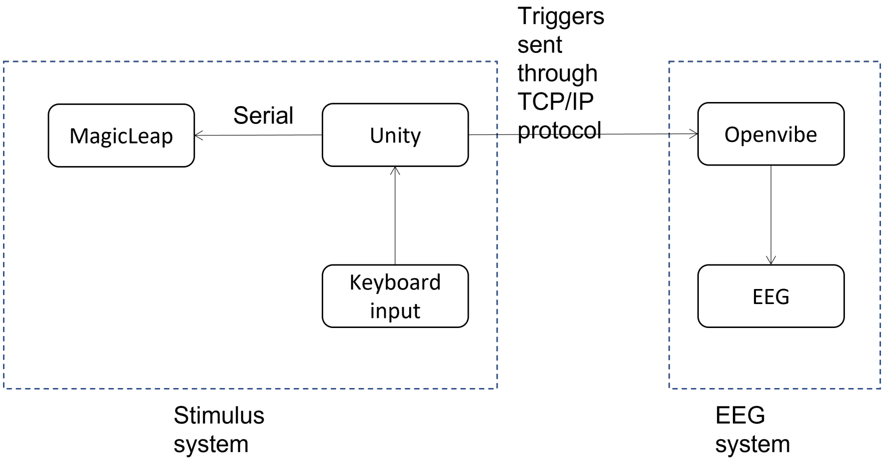


Figure 7.7: Flowchart of AR environment setup

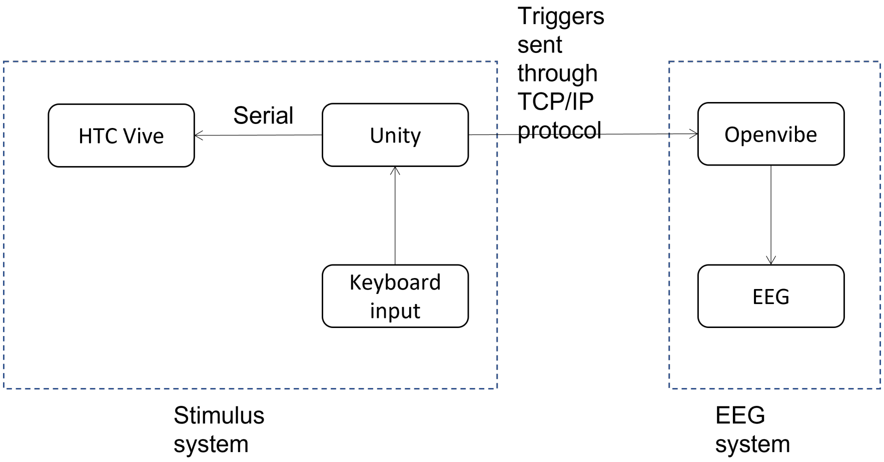


Figure 7.8: Flowchart of VR environment setup

turn on or off an RGB LED attached to the back of a cube made out of frosted acrylic sheets.

The stimuli of both AR and VR environment setup are controlled by Unity, which is capable of communicating with both MagicLeap and HTC Vive via serial connection.

To ensure that confounding factors are kept to a minimum, the input method were kept constant for every environment. To achieve this, the software for each environment were programmed to take in inputs from a keyboard. Participants would need to press the spacebar to respond to the stimuli.

At the start and end of the experiment, the start and end of each stimuli and during each key press, a trigger will be sent by either Psychopy or Unity, depending on the environment, from the stimulus system to the EEG system. The EEG system will receive and record the trigger timings on Openvibe, which will also be receiving the signal from the EEG.

7.2.2 Software

PsychoPy and Unity were use to build the simulation. On both platforms, three main experimental phases were introduced:

- Instruction phase, where participants were instructed to press space as soon as the stimulus appears. Participants can start the experiment by pressing the spacebar.
- Preparation (fixation) phase, where a fixation point is presented to the participant and the participant were not required to do anything.
- Stimulus-response trial phase, where the participant were to respond as quickly as they can to the stimulus by pressing the spacebar.

In both cases, trials are looped and randomized to prevent practice effect [235]. The reaction times are saved for later analysis, and a trigger is sent to the EEG system at the same time.

Both the script for Psychopy and Unity are designed to be modular and adaptable to the needs of other neuroscience and psychology researchers. Researchers can modify the number of trials, stimulus properties (e.g., shape, size, colour, position), timing parameters (e.g., fixation jitter, stimulus duration), and response mappings directly through configuration variables or external condition files without altering the core structure of the experiment. The modular routine design (instruction, fixation, stimulus–response) allows additional experimental phases or conditions to be inserted with minimal effort. Crucially, the use of TCP/IP–based event markers enables straightforward integration with common neurophysiological recording systems (e.g., EEG, MEG, fNIRS), making the paradigm suitable for multimodal studies that require precise temporal synchronisation. As a result, the script provides a generalizable template for reaction-time and stimulus–response paradigms that can be tailored to diverse research questions while maintaining reproducibility and timing precision.

Building on this general adaptability, Unity further extends the usefulness of the paradigm for neuroscience and psychology researchers by enabling seamless integration with augmented and virtual reality hardware, such as HTC Vive and Magic Leap. Whereas the core task structure can be readily modified to address different experimental questions, Unity's native support for AR/VR devices allows these adaptations to be implemented in fully three-dimensional, immersive environments. Researchers can directly control stimulus position, depth, size, and motion relative to the participant's head and body, and can incorporate naturalistic interactions via tracked controllers or hand movements. This makes it possible to translate conventional reaction time and stimulus–response paradigms into ecologically valid spatial contexts, facilitating investigations of perception, attention, sensorimotor integration, and embodied cognition while preserving precise timing, reproducibility, and synchronization with neurophysiological recordings.

7.3 Hardware testing

With the use of multiple hardware and software platforms, there are many confounding factors which may affect the reaction time results, including luminescence of stimuli, latency between stimuli response and response time recorded, latency between the TCP communication between the stimuli system and the EEG system, noise from the XR headsets and background electromagnetic noise. Hence, to ensure the accuracy of the reaction time measured, it is paramount to measure each of these confounding factors and adjust the reaction time measured accordingly.

Empirical research has consistently demonstrated the existence of an inverse relationship between stimulus luminance and human reaction time [236, 237, 238]. This effect has been observed across a wide range of experimental paradigms, including simple and choice reaction time tasks, and under varying levels of background illumination and contrast. The relationship is often described as non-linear, with reaction time decreasing rapidly at low luminance levels and approaching an asymptote as luminance increases, suggesting diminishing returns at higher intensities. These findings indicate that luminance primarily influences early stages of sensory processing by enhancing signal strength and perceptual salience, thereby facilitating faster stimulus detection and response initiation. Therefore, a luminescence analysis is conducted so that the final reaction time recorded can be adjusted accordingly.

Many studies have previously examined the latency introduced by experimental equipment and software in reaction time paradigms, demonstrating that system-induced delays can meaningfully bias measured response times if left unaccounted for [239, 240]. Such latencies may arise at multiple stages of the stimulus–response pipeline, including stimulus rendering, signal transmission, input device polling, and data logging. Importantly, the magnitude and variability of these delays are highly dependent on several factors, including processor architecture, operating system scheduling, graphics pipelines, communication protocols, and the computational complexity of the virtual models and scenes being rendered. Consequently, latency values reported in prior studies may not generalize to the specific combination of hardware, software, and experimental configurations employed in the present study. Therefore, dedicated latency measurements were conducted using the exact hardware processors, software environment, and stimulus models used in the experiment. These empirically derived latency values were then used to correct the recorded reaction times.

Chapter 5 outlined how various sources of noise, such as physiological artifacts and environmental electrical interference, can contaminate EEG recordings and how these can be addressed in post-processing. However, noise reduction can also be implemented at the hardware level, which can substantially improve the signal quality prior to any digital artifact correction. Because EEG records very low-amplitude electrical activity (on the order of microvolts), it is inherently sensitive to electromagnetic interference (EMI) from nearby electrical equipment and ambient electrical fields, including mains line noise at 50/60 Hz and higher-frequency emissions from consumer electronics. Shielding solutions such as Faraday cages or conductive enclosures are widely used in clinical and research settings to attenuate such interference by surrounding the recording setup with a grounded conductive barrier that blocks external electrical fields from reaching the electrodes and amplifiers. For example, fully shielded EEG rooms and enclosures based on Faraday technology are standard in many laboratories to create an electrically quiet recording environment.

In the context of combined EEG and extended reality (XR) experiments, hardware shielding becomes particularly important. Recent structured testing has shown that head-mounted displays (HMDs), including commercially available virtual reality devices, can introduce consistent electromagnetic artifacts into EEG recordings, especially at their refresh-rate frequencies and mains harmonics, although much of the primary EEG frequency range (¡50 Hz) may remain relatively unaffected. Because participants must wear an XR headset directly over the EEG cap, the close proximity of the device's electronics, displays, sensors and power systems can introduce additional noise sources that are difficult to fully eliminate in post-processing alone. Hardware measures such as shielding of electrode leads, the use of conductive materials around sensitive components, and even purpose-built enclosures have been shown to help reduce the coupling of external electromagnetic fields into EEG sensors. Therefore, in this chapter, an aluminium foil cap, placed on top of the EEG cap, is tested for it's effectiveness as a Faraday's cage.

This section will discuss the methods and results of the various hardware tests conducted in preparation for the visual perception experiment, namely luminescence, latency and noise.

7.3.1 Methods

7.3.1.1 Setup

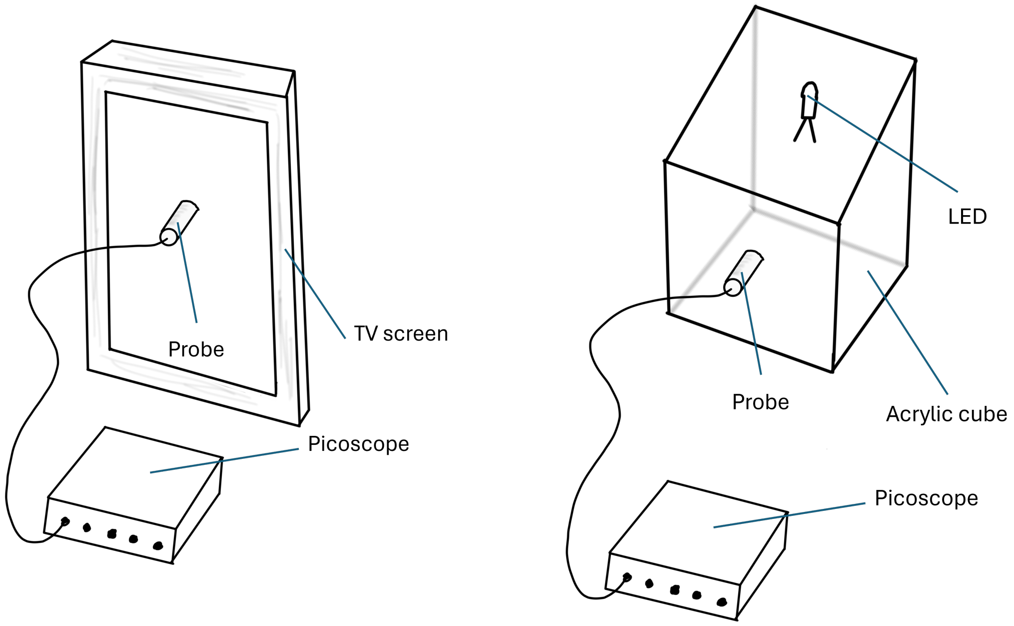
Luminescence and stimulus latency (the latency between the time the stimulus was actually presented and the recorded stimulus time). Both factors were measured with the use of a photosensor, which is attached to an oscilloscope. The oscilloscope, PicoScope, was used in this experiment. The photosensor, which is fixed on the display, detects the change in luminosity when the stimuli appear or disappear, then the PicoScope registers this change in luminosity as a change in voltage. The signal was recorded at a frequency of 30 μHz. Figure 7.9 shows how this experiment was set up for each environment.

For this luminescence and latency test, the simulation was modified so that the stimuli appear every 4 second for 2 second. After the simulation was completed, the timing of voltage changes recorded on the PicoScope was compared with the timing of the stimuli appearance and disappearance, as captured by the computer.

For the TCP latency test, a simple server and client python script which recorded the time a client request was sent, the time the server received the request, and the time the client received the server response. The client script was run from the stimuli system, while the server script was run from the EEG system.

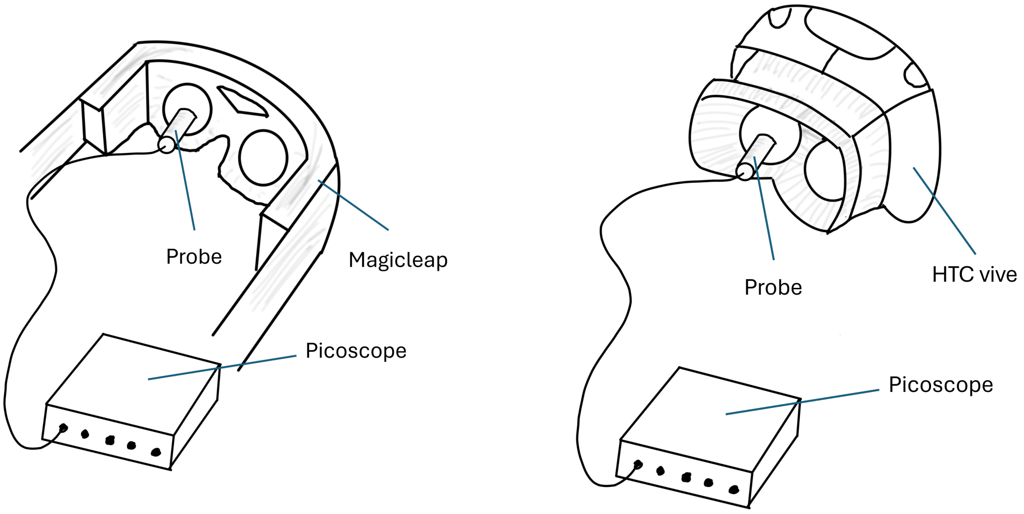
For the noise analysis, an oscilloscope was used to detect and record the electromagnetic wave emitted by a mobile phone placed near it. An aluminium foil cap was then placed over the oscilloscope probe, and electromagnectic signal was recorded again. This was repeated with the aluminium foil connected to the ground. For the control condition, the experiment was repeated with the mobile phone turned off.

To test if the findings from this experiment can be translated to EEG signals, the same aluminium foil cap was placed on a participant wearing an EEG cap, over the EEG cap, and background electromagnectic signal was recorded by the EEG



(a) 2D environment

(b) Real world environment



(c) AR environment

(d) VR environment

Figure 7.9: Hardware test setup

electrodes. The participant was not instructed to do anything. The experiment was then repeated with the foil cap grounded, and then again with the foil cap removed.

Next, to measure the noise level of XR systems, the XR systems used for the visual perception experiment, namely Magicleap One and HTC Vive, was placed over the EEG cap, without the aluminium foil cap. For completion, other XR systems were also included, namely Mobile VR, Meta Quest and Oculus Rift. The experiment was repeated with the XR systems turned off.

7.3.1.2 Luminence analysis script

Each PicoScope recording was stored as a single comma-separated values (CSV) file containing three concatenated vectors: photodiode voltage, sample time in nanoseconds, and Unix time. The voltage and time vectors were extracted and reshaped into one-dimensional arrays. Sample timestamps were converted from nanoseconds to seconds to facilitate alignment with stimulus timing information.

Stimulus presentation times were provided as an array of onset–offset pairs corresponding to individual stimuli. A binary mask was constructed to identify samples occurring during any stimulus interval. All samples not belonging to a stimulus interval were classified as baseline (non-stimulus) samples. This ensured that baseline estimates were derived exclusively from periods in which the display was not actively presenting a stimulus.

Baseline photodiode samples were inspected for transient artifacts arising from electrical noise or acquisition irregularities. To obtain a robust estimate of the baseline luminance level, outliers were removed using a median absolute deviation (MAD) criterion. Samples deviating by more than five MADs from the baseline median were excluded. The baseline luminance level was then estimated as the median of the remaining baseline samples, providing robustness against skewed noise distributions.

The estimated baseline level was subtracted from the entire photodiode signal to isolate stimulus-evoked luminance changes. Following baseline subtraction, small negative values attributable to measurement noise were clipped to zero. This step ensured that subsequent brightness estimates reflected stimulus luminance rather than noise around the display's black level.

For each stimulus interval, the baseline-corrected photodiode signal was segmented using the corresponding onset and offset times. Stimulus brightness was quantified as the mean photodiode voltage within each segment. This metric reflects the time-averaged luminance of the stimulus and is independent of stimulus duration. Root-mean-square (RMS) voltage was calculated and operationalised as the luminance value.

7.3.1.3 Latency analysis script

To quantify stimulus–response timing and system latency, a custom signal processing pipeline was implemented in MATLAB. The analysis comprised two main stages: (i) preprocessing and temporal alignment of recorded signals with system timestamps, and (ii) event detection and latency estimation based on signal energy changes.

Raw experimental data were organised into three streams: the recorded signal of interest, a high-resolution time vector converted from nanoseconds to seconds, and an additional system timestamp used to determine the total recording duration. In parallel, system-level timing information corresponding to stimulus presentation and response events was loaded from experiment log files. Two timing formats were supported to accommodate different experimental configurations. All stimulus timestamps were converted to seconds and temporally realigned so that time zero corresponded to the start of the experiment. This ensured a common time reference between the recorded signal and the stimulus markers.

To facilitate robust event detection, the recorded signal was transformed into an RMS representation. Prior to this step, the signal was visually inspected and manually offset by a constant value to correct baseline shifts and ensure consistent polarity across trials. The signal was then segmented into consecutive, non-overlapping windows of fixed length (100 samples), and the RMS value of each segment was computed. This procedure produced a smoothed amplitude envelope that reduced sensitivity to high-frequency noise while preserving transient changes associated with stimulus events. Each RMS value was associated with the temporal midpoint of its corresponding window.

The RMS envelope was further refined by restricting analysis to a user-defined temporal region containing the stimuli of interest. Within this region, event detection was performed using a manually selected amplitude threshold. RMS values crossing this threshold were identified as candidate signal responses. An interactive procedure allowed the threshold to be iteratively adjusted until the detected events accurately corresponded to visible signal changes. This manual verification step ensured that thresholding was robust to noise and inter-trial variability.

Detected threshold crossings were grouped into discrete response events based on their temporal separation. A minimum inter-event interval was enforced to distinguish independent stimulus responses from continuous or overlapping signal activity. Each group of threshold crossings was classified as the onset or offset of a response. This resulted in a sequence of detected signal events aligned with each stimulus presentation.

Finally, detected signal response times were compared with the known stimulus start and end times obtained from the system logs. For each stimulus, the temporal difference between the stimulus marker and the corresponding detected signal event was computed. These differences provided estimates of system latency and response timing accuracy. The resulting latency measures were returned for further statistical analysis and visual inspection. Additionally, the signal, stimulus markers, and detected events were plotted together to allow qualitative validation of the alignment between stimuli and signal responses.

7.3.1.4 Noise analysis script

Continuous EEG data were imported into MATLAB using the FieldTrip toolbox (Oostenveld et al., 2011). To ensure approximate stationarity of noise characteristics, the continuous EEG signal was segmented into non-overlapping 1-second windows. This segmentation length provided sufficient frequency resolution (1 Hz) for characterizing line noise while minimizing the impact of non-stationary artifacts such as electrode drift or muscle activity.

Power spectral density (PSD) estimates were computed separately for each 1-second window using a fast Fourier transform (FFT). A boxcar taper was applied to obtain unbiased power estimates. Spectral power was calculated over the frequency range 1–80 Hz, encompassing both line noise and high-frequency muscle artifacts. Trial-wise spectral estimates were preserved to allow quantification of noise variability.

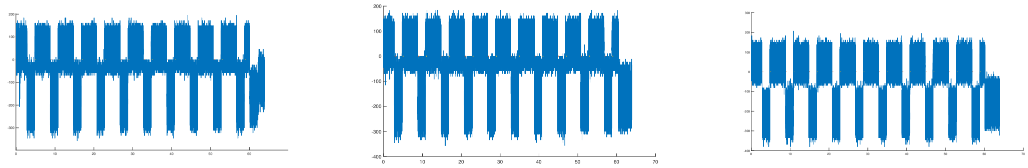
For each channel, mean PSD was computed across all segmented windows, providing an estimate of the overall noise floor. Noise was operationalized as the sum of the mean PSD across all frequencies for each channel. Channel 1 was removed for all conditions due to it being abnormally noisy as the reference channel.

7.3.2 Results

7.3.2.1 Picoscope trace

Figure 7.10 shows the trace output from the Picoscope for each environment and each colour, with the dots indicating the start and end time recorded by the stimulus program.

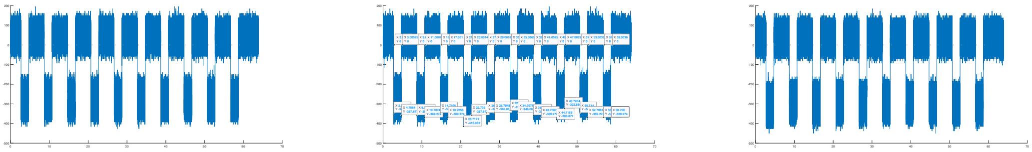
From Figure 7.10, visual inspection of the PicoScope traces indicated considerable variability in the magnitude of the voltage change across both environment and colour, reflecting differences in luminance across all conditions. Despite this variability in signal amplitude, the temporal onset of the signal was highly con-



(a) 2D, Green

(b) 2D, Red

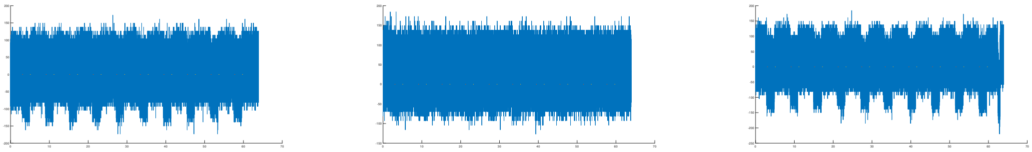
(c) 2D, Yellow



(d) Real world, Green

(e) Real world, Red

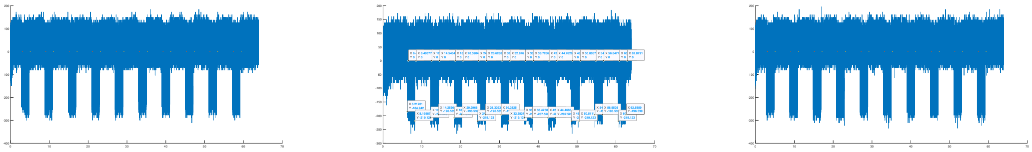
(f) Real world, Yellow



(g) AR, Green

(h) AR, Red

(i) AR, Yellow



(j) VR, Green

(k) VR, Red

(l) VR, Yellow

Figure 7.10: Trace output from Picoscope

sistent within each condition, with minimal variation in latency relative to the trigger.

7.3.2.2 Luminance testing

Table 7.3 shows the descriptive statistics for the luminance for each environment, while Table 7.4 shows the descriptive statistics for the luminance of each colour.

Table 7.3: Table for descriptive statistics for each environment

Environment	Mean	Sample size	Std. deviation
2D	220.732	30	26.250
Real world	325.072	30	8.490
AR	41.071	30	8.252
VR	102.191	30	13.448
Total	172.267	120	110.914

Table 7.4: Table for descriptive statistics for each colour

Colour	Mean	Sample size	Std. deviation
Green	170.000	40	109.362
Red	157.988	40	110.894
Yellow	188.812	40	113.067
Total	172.267	120	110.914

A two-way independent ANOVA was performed to analyze the effect of both the environment (2D, real world, AR or VR) and the colour of stimuli (red, green or yellow) on luminance. Levene's test cannot confirm that the assumption of homogeneity of variance had been met ($F(11,108) = 9.473$, $p < 0.001$). As the assumption of homogeneity of variance cannot be confirmed, a non-parametric test, Kruskal-Wallis H test was conducted. Table 7.5 and Table 7.6 show the results of the Kruskal-Wallis H.

From Table 7.5, the test revealed statistically significant differences across the four levels of Environment, $\chi^2(3) = 111.570$, $p < 0.001$. Median values were 205.238 for 2D ($n = 30$), 39.099 for AR ($n = 30$), 326.989 for RW ($n = 30$),

Table 7.5: Table for results of Kruskal-Wallis H. test comparing the effect of environment on luminance

Factor	χ^2	df	p	Median 2D	Median AR	Median RW	Median VR
Environment	111.570	3	<0.001	205.238	39.099	326.989	108.911

Table 7.6: Table for results of Kruskal-Wallis H. test comparing the effect of environment on luminence

Factor	χ^2	df	p	Median Green	Median Red	Median Yellow
Colour	6.290	2	0.043	155.852	141.420	184.696

and 108.911 for VR (n = 30). Post hoc pairwise comparisons with Bonferroni correction indicated that all pairs, AR-VR ($\chi^2(3)$ = -30.00, p <0.001), AR-2D ($\chi^2(3)$ = 60, p <0.001), AR-RW ($\chi^2(3)$ = -90.00, p <0.001), VR-2D ($\chi^2(3)$ = 30.00, p <0.001), VR-RW ($\chi^2(3)$ = 60.00, p <0.001), 2D-RW ($\chi^2(3)$ = -30.00, p <0.001), were significant.

Similarly, from Table 7.6, the test showed statistically significant differences across the three levels of Colour, $\chi^2(2)$ = 6.290, p = 0.043. Median values were 155.852 for green (n = 40), 141.420 for red (n = 40), and 184.696 for yellow (n = 40). Post hoc pairwise comparisons with Bonferroni correction indicated that all pairs, AR-VR ($\chi^2(3)$ = -30.00, p <0.001), AR-2D ($\chi^2(3)$ = 60, p <0.001), AR-RW ($\chi^2(3)$ = -90.00, p <0.001), VR-2D ($\chi^2(3)$ = 30.00, p <0.001), VR-RW ($\chi^2(3)$ = 60.00, p <0.001), 2D-RW ($\chi^2(3)$ = -30.00, p <0.001), were significant.

These results indicate that both environment and colour elicited differential changes in luminence.

7.3.2.3 Latency testing

Table 7.7 shows the descriptive statistics for the latency time for each environment, while Table 7.8 shows the descriptive statistics for the latency time for each colour.

Table 7.7: Table for descriptive statistics for each environment

Environment	Mean	Sample size	Std. deviation
2D	0.00220	30	0.0456
Real world	-0.00114	30	0.0715
AR	-0.00291	30	0.0369
VR	-0.000513	30	0.00460
Total	-0.000593	120	0.0293

Table 7.8: Table for descriptive statistics for each colour

Colour	Mean	Sample size	Std. deviation
Green	-0.00275	40	0.0103
Red	0.00138	40	0.0479
Yellow	-0.000407	40	0.0148
Total	-0.000593	120	0.0293

A two-way independent ANOVA was performed to analyze the effect of both the environment (2D, real world, AR or VR) and the colour of stimuli (red, green or yellow) on latency time. Levene's test cannot confirm that the assumption of homogeneity of variance had been met (F(11,108) = 5.812, $p < 0.001$). As the assumption of homogeneity of variance cannot be confirmed, a non-parametric test, Kruskal-Wallis H test was conducted. Table 7.9 and Table 7.10 show the results of the Kruskal-Wallis H.

Table 7.9: Table for results of Kruskal-Wallis H. test comparing the effect of environment on latency

Factor	χ^2	df	p	Median 2D	Median AR	Median RW	Median VR
Environment	1.021	3	0.796	0.000	0.000	0.000	0.000

From Table 7.9, the test revealed no statistically significant difference across the four levels of Environment, $\chi^2(3) = 1.021$, $p = 0.796$. Median values were 0.000 for 2D ($n = 30$), 0.000 for AR ($n = 30$), 0.000 for RW ($n = 30$), and 0.000 for VR ($n = 30$).

Table 7.10: Table for results of Kruskal-Wallis H. test comparing the effect of environment on latency

Factor	χ^2	df	p	Median Green	Median Red	Median Yellow
Colour	0.149	2	0.928	0.000	0.000	0.000

Similarly, from Table 7.10, the test showed no statistically significant difference across the three levels of Colour, $\chi^2(2) = 0.149$, p = 0.928. Median values were 0.000 for green (n = 40), 0.000 for red (n = 40), and 0.000 for yellow (n = 40).

These results indicate that neither environment nor colour elicited differential changes in latency time.

7.3.2.4 TCP latency

Table 7.11 shows the descriptive statistics of the round-trip time (RTT) and clock-offset timing between the EEG PC and the stimulus PC.

Table 7.11: Table for descriptive statistics of the round-trip time (RTT) and clock-offset timing between the EEG PC and the stimulus PC

	Mean	Median	SD	Min	Max	N
RTT(s)	4.860E-3	5.015E-3	1.409E-3	1.010E-3	7.000E-3	20
Clock-offset(s)	1.774	1.774	5.36E-4	1.773	1.775	20

From Table 7.11, across all samples, RTT ranged from 0.1 ms to 0.7 ms, with a mean of 0.486 ms and a standard deviation of 0.502 ms. These results indicate stable low-latency communication during the measurement period. Clock offset estimation showed a consistent client-server skew of approximately 1.77 s, suggesting stable but unsynchronized system clocks.

7.3.3 Results

7.3.3.1 Foil on scope test

Table 7.12 shows the descriptive statistics for the voltage recorded both with mobile phone turned on and off, while Table 7.13 shows the descriptive statistics for the voltage recorded with no foil, foil ungrounded and foil grounded.

Table 7.12: Table for descriptive statistics for mobile phone turned on and off

	Mean	Sample size	Std. deviation
Mobile off	28.640	15	7.915
Mobile on	28.907	15	8.944
Total	28.773	30	8.2992

Table 7.13: Table for descriptive statistics for no foil, foil ungrounded and foil grounded

	Mean	Sample size	Std. deviation
No foil	32.000	10	3.617
Ungrounded	36.240	10	2.416
Grounded	18.080	10	1.564
Total	28.773	30	8.2992

A two-way independent ANOVA was performed to analyze the effect of both mobile power and foil presence on voltage recorded. Levene's test cannot confirm that the assumption of homogeneity of variance had been met (F(5,24) = 3.461, p = 0.017). As the assumption of homogeneity of variance cannot be confirmed, non-parametric tests were conducted. Mann-Witney U. test was conducted for mobile power and Kruskal-Wallis H test for foil presence. Table 7.14 and Table 7.15 show the results of the Non-parametric tests.

From Table 7.14, the test revealed no statistically significant difference between mobile off and mobile on (U = 106.000, z = -0.271, p = 0.806, r = 19.353). Median values were 31.200 for mobile off and 33.600 for mobile on. These re-

Table 7.14: Table for results of Mann-Whitney U. test comparing the effect of environment on latency

Factor	U	z	p	r (Eff size)	Median Pow off	Median Pow on
Mobile power	106.000	-0.271	0.806	19.353	31.200	33.600

Table 7.15: Table for results of Kruskal-Wallis H. test comparing the effect of environment on latency

Factor	χ^2	df	p	Median No foil	Median Foil ungrounded	Median Foil grounded
Foil presence	22.893	2	<0.001	32.800	35.200	18.400

sults indicate that power difference does not elicit differential changes in voltage recorded.

However, from Table 7.15, the test showed statistically significant difference across the three foil presence conditions, $\chi^2(2) = 22.893$, $p < 0.001$. Median values were 32.800 for no foil (n = 10), 35.200 for foil ungrounded (n = 10), and 18.400 for foil grounded (n = 10). Post hoc pairwise comparisons with Bonferroni correction indicated that the pairs, foil grounded - no foil ($\chi^2(3) = 11.400$, $p = 0.004$) and foil grounded - foil ungrounded ($\chi^2(3) = 60$, $p < 0.001$), were significant. These results indicate that the presences of a grounded foil elicit differential changes in voltage recorded compared to both no foil and ungrounded foil.

7.3.3.2 Foil on EEG test

Table 7.16 shows the descriptive statistics for the power recorded with no foil, foil ungrounded and foil grounded.

A one-way independent ANOVA was performed to analyze the effect of foil presence on power recorded. Levene's test cannot confirm that the assumption of homogeneity of variance had been met ($F(2,90) = 4.268$, $p = 0.017$). As

Table 7.16: Table for descriptive statistics for no foil, foil ungrounded and foil grounded

	Mean	Sample size	Std. deviation
No foil	6.122E4	31	6.551E3
Ungrounded	3.513E6	31	1.154E7
Grounded	2.975E4	31	2.692E4
Total	1.201E6	93	6.796E6

such, Games-Howell was used for post hoc comparisons instead of Bonferroni. Table 7.17 shows the results of the one-way independent ANOVA.

Table 7.17: Table for results of one-way independent ANOVA comparing the effect of foil presence on power recorded

	df	df error	F	Sig.	n_p^2
Foil presence	2	90	1.243E14	0.066	0.058

From Table 7.17, the main effect of foil presence was not statistically significant (F(2, 90) = 1.243E14, p = 0.058, $n_p^2 = 0.058$). Since the comparison was not statistically significant, no post-hoc test was performed. These results indicate that there is not enough evidence to prove that foil presence elicit differential changes in power recorded.

7.3.3.3 XR systems comparison

Table 7.18 shows the descriptive statistics for the log scaled power recorded on each XR system, along with the baseline value taken from the no foil descriptive statistics in the earlier foil on EEG test.

A one-way independent ANOVA was performed to analyze the effect of the type of XR system on log scaled power recorded. Levene's test confirms that the assumption of homogeneity of variance had been met (F(5,180) = 1.219, p = 0.302). As such, Bonferroni was used for post hoc comparisons. Table 7.19 shows the results of the one-way independent ANOVA.

Table 7.18: Table for descriptive statistics for XR systems

	Mean	Sample size	Std. deviation
Baseline	4.7829	31	0.0682
Magicleap	3.524	31	0.0343
Mobile	3.379	31	0.495
Quest	4.853	31	0.125
Rift	3.887	31	0.0902
Vive	4.468	31	0.128
Total	4.149	186	0.593

Table 7.19: Table for results of one-way independent ANOVA comparing the effect of foil presence on power recorded

	df	df error	F	Sig.	n_p^2
System	5	180	1577.278	<0.001	0.978

From Table 7.19, the main effect of foil presence was statistically significant (F(5, 180) = 1577.278, $p < 0.001$, $n_p^2 = 0.058$). Post hoc pairwise comparisons with Bonferroni correction indicated that all pairs were significant, with $p<0.05$. These results indicate that different XR systems elicit differential changes in log scaled power recorded.

7.3.3.4 XR systems power comparison

Table 7.20 shows the descriptive statistics for the log scaled power recorded on each XR system, while Table 7.21 shows the descriptive statistics for the log scaled power recorded both with the XR systems turned on and off

A two-way independent ANOVA was performed to analyze the effect of the type of XR system and power on log scaled power recorded. Levene's test confirms that the assumption of homogeneity of variance had been met (F(9,300) = 0.986, p = 0.451). As such, Bonferroni was used for post hoc comparisons. Table 7.22 shows the results of the two-way independent ANOVA.

From Table 7.22, the main effect of System was statistically significant (F(4,

Table 7.20: Table for descriptive statistics for XR systems

	Mean	Sample size	Std. deviation
Magicleap	3.388	62	0.158
Mobile	3.593	62	0.220
Quest	4.594	62	0.300
Rift	3.878	62	0.0793
Vive	4.122	62	0.362
Total	3.915	310	0.487

Table 7.21: Table for descriptive statistics for XR systems

	Mean	Sample size	Std. deviation
Off	3.808	155	0.358
On	4.022	155	0.570
Total	3.915	310	0.487

Table 7.22: Table for results of two-way independent ANOVA comparing the effect of the type of XR system and power on log scaled power recorded

	df	df error	F	Sig.	n_p^2
System	4	300	1510.761	<0.001	0.953
Power	1	300	392.373	<0.001	0.567
Sys × Pow	4	300	330.444	<0.001	0.815

300) = 1510.761, p < 0.001, $n_p^2 = 0.953$). Similarly, significant main effect of Power was observed (F(1, 300) = 392.373, p < 0.001, $n_p^2 = 0.567$) and the System × Power interaction was also significant (F(4, 300) = 330.444, p < 0.001, $n_p^2 = 0.815$). Post hoc pairwise comparisons with Bonferroni correction indicated that all pairs were significant, with p<0.05. These results indicate that both different XR systems and power elicit differential changes in log scaled power recorded.

7.4 Discussion

The aim of this chapter was to characterise and quantify hardware-related confounds, namely luminance, latency, and electromagnetic noise across different experimental environments (2D screen, real world, AR, and VR), in order to ensure that subsequent behavioural and EEG results can be interpreted meaningfully. The results demonstrate that while some confounds can be corrected for analytically, others represent intrinsic differences between environments that must be explicitly acknowledged in the interpretation of later experiments.

The luminance analysis revealed statistically significant differences between all experimental environments, as well as smaller but significant differences between stimulus colours. These findings indicate that luminance was not successfully matched across conditions, despite identical nominal stimulus colours being specified in software.

Importantly, the observed differences are not unexpected. Each environment relies on fundamentally different display technologies: emissive plasma pixels in the 2D condition, reflected LED light in the real-world condition, waveguide-based optical projection in AR [231], and stereoscopic OLED/LCD panels in VR [228]. These systems differ not only in peak luminance but also in spectral output, optical attenuation, and viewing geometry. As a result, even when software parameters are matched, the physical light reaching the participant's retina varies substantially across environments.

Given the well-established inverse relationship between stimulus luminance and reaction time [236, 237, 238], these differences are likely to systematically bias behavioural responses if left unaccounted for. The fact that all environment pairs differed significantly suggests that luminance should not be treated as random noise but rather as a systematic confound. Consequently, luminance measurements from this chapter should be used as covariates or correction factors when interpreting reaction time differences in the subsequent visual perception experiment.

In contrast to luminance, latency measurements showed no statistically significant differences across environments or stimulus colours. Median latencies were effectively zero in all conditions, and variability was minimal. Together with the TCP round-trip measurements, these results indicate that the experimental pipeline—comprising stimulus rendering, trigger generation, network transmission, and EEG recording—operates with stable and negligible temporal offsets relative to the timescales relevant for reaction time and EEG analysis.

This finding has two important implications. First, no environment-specific latency correction is required when analysing reaction times in the next chapter. Reaction time differences can therefore be attributed to participant behaviour rather than hardware-induced delays. Second, the temporal precision of the trigger–EEG alignment is sufficient for time-locked EEG analysis, including event-related potentials and time–frequency measures. This simplifies both the statistical modelling and the interpretation of cross-environment differences, strengthening confidence that observed effects reflect genuine perceptual or cognitive processes rather than technical artefacts.

The noise analysis yielded more nuanced results. The oscilloscope-based tests demonstrated that a grounded aluminium foil enclosure significantly reduced recorded electromagnetic interference, whereas an ungrounded foil increased noise levels relative to no foil. This is consistent with Faraday cage principles: without grounding, the foil can act as an antenna rather than a shield [241].

However, when the foil was applied over an EEG cap, no statistically significant reduction in broadband EEG power was observed. Several factors may explain this apparent discrepancy. First, EEG noise is dominated by multiple sources, including electrode–skin impedance fluctuations, amplifier noise, and physiological artefacts, which may mask modest reductions in environmental EMI. Second, the foil configuration used here does not constitute a fully enclosed or uniformly grounded Faraday cage, limiting its effectiveness [241]. Third, the use of raw broadband power as the outcome measure may lack sensitivity to narrowband EMI reductions.

By contrast, the XR system comparison revealed robust and large differences in log-scaled EEG power between devices, as well as significant interactions with power state (on vs off). These results indicate that XR headsets themselves are substantial sources of electromagnetic noise, and that this noise differs qualitatively and quantitatively between devices. Importantly, these effects persisted even when stimuli were not being presented, suggesting that headset electronics alone can influence EEG recordings.

Taken together, the noise results suggest that post-processing strategies are likely to be more effective than simple external shielding in mitigating XR-induced EEG contamination. Moreover, headset-specific noise profiles should be considered when comparing EEG results across XR platforms.

Here, raw power was used for the foil-on EEG test and log-transformed power for the XR system comparisons. This is because raw EEG power is highly skewed and sensitive to extreme values, as reflected in the large variances observed in the foil test. In that context, raw power was appropriate for detecting gross changes potentially introduced by shielding.

In contrast, comparisons across XR systems required stable variance and sensitivity to relative differences across conditions and devices. Log transformation is standard practice in EEG analysis for precisely this reason, as it normalises distributions and improves statistical robustness. Therefore, the use of different power representations reflects different analytical goals rather than an inconsistency in methodology.

Although aluminium foil grounding significantly reduced electromagnetic interference at the oscilloscope level, it did not produce a statistically reliable reduction in EEG power when applied over the EEG cap. Given the absence of a robust shielding effect at the level of the recorded neural signal, the foil condition was not included in subsequent XR system comparisons. Excluding the foil factor also avoided introducing substantial random variables arising from variance differences between shielding conditions that did not meaningfully impact EEG measurements.

7.5 Conclusion

This chapter established a validated hardware and software platform for conducting EEG experiments across 2D, real-world, AR, and VR environments, and provided a detailed characterisation of key technical confounds. Luminance differed significantly across all environments, necessitating careful interpretation or correction of reaction time results. In contrast, system latency was minimal and consistent, supporting reliable behavioural and EEG timing analysis. Noise analysis demonstrated that XR headsets are non-negligible sources of electromagnetic interference, with effects that vary by device and power state, while simple foil-based shielding offered limited benefit at the EEG level.

Together, these findings define the constraints and assumptions under which the visual perception experiment in the next chapter should be interpreted. By explicitly quantifying and contextualising these hardware factors, the platform enables more robust conclusions about perceptual and neural differences between real, virtual, and augmented environments, rather than conflating such effects with uncontrolled technical artefacts.

Chapter 8

EEG Reaction Time Experiment

8.1 Introduction

In the previous chapters, we have discussed topics ranging from mathematics to technology to neuroscience. In this chapter, we put all these knowledge together in an attempt to push the boundaries of human knowledge on visual perception.

Our visual perception is a complex part of our cognition, including pattern recognition, object identification, depth estimation and motion sensing. Two of the early theories of perception are the Gibson's direct theory of perception, also known as the 'bottom-up' theory [242], and Gregory's constructivist (indirect) theory of perception, also known as the 'top-down' theory [243]. In recent years, with the advancement of technology, researchers have been trying to pinpoint the neural correlates of specific visual functions. One neural pathway that was proposed was a pathway that originate from the striate cortex, and separated into a dorsal stream, which reaches the posterior parietal cortex, and a ventral stream, which reaches the inferior temporal region [244]. Later, an fMRI experiment was conducted to investigate the functions of these two streams, which concluded that the dorsal pathway relates to visual spatial processing and the ventral pathway relates to visual form processing. [245].

Past research have found that real and virtual visual information are processed differently in our visual cortex. In an experiment conducted by Richardson, Montello and Hegarty to investigate the difference in spatial knowledge acquisition from maps, real environment and virtual environment, it was found that there were significant differences in spatial representations found after learning the environment from the real environment and from the virtual environment [246]. While this study have shone light on the difference in the cognitive mechanisms involved in processing real and virtual environments, one drawback of the study is that the virtual environment was presented to the participants on a 2D screen, which may have affected factors such as the participants' ability to perceive depth, and the level of immersion in the environment.

8.2 Real vs digital visual stimuli

The visual perception of digital visual stimuli and real visual stimuli involves the same neurological pathways. However, there are subtle differences between the nature of the stimuli and their interaction with the human visual system, due to the medium of presentation.

Light of different wavelengths are emitted or reflected differently by real and digital stimuli. As Babilion and Khanh (2018) have pointed out, for real objects, colour perception is influenced by physical properties such as texture, shape, and reflectance, combined with ambient lighting conditions and chromatic adaptation [247]. Real objects provide consistent sensory information that aligns with memory colours—typical colours our brain associates with familiar objects based on long-term memory. In contrast, virtual objects are displayed through screens or projection systems, where their colour appearance is heavily dependent on the digital rendering processes, display calibration, and the viewing environment [248]. Virtual stimuli often lack the textural and depth cues present in real-world objects, which can affect the perception of chromaticity and lightness [249].

Another difference between real and digital stimuli is that digital stimuli are typically presented on a 2D screen, hence requiring the visual system to interpret artificial depth cues, such as size, perspective, or shading, to simulate three-dimensionality. Even in a virtual environment, where 3D stimuli can be presented, there can still be differences in how people perceive depth in these environments. This is due to limitations in reproducing certain visual and physiological depth cues in virtual environment. In the real-world, depth perception relies on a combination of binocular (e.g., stereopsis) and monocular (e.g., texture gradient, motion parallax) cues, as well as proprioceptive factors like accommodation and convergence, which together allow accurate spatial judgments [250]. However, in virtual environments, particularly when using head-mounted displays (HMDs), these depth cues are often simulated or incomplete. For example, stereoscopic rendering in HMDs decouples convergence and accommodation, leading to vergence-accommodation conflicts that impair depth perception and cause discomfort. Additionally, users in VEs frequently underestimate distances due to restricted fields of view, unrealistic textures, and the absence of consistent depth cues, such as ground continuity [250].

There are other issues with virtual stimuli, such as limitations by screen resolution and refresh rates. A low screen resolution or refresh rate can introduce motion artifacts, such as stuttering or aliasing, and affect the perception of smooth motion. Also, digital stimuli are primarily visual or auditory. They do not provide the same multisensory input from real stimuli. Although VR and AR technologies are attempting to bridge this gap by incorporating more sensory modalities, they still cannot fully replicate the complex sensory input provided by real-world stimuli.

While the human visual system can adapt to interpret both types of stimuli, the inherent differences in presentation and sensory fidelity create distinct perceptual experiences. The current research attempts to uncover more about the mechanisms behind this difference in perception.

8.2.1 Reaction time studies

Reaction time studies are a great approach to help us understand how we react to visual stimulus. In traditional reaction time studies, stimuli are usually shown on a 2D screen, and participants would have a button or a keyboard, which they can press in response to the visual stimuli. For example, Perdices and Cooper ran a reaction time study on patients with human immunodeficiency virus (HIV), where the stimulus was presented on a 2D computer screen, and the participants were asked to respond by pressing a keyboard key [40]. In another study, Goswami et. al. wanted to study the effect of coorelated colour temperature (CCT) of an LED light on reaction time. They built a system where the LED light stimulus was connected to a response button through a microcontroller [41].

The use of AR and VR systems expand the scope of neuroscience experiments. These systems allow the experimenter to present any stimulus they want to the participant, anywhere within the participant's visual field, and the potential response from the participant is expanded beyond a key or button press. Antão Et. Al., studied the use of AR task in improving the reaction time of people with autism spectrum disorder (ASD) [251]. In one of the task, alphabets were projected onto the real world, rendered on a computer screen, and the participants were asked to reach the same symbol, causing the virtual object to react to a gesture. It was concluded that AR can provide a novel and beneficial experiment for people with ASD in education.

8.2.2 The use of EEG in the study of perception

Electroencephalogram (EEG) has been used to study a wide variety of neuroscience phenomena, including cognition, affect and neuropathologies. Visual processing have also been studied using EEG, as can be seen from the studies discussed above [224, 226]. The advantage of using EEG over other neuroimaging techniques include mobility, high temporal resolution and easy to maintain. Furthermore, in AR and VR studies, AR and VR headsets can be fitted over an EEG cap, thus expanding the types of virtual stimuli that can be shown to the participants.

Several studies have identified the event-related potential of visual perception with the use of EEG. [252] recorded the EEG of participants as they are perceiving known and unknown visual stimuli. They found that the amplitude of event-related gamma oscillations during visual perception correlates significantly with the phase of theta oscillations occurring at the same time. Bush, Dubois and VanRullen also recorded the EEG of participants as brief flashes of light, at the luminance threshold of each participant [253]. They found that detection probability was highly dependent on the phase of the participants' spontaneous EEG oscillations in the low alpha and theta bands just before stimulus onset.

8.2.3 Current study

The purpose of this research is to investigate how one's reaction time and brain activities will be affected when stimulus is being presented from different environments, namely AR, VR and 2D and real-world environments. We hypothesise that there will be a difference in terms of reaction time and brain activity when participants are placed in these different environments.

8.3 Methods

8.3.1 Participants

This study was approved by the Ethics Committee University of York, Department of Electronics Engineering. 8 participants were recruited for the study. All participants signed the consent form for the experiment. However, 2 of the participants were rejected due to existing neurological conditions, one of whom had Irlen Syndrome, which will be explored in the next chapter. 3 of the data set recorded were found to be too noisy and hence rejected.

8.3.2 Procedure

The 2D environment was built based on Perdices and Cooper's research, where patients with human immunodefeciency virus (HIV) were recruit in a two-part study, the first part being a simple reaction time (RT) experiment and the second part being a choice RT experiment [40]. The experiments were done using a computer, which participants use to respond to target stimuli. Only target stimuli were shown in the simple RT experiment, while both target and non-target stimuli were shown in the choice RT experiment.

In this study, Psychopy was used to build a similar experiment for the 2D environment. In the simple RT experiment, participants were shown a yellow square, as shown in Figure 8.1, and the participants were asked to press the space bar when the cube appears. In the choice RT experiment, participants were shown either a red or a green square, as shown in Figure 8.2. Participants were asked to press the space bar only when a green square is displayed. Participants sat at a distance of about 1.5 m away from the TV screen, a Samsung PS50C450B1W.

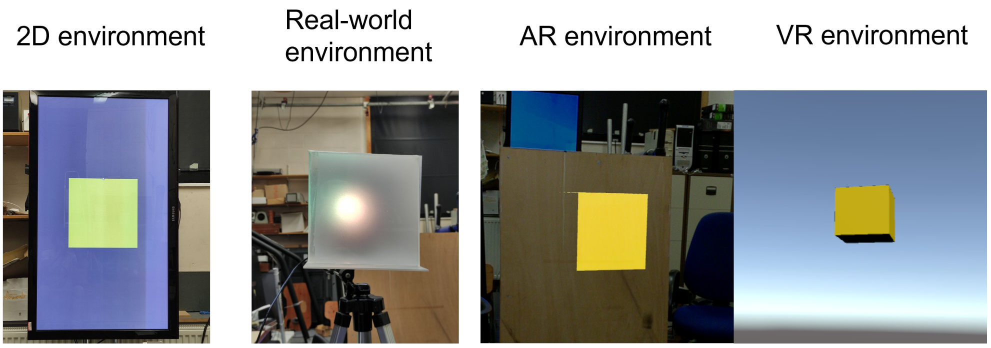


Figure 8.1: Stimuli shown to participants during simple reaction time experiment

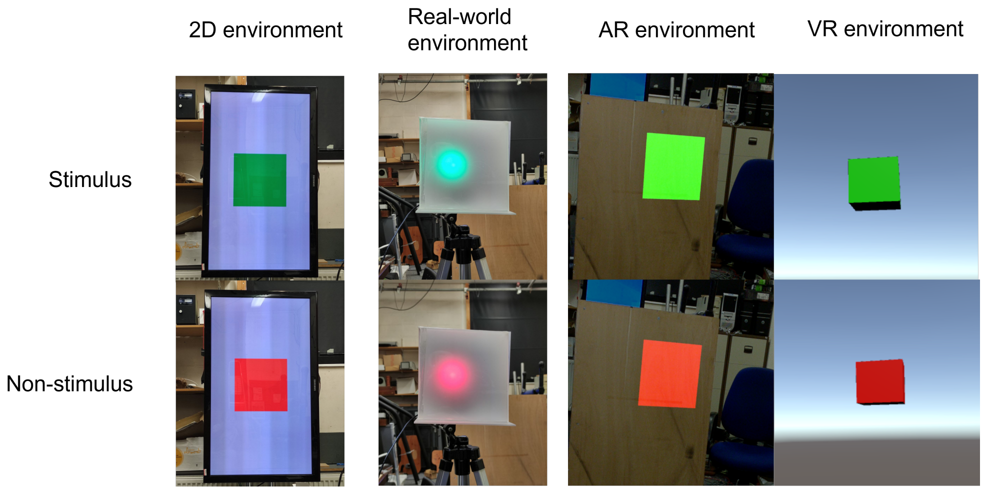


Figure 8.2: Stimuli shown to participants during choice reaction time experiment

The trials conducted in the other environments were similar to that in the 2D environment, except that a cube was shown in place of a square, as can be seen in Figure 8.1 and Figure 8.2. In the real world environment, an LED light was placed in a cube built using frosted acrylic boards, measuring 25cm x 25cm x 25cm. In the virtual world environments (AR and VR), a virtual cube was generated.

8.3.3 EEG recording

EEG was recorded using the Brainvision ActiCHamp system (Brain Products GmbH, Gilching, Germany). All tasks were recorded with 32-channel EEG through ActiCap Slim active Ag/AgCl electrodes, with a sampling rate of 500 Hz. Electrodes were placed in the international 20/20 system, embedded in the ActiCap. Data recording was controlled using Openvibe, and impedances were kept below 20 kOhm.

8.3.4 Analysis

The analysis for the EEG data collected in this experiments follows the pipeline introduced in Chapter 6. Howerver, the CSP function, wavelet function and statistical analysis steps were modified to better fit the nature of the data collected in this experiment.

8.3.4.1 Updated CSP function

For the main experiment, the CSP pipeline introduced in Chapter 6 was modified to computed CSP at the trial level, allowing spatial filters to be derived separately for each trial or trial pair. This approach was adopted due to the paired structure of the data. Specifically, each experimental trial comprised a baseline segment immediately preceding stimulus onset and a post-stimulus segment immediately following stimulus onset. These two segments form a linked pair originating from the same trial, recording session, and ongoing brain state.

For each trial, data from the two classes were arranged into channel-by-time matrices and submitted to a CSP decomposition. Spatial filters were obtained by solving the generalized eigenvalue problem between the two trial-specific covariance matrices, yielding a set of spatial filters that maximized variance for one class while minimizing variance for the other. The resulting demixing matrix, spatial patterns, and eigenvalues were stored on a per-trial basis. These spatial filters were then applied to the corresponding trial data to produce spatially filtered signals for each condition.

Unlike the CSP implementation in Chapter 6, this version did not perform explicit covariance normalization, shrinkage regularization, or class-level averaging across trials. As a result, the spatial filters reflect trial-specific variance structure rather than a single set of global class-level spatial patterns. This design choice prioritizes flexibility and exploratory analysis over statistical robustness, and is therefore more sensitive to trial-wise variability and noise.

8.3.4.2 Updated wavelet function

For the main experiment, the wavelet analyses were performed at the participant level. Time–frequency representations (TFRs) of oscillatory power were computed using a sliding-window Fourier transform approach implemented in FieldTrip [12]. Spectral estimation was performed using a Hanning taper with a fixed temporal window length of 0.5 s for all frequencies.

Power spectra were computed for frequencies ranging from 2 to 40 Hz in steps of 2 Hz. Time–frequency estimates were calculated across all available time points in each trial. The output of this step was a power-based TFR for each participant and condition.

Scalp topographies of oscillatory power were generated from the time–frequency data. For visualization, power was averaged within a predefined time window (0.3–0.6 s post-event) and frequency band (8–12 Hz), corresponding to the alpha band. Color scaling was set to maximum absolute values to allow symmetric visualization of power differences across the scalp.

All topographical plots used a standard 64-channel EEG layout (ActiCAP) and included a colorbar for interpretability.

8.3.4.3 Reaction time preprocessing

Reaction times (RTs) were preprocessed in two sequential steps: first, a Piéron-law–based luminance correction was applied to account for systematic differences in stimulus intensity across experimental environments. Second, extreme RT values were removed using a conservative outlier-filtering pipeline.

RTs were corrected using a Piéron-law model ($RT = \alpha + \beta \cdot L - k$, where L is the physical stimulus luminance (cd/m^2), α is the irreducible (non-sensory) processing time, β is the scaling constant and k is the luminance exponent) [254], which captures the well-established inverse relationship between reaction time and stimulus intensity [236, 237, 238]. Stimulus intensity was indexed using the RMS voltage of a photodiode sensor, which had baseline illumination subtracted. Model parameters (α, β, k) were estimated via nonlinear least-squares fitting of the observed RTs to the measured stimulus intensity. Corrected RTs were then calculated by removing the luminance-dependent component and re-expressing each trial relative to a common reference intensity (the mean effective voltage across trials). This procedure ensures that differences in RT reflect perceptual or cognitive processing rather than differences in physical stimulus intensity.

After luminance correction, RTs were filtered to remove extreme values that likely reflect lapses or anticipatory responses. Trials with RTs below 150 ms or above 3000 ms were first removed. Subsequently, a participant-wise, standard deviation-based trimming procedure was applied, excluding any trial with a corrected RT more than 3 standard deviations from the participant's mean. This two-step procedure preserves the central tendency and variability of the RT distribution while minimizing the influence of extreme values [255, 256].

8.3.4.4 Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics for Windows, version 31 (Armonk, NY: IBM Corp). A one-way repeated-measures ANOVA was used because all experimental conditions were measured within the same participants, making the observations across conditions statistically dependent. This analysis treats participants as their own control, thereby accounting for between-subject variability and isolating the effect of the experimental manipulation. Given the small number of participants, this within-subject approach is particularly appropriate, as it avoids pseudoreplication that would arise from treating individual trials as independent observations and ensures that statistical inference is based on variability between participants rather than within-participant noise.

Because only three participants provided viable EEG data, formal statistical inference on the EEG signals was not performed. Statistical analysis at the channel or source level would be severely underpowered and risk producing unreliable or misleading results. Instead, the EEG data were analyzed descriptively by computing and visualizing topographical maps of average signal power following Common Spatial Pattern (CSP) spatial filtering. For each condition, CSP-filtered signals were obtained and spectral power was averaged across trials. The resulting mean power values were then projected onto scalp topographies to illustrate the spatial distribution of neural activity associated with the experimental conditions. This approach allows qualitative assessment of consistent spatial patterns across participants while avoiding inappropriate statistical testing on a sample size insufficient for reliable inferential analysis. Accordingly, these topographical maps are interpreted as descriptive summaries of the data rather than as statistically validated effects.

8.4 Results

8.4.1 Reaction time

8.4.1.1 Simple reaction time experiment

Table 8.1 shows the descriptive statistics for the reaction times of the participants in the four different environments, during the simple reaction time experiment.

Table 8.1: Table for descriptive statistics for the reaction times of the participants in the four different environments, during the simple reaction time experiment

	Mean	Sample size	Std. deviation
2D screen	0.318	3	0.0570
Real world	0.268	3	0.0369
AR set	0.449	3	0.0894
VR set	0.392	3	0.0368

Figure 8.3 shows the box plot of the reaction time recorded in each environment in the simple RT experiment.

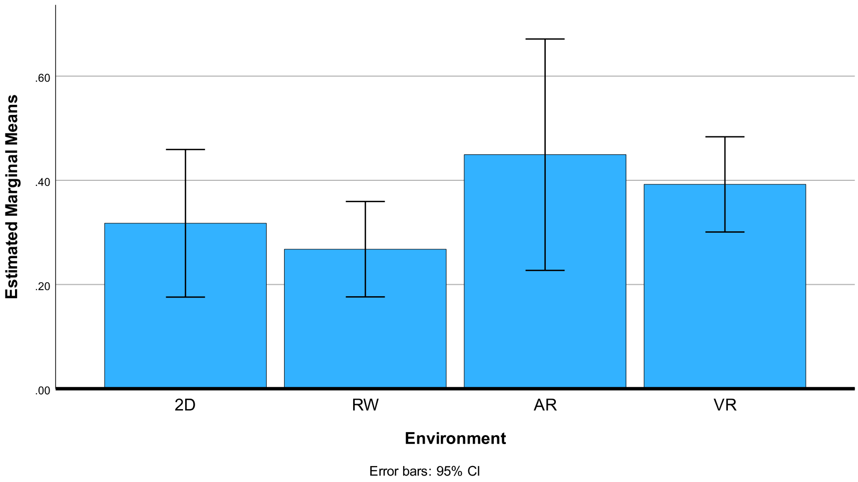


Figure 8.3: Bar graph of simple reaction time experiment

From Figure 8.3, the order of the reaction time, from fastest to slowest, is as follows: real world (0.268 ± 0.0369), 2D screen (0.318 ± 0.0570), VR set (0.449 ± 0.0894) and AR set (0.392 ± 0.0368). However, this order cannot be taken as statistically significant without a formal statistical test.

A one-way repeated-measures ANOVA was performed to analyze the effect of the reaction times of the participants in the four different environments, during the simple reaction time experiment. Mauchly's test cannot be performed due to the low number of subjects. Hence, the assumption of homogeneity of variance cannot be confirmed. As such, Greenhouse-Geisser correction was used. Table 8.2 shows the results of the one-way independent ANOVA.

Table 8.2: Table for results of one-way repeated-measures ANOVA comparing the effect of environment on reaction time

	df	df error	F	Sig.	n_p^2
Environment	1.059	2.119	6.679	0.116	0.770

From Table 8.2, the main effect of environment was not statistically significant (F(1.059, 2.119) = 6.679, p = 0.116, $n_p^2 = 0.770$). Since the comparison was not statistically significant, no post-hoc test was performed. These results indicate that there is not enough evidence to prove that environment elicit differential changes in reaction time.

8.4.1.2 Choice reaction time experiment

Table 8.3 shows the descriptive statistics for the reaction times of the participants in the four different environments, during the choice reaction time experiment.

Table 8.3: Table for descriptive statistics for the reaction times of the participants in the four different environments, during the choice reaction time experiment

	Mean	Sample size	Std. deviation
2D screen	0.436	3	0.0453
Real world	0.365	3	0.0320
AR set	0.572	3	0.0439
VR set	0.463	3	0.0347

Figure 8.4 shows the box plot of the reaction time recorded in each environment in the choice RT experiment.

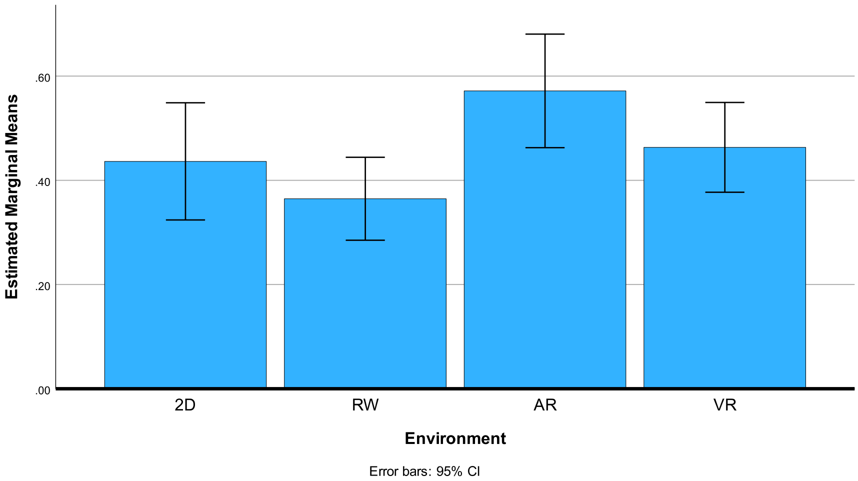


Figure 8.4: Bar graph of choice reaction time experiment

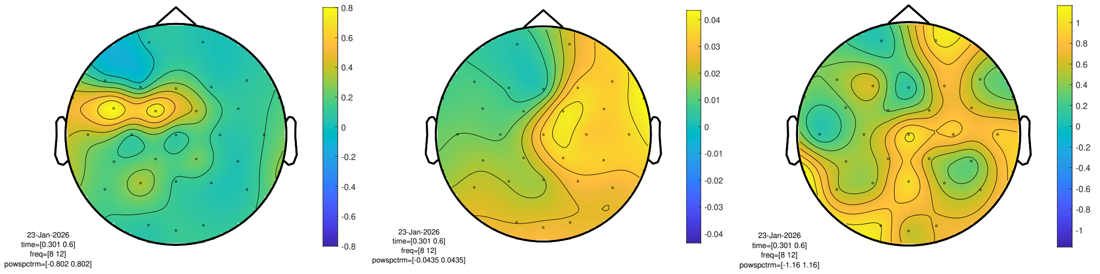
From Figure 8.4, the order of the reaction time, from fastest to slowest, is as follows: real world (0.365 ± 0.0320), 2D screen (0.436 ± 0.0453), VR set (0.463 ± 0.0347) and AR set (0.572 ± 0.0439). However, this order cannot be taken as statistically significant without a formal statistical test.

A one-way repeated-measures ANOVA was performed to analyze the effect of the reaction times of the participants in the four different environments, during the simple reaction time experiment. Mauchly's test cannot be performed due to the low number of subjects. Hence, the assumption of homogeneity of variance cannot be confirmed. As such, Greenhouse-Geisser correction was used. Table 8.4 shows the results of the one-way independent ANOVA.

Table 8.4: Table for results of one-way repeated-measures ANOVA comparing the effect of environment on reaction time

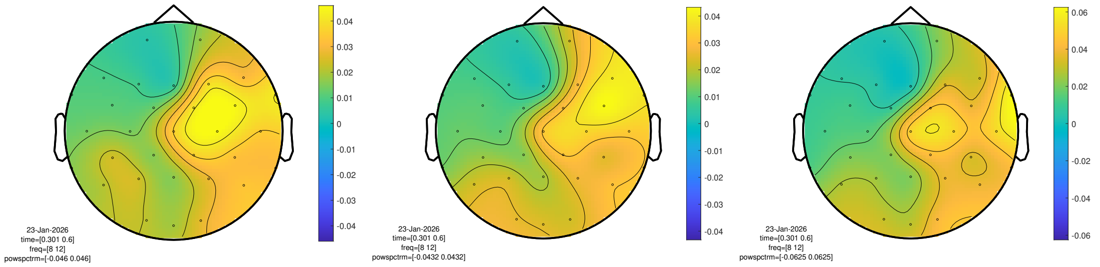
	df	df error	F	Sig.	n_p^2
Environment	1.721	3.443	11.220	0.032	0.849

From Table 8.4, the main effect of foil presence was statistically significant (F(1.721, 3.443) = 11.220, p = <0.032, $n_p^2 = 0.849$). However, post hoc pairwise comparisons with Bonferroni correction did not reveal any significant differences between individual time points (all p > 0.05).



(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.5: EEG topographical map of participants during simple reaction time on 2D screen



(a) Participant 1 (b) Participant 2 (c) Participant 3

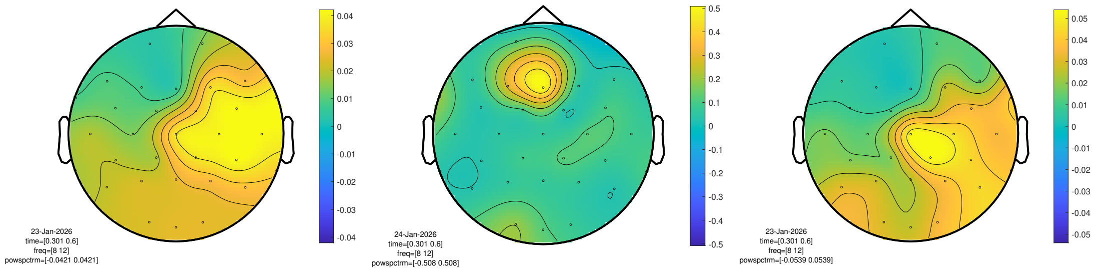
Figure 8.6: EEG topographical map of participants during simple reaction time in real world environment

8.4.2 EEG topographical map

8.4.2.1 Simple reaction time

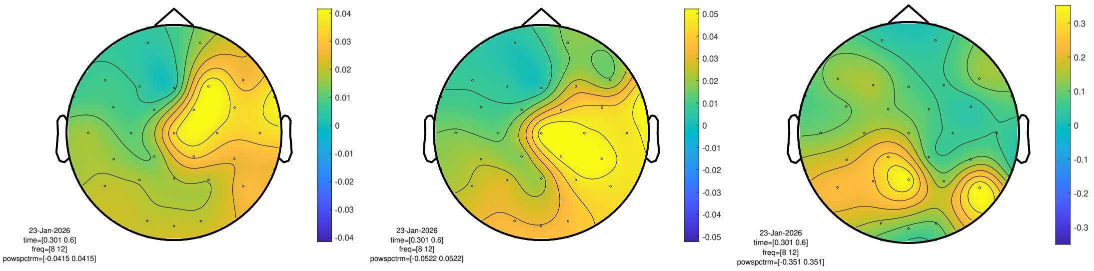
Figures 8.5, 8.6, 8.7, 8.8 show the alpha-band (8–12 Hz) CSP-filtered EEG topographies during the simple reaction time task.

Visual inspection of the topographies during the simple reaction time task did not reveal any consistent spatial patterns across participants or experimental en-



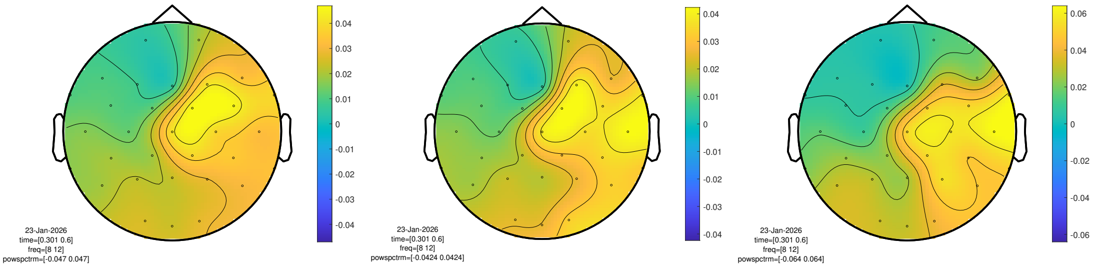
(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.7: EEG topographical map of participants during simple reaction time in AR environment



(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.8: EEG topographical map of participants during simple reaction time in VR environment



(a) Participant 1 (b) Participant 2 (c) Participant 3

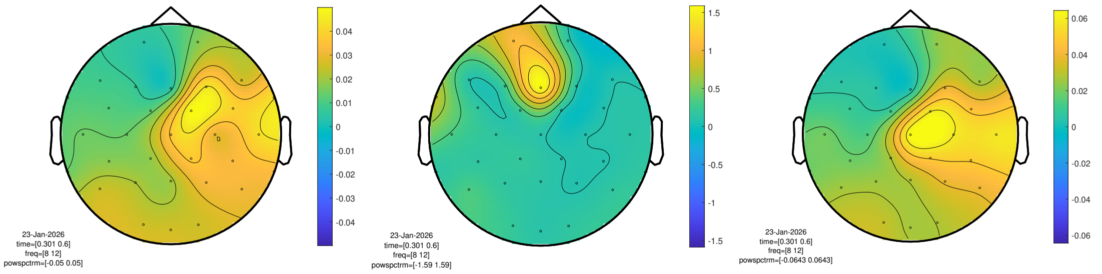
Figure 8.9: EEG topographical map of participants during simple reaction time on 2D screen

vironments. While individual participants exhibited localized changes in spectral power following stimulus onset, the spatial distribution, polarity, and magnitude of these effects varied substantially between subjects. No reproducible topographical features were observed that were common across all three participants for any environment (2D, real-world, AR, or VR). These results indicate a high degree of inter-individual variability in neural responses during simple reaction time tasks under the present experimental conditions.

8.4.2.2 Choice reaction time - reacted

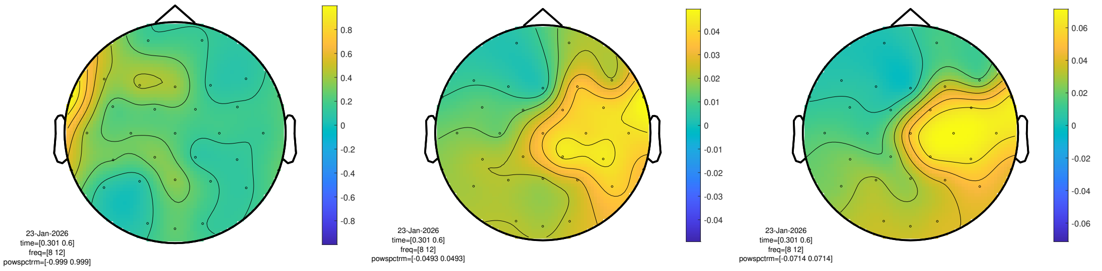
Figures 8.9, 8.10, 8.11, 8.12 show the alpha-band (8–12 Hz) CSP-filtered EEG topographies during the choice reaction time task, when the green stimulus was shown and the participant reacted to it.

Similar to the simple reaction time experiment, EEG topographical maps corresponding to reacted trials in the choice reaction time task did not show consistent or systematic spatial patterns across participants. Although condition-



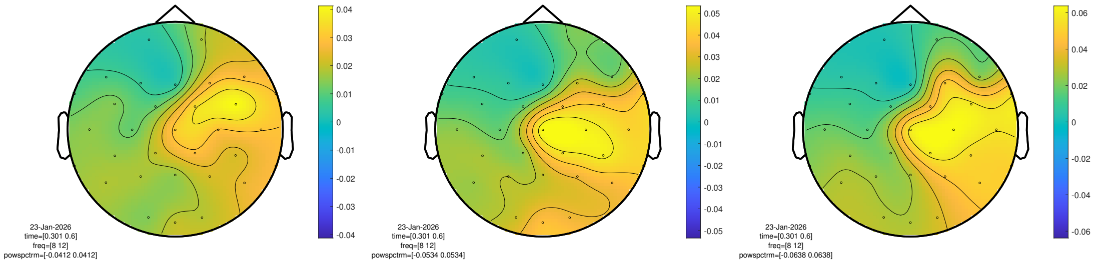
(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.10: EEG topographical map of participants during simple reaction time in real world environment



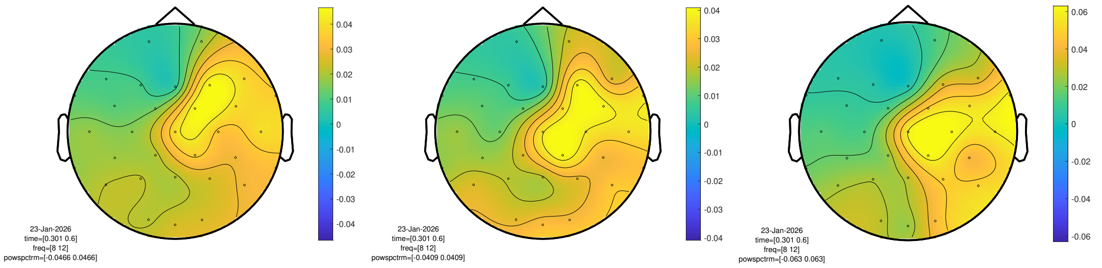
(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.11: EEG topographical map of participants during simple reaction time in AR environment



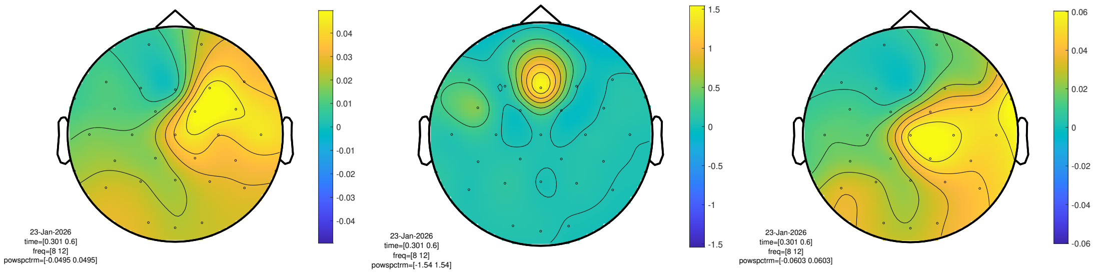
(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.12: EEG topographical map of participants during simple reaction time in VR environment



(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.13: EEG topographical map of participants during simple reaction time on 2D screen



(a) Participant 1 (b) Participant 2 (c) Participant 3

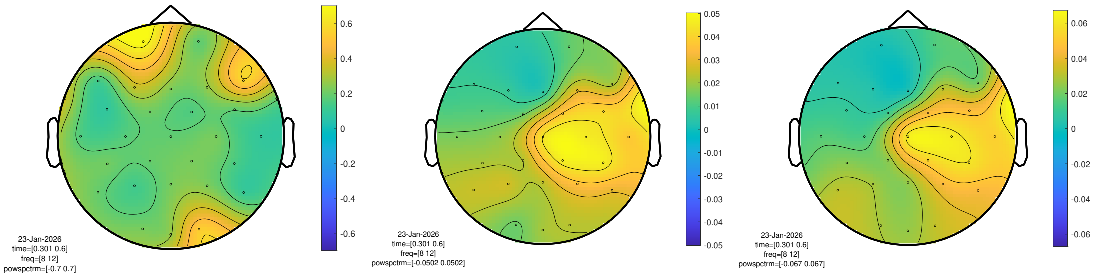
Figure 8.14: EEG topographical map of participants during simple reaction time in real world environment

specific differences could be observed within individual participants, these effects were not aligned in scalp distribution or direction across subjects. As such, no common neural signature associated with successful target detection could be identified at the group level.

8.4.2.3 Choice reaction time - not reacted

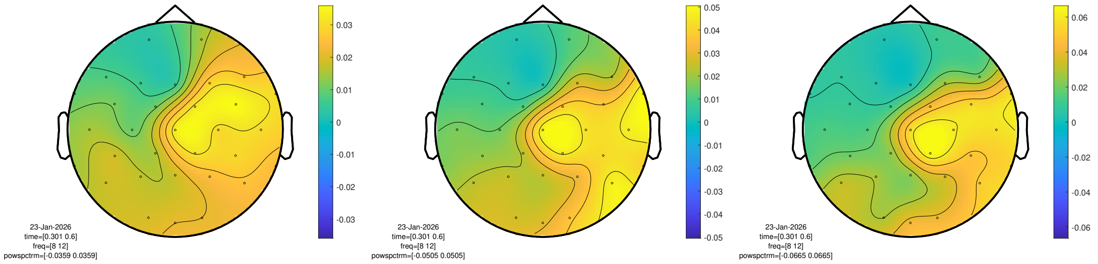
Figures 8.13, 8.14, 8.15, 8.16 show the alpha-band (8–12 Hz) CSP-filtered EEG topographies during the choice reaction time task, when the red stimulus was shown and the participant avoided it.

For avoided trials, alpha-band topographies again demonstrated substantial variability across participants and conditions. No shared spatial features were evident that differentiated non-reacted trials across environments or participants. The absence of consistent patterns suggests that neural activity during missed responses is strongly influenced by individual-specific factors and ongoing brain states, rather than by the experimental manipulation alone.



(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.15: EEG topographical map of participants during simple reaction time in AR environment



(a) Participant 1 (b) Participant 2 (c) Participant 3

Figure 8.16: EEG topographical map of participants during simple reaction time in VR environment

8.5 Discussion

The present study investigated how reaction time and EEG activity differ when visual stimuli are presented in four environments: a 2D screen, the real world, an augmented reality (AR) environment, and a virtual reality (VR) environment. Across both behavioural and electrophysiological measures, the results suggest that the mode of stimulus presentation influences perceptual processing, although the extent and consistency of these effects are constrained by the limited sample size.

In the reaction time results, participants generally responded more slowly to stimuli presented in virtual environments compared to those presented in the real world or on a 2D screen. This trend was observed in both the simple and choice reaction time tasks, although statistical significance was limited. One possible explanation for this effect relates to ecological validity, whereby perceptual systems are more finely tuned to stimuli that resemble real-world, naturalistic settings [257]. When participants are placed in VR or presented with virtual objects in AR, additional perceptual and cognitive resources may be required to interpret artificial depth cues, resolve discrepancies between visual and proprioceptive information, or adapt to the virtual context, resulting in longer response times.

Interestingly, reaction times in AR were often slower than in the real-world condition but not consistently slower than in VR. This pattern may reflect differences in adaptation demands between the two immersive technologies. In VR, participants are fully immersed in a coherent virtual environment, which may facilitate rapid contextual adaptation once immersion is established. In contrast, AR requires participants to integrate virtual objects with a real-world background, potentially increasing cognitive load as the brain reconciles competing perceptual frameworks [258, 259]. While this interpretation remains speculative, it highlights a potential distinction between AR and VR that warrants further investigation.

The EEG results were analysed descriptively due to the very small number of participants providing usable EEG data ($n = 3$). Visual inspection of CSP-filtered alpha-band (8-12 Hz) topographical maps did not reveal consistent spatial patterns across participants, task conditions, or experimental environments. Although individual participants exhibited localized changes in spectral power following stimulus onset, the spatial distribution, polarity, and magnitude of these effects varied substantially between subjects. As a result, no reproducible group-level neural signatures could be identified for any condition.

The absence of consistent EEG patterns is likely attributable to several factors. First, EEG signals are known to exhibit strong inter-individual variability, particularly in oscillatory activity and scalp topographies [260]. Second, the use of trial-level Common Spatial Pattern (CSP) analysis without covariance regularization prioritises sensitivity to trial-specific variance at the expense of cross-trial and cross-subject stability. Finally, the extremely limited sample size severely restricts statistical power and precludes reliable inference at the group level. Consequently, the EEG findings should be interpreted strictly as qualitative, descriptive summaries rather than evidence of systematic neural differences between environments.

Several methodological limitations of the present study should be acknowledged. Most notably, the small number of participants limits the generalisability of both the behavioural and electrophysiological results. While within-subject statistical analyses were applied to the reaction time data, the low number of subjects reduces sensitivity to detect reliable effects and increases the likelihood that observed trends reflect individual differences rather than population-level phenomena. For the EEG data, the limited sample size entirely precluded inferential statistical analysis.

Additionally, the visual stimuli used in this study were intentionally simple, consisting only of red, green, and yellow cubes. While this design choice reduced stimulus complexity and experimental confounds, it also limits the ecological and perceptual richness of the task. Future studies could explore how variations in shape, texture, colour, and semantic content influence perception and neural processing across real and virtual environments.

8.6 Conclusion

This study examined how visual stimuli presented in real-world, 2D, AR, and VR environments influence reaction time and EEG activity. The behavioural results suggest that stimulus presentation environment can affect reaction time, with a tendency toward slower responses in immersive virtual environments. However, these effects were modest and should be interpreted cautiously due to the limited sample size.

The EEG analyses, conducted descriptively, did not reveal consistent spatial patterns across participants or conditions. This lack of reproducible EEG signatures highlights the substantial inter-individual variability in neural responses and underscores the challenges of drawing group-level conclusions from small-sample EEG studies, particularly when using exploratory, trial-level spatial filtering approaches.

Despite these limitations, the study demonstrates the feasibility of combining EEG with AR and VR paradigms and provides preliminary insights into how different presentation environments may shape perceptual processing. The findings highlight the importance of adequately powered studies and carefully controlled stimulus designs in future work aimed at understanding the neural mechanisms underlying perception in immersive and mixed-reality environments.

Chapter 9

Case study on Irlen Syndrome and digital stimuli

9.1 Introduction

Meares-Irlen syndrome, or simply Irlen syndrome (IS) is a condition of perceptual dysfunction, characterized by visual stress and distortions while reading, such as blurriness, double vision, or color changes around text. This is due to poor adaption to light source, luminance, intensity, wavelength and colour contrast. This has resulted in individuals with this syndrome being slow and ineffective readers with low comprehension. It can also cause headaches, eye strain, and difficulty focusing, which interfere with reading and learning [261]. Figure 9.1 shows some of the types of visual distortions people with IS experience [17].

Figure 9.1: Types of visual distortions with IS [17]

The current treatment recommends the use of filters and coloured lenses in order to reshape the light spectrum, reducing the contrast between light and dark to facilitate visual and retinal photoreceptor adaptation. However, research on IS is still limited. Most of the research conducted on IS focusing on the subject's reading ability. However, little is known about how individuals with IS react to digital stimuli. Reaction time studies are a good way to help us understand how individuals with IS react to visual stimuli, both with filter and without.

Chouinard et. al. (2012) conducted a behavioural and neuroimaging study on a case study participant with Irlen Syndrome, using a language experiment with varied proportions of regular and exception words[122]. The researchers found that in the case study participant, there was a significant increasing trend of impulse response function (IRF), a function a system's response to an impulse, in BA17 to BA19 (visual cortex), BA6 (premotor and suplementary motor cortex) and postcentral gyrus. No such increasing trend of IRF were found in the control subject. This demonstrates that in the case study participant, there was hyperexcitablility in the visual cortex, which spread to the other areas of the brain. While there are many limitations with this study, it demonstrated the possible existence of neural correlates of Irlen Syndrome, prompting further research into the neuro-mechanisms of this condition.

In a larger neuroimaging study with 15 Irlen syndrome patients, Kim et. al. (2015) reported significant activation in the superior temporal gyri, specifically the left middle and superior temporal gyri, during sentence reading after wearing color-tinted lenses, compared to reading unaided [262]. Interestingly, superior temporal gyrus is primarily associated with auditory processing, particularly in speech recognition and comprehension, but not in the visual processing of written language. However, the Left middle temporal gyrus is involved in semantic processing and the integration of language information, which are amodal processes [262]. This suggests that the root cause of Irlen Syndrome might not be with the primary visual processing of written words, but with the visual-auditory integration of language information.

As discussed briefly in Chapter 3, Irlen syndrome is a condition that has been closely associated with dyslexia and is most commonly studied in the context of reading. Consequently, much of the existing research on Irlen syndrome and visual stress has focused on printed text, with relatively little attention given to how digital stimuli may affect individuals who experience visual stress. Chapter 8 demonstrated that digital stimuli can influence visual perception even in neurotypical participants. On this basis, it is plausible that individuals with Irlen syndrome may experience heightened visual stress when interacting with digital displays, although this interpretation remains speculative. This discrepancy highlights a gap in the current literature and suggests an avenue for further investigation.

In the previous chapter, a participant was not able to participate in the experiment due to Irlen Syndrome. Irlen Syndrome is a condition to which there is limited research. Hence, it served as an inspiration to do a case study on that participant to see how people with Irlen syndrome interact with digital stimuli. The purpose of this research is to study how fast an individual with Irlen Syndrome can react to a digital stimuli, with filters and without. The timing is then compared with the reaction time of healthy subjects.

9.2 Methods

This experiment builds on the experiment discussed in Chapter 8. As a reminder, The experiment consists of two parts, the first part being a simple reaction time (RT) experiment and the second part being a choice RT experiment. In the simple RT experiment, subjects will be shown a yellow cube, and the subjects are to press the bumper button on the controller when the cube appears. In the choice RT experiment, subjects will be shown either a red or a green cube. Subjects are to press the button only when a green cube is displayed.

In contrast with the experiment in Chapter 8, this experiment is conducted on a 2D screen, without the participant hooked up to an EEG machine, in order to reduce methodological complexity and improve interpretability in a single-participant case study. Findings from the previous chapter showed that while reaction time measures produced interpretable behavioural trends, EEG data did not yield consistent or reproducible neural signatures due to high inter-individual variability and extremely limited sample sizes. Including EEG in the present study would therefore add analytical burden without providing meaningful explanatory value. Moreover, restricting the experiment to a 2D display minimised potential confounds arising from immersive environments or recording equipment, which is particularly important when studying individuals with Irlen Syndrome who may be sensitive to visual and sensory load. Finally, the use of a standard screen enhances ecological relevance, as digital displays are the most common medium through which individuals with Irlen Syndrome encounter visual stimuli in everyday contexts.

The program used to generate the stimuli for the experiment in 8 was modified so that it does not look for TCP servers, and it allows for digital overlays to be placed on top of it. For this experiment, Colorveil is used to generate a screen overlay. It allows the researcher to choose the overlay colour that the participant prefers, and also another one that the participant don't. The preferred colour for the filter was based on the prescribed colour from the participant's Irlen Syndrome diagnosis. The non-preferred colour for the filter was chosen to be contrasting from the prescribed colour.

9.2.1 Data analysis

Statistical analysis was performed using IBM SPSS Statistics for Windows, version 27 (Armonk, NY: IBM Corp). In this study, we want to find out how having a filter can affect the reaction time of a participant with Irlen Syndrome reacting to a digital stimuli. Hence, we can perform an independent ANOVA, comparing the reaction time when the participants have different filter conditions.

9.3 Results

9.3.1 Simple reaction time experiment

Table 9.1 shows the descriptive statistics for the participant's reaction time in each filter condition (no filter, non-preferred filter, preferred filter), during the simple reaction time experiment:

Table 9.1: Table for descriptive statistics for participant's reaction time in each filter condition, during the simple reaction time experiment

	Mean	Sample size	Std. deviation
No filter	0.273	30	0.0325
Non-preferred filter	0.266	30	0.0331
Preferred filter	0.256	30	0.0320
Total	0.265	90	0.0329

Figure 9.2 shows the box plot of the reaction time recorded with each filter in the simple RT experiment.

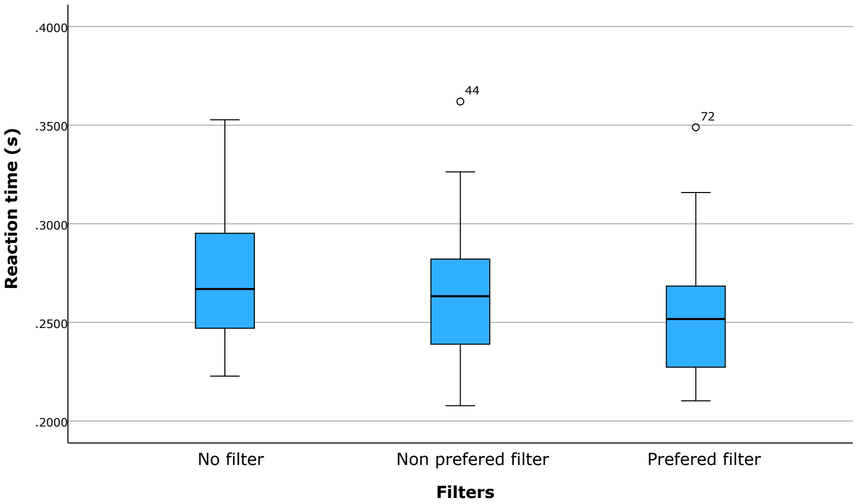


Figure 9.2: Box plot of simple reaction time experiment

From Figure 9.2, the order of the reaction time, from fastest to slowest, is as follows: preferred filter (0.256 ± 0.0320), non-preferred filter (0.266 ± 0.0331) and no filter (0.273 ± 0.0325). However, this order cannot be taken as statistically significant without a formal statistical test.

A one-way independent ANOVA was performed to analyze the effect of the participant's reaction time in each filter condition, during the simple reaction time experiment. Levene's test confirms that the assumption of homogeneity of variance had been met (F(2,87) = 0.035, p = 0.966). As such, Bonferroni was used for post hoc comparisons. Table 9.2 shows the results of the one-way independent ANOVA.

Table 9.2: Table for results of one-way independent ANOVA comparing the effect of participant's reaction time in each filter condition

	df	df error	F	Sig.	n_p^2
System	2	87	1.975	0.145	0.043

From Table 9.2, the main effect of filter condition was not statistically significant (F(2, 87) = 1.975, p = 0.145, $n_p^2 = 0.043$). Since the comparison was not statistically significant, no post-hoc test was performed. These results indicate that there is not enough evidence to prove that filter condition elicit differential changes in reaction time.

9.3.2 Choice reaction time experiment

Table 9.3 shows the descriptive statistics for the participant's reaction time in each filter condition (no filter, non-preferred filter, preferred filter), during the choice reaction time experiment:

Table 9.3: Table for descriptive statistics for participant's reaction time in each filter condition, during the choice reaction time experiment

	Mean	Sample size	Std. deviation
No filter	0.374	30	0.0515
Non-preferred filter	0.375	30	0.0516
Preferred filter	0.397	30	0.0573
Total	0.382	90	0.0540

Figure 9.3 shows the box plot of the reaction time recorded with each filter in the simple RT experiment.

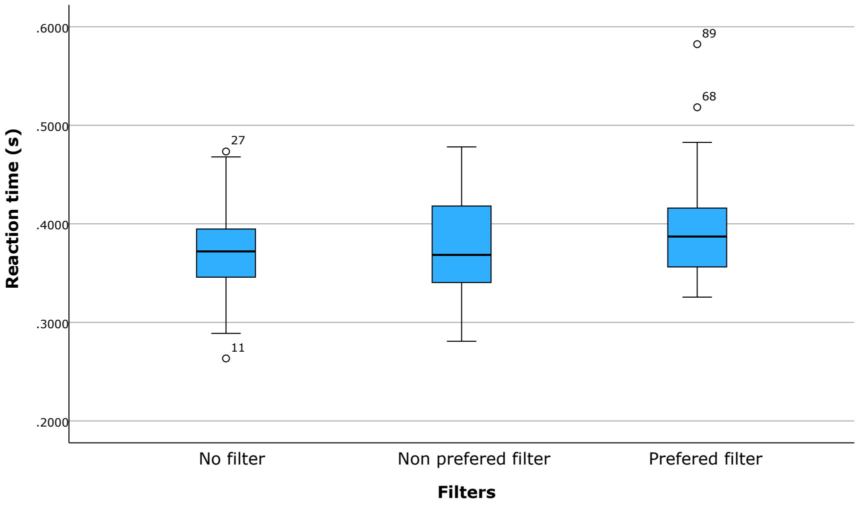


Figure 9.3: Box plot of choice reaction time experiment

From Figure 9.3, the order of the reaction time, from fastest to slowest, is as follows: no filter (0.374 ± 0.0515), non-preferred filter (0.375 ± 0.0516) and preferred filter (0.397 ± 0.0573). However, this order cannot be taken as statistically significant without a formal statistical test.

A one-way independent ANOVA was performed to analyze the effect of the participant's reaction time in each filter condition, during the choice reaction time experiment. Levene's test confirms that the assumption of homogeneity of variance had been met (F(2,87) = 0.197, p = 0.822). As such, Bonferroni was used for post hoc comparisons. Table 9.4 shows the results of the one-way independent ANOVA.

Table 9.4: Table for results of one-way independent ANOVA comparing the effect of participant's reaction time in each filter condition

	df	df error	F	Sig.	n_p^2
System	2	87	1.775	0.176	0.039

From Table 9.4, the main effect of filter condition was not statistically significant (F(2, 87) = 1.775, p = 0.176, $n_p^2 = 0.039$). Since the comparison was not statistically significant, no post-hoc test was performed. These results indicate that there is not enough evidence to prove that filter condition elicit differential changes in reaction time.

9.4 Discussion

The present case study examined whether digital colour filters influence reaction time in an individual with Irlen Syndrome during simple and choice reaction time tasks. While descriptive trends suggested faster responses when filters were applied, particularly when the participant's preferred filter was used, the repeated-measures ANOVA did not reveal a statistically significant main effect of filter condition for either task. These results indicate that, within the limits of this dataset, no reliable behavioural effect of filtering can be inferred.

Despite the absence of a significant main effect, the descriptive pattern observed in the simple reaction time task is noteworthy. Mean reaction times were lowest when the preferred filter was applied, followed by the non-preferred filter, and highest in the no-filter condition. In simple reaction time paradigms, where performance is primarily driven by perceptual detection and motor execution, it is plausible that reducing visual stress through filtering could facilitate more efficient stimulus processing. However, given the non-significant ANOVA result and the single-participant design, this pattern should be interpreted as exploratory rather than confirmatory.

In contrast, no consistent trend was observed in the choice reaction time task, and reaction times remained relatively stable across filter conditions. The additional cognitive demands associated with stimulus discrimination and response inhibition may have overshadowed any subtle perceptual benefits provided by the filters. This interpretation aligns with the notion that as task complexity increases, higher-level cognitive processes exert a stronger influence on reaction time than early-stage perceptual factors.

The interpretability of these findings is inherently limited by the use of a single participant. Although repeated-measures ANOVA was used to account for within-subject variability, inferential statistics derived from a single individual cannot be generalised to the broader Irlen Syndrome population. Moreover, the statistical power to detect small effects is limited, increasing the likelihood of Type II error. As such, the absence of a significant main effect does not necessarily imply the absence of a true underlying effect, but rather reflects the constraints of the study design.

Overall, the findings suggest that while digital colour filtering may influence reaction time at a descriptive level, particularly in low cognitive load tasks, robust conclusions cannot be drawn without larger samples and more comprehensive experimental designs. The present study should therefore be viewed as a preliminary exploration that motivates further investigation rather than as definitive evidence for the efficacy of digital filters in improving reaction time performance.

9.5 Conclusion

This chapter presented an exploratory behavioural case study investigating the effect of digital colour filters on reaction time in an individual with Irlen Syndrome. Using simple and choice reaction time tasks, the study examined whether screen-based filtering influenced performance under varying levels of cognitive demand. Although descriptive trends suggested faster responses in the simple reaction time task when filters were applied, particularly when the participant's preferred filter was used, statistical analysis using repeated-measures ANOVA did not reveal a significant main effect of filter condition.

The absence of a statistically significant effect highlights the limitations inherent in single-participant designs, including reduced statistical power and restricted generalisability. As such, the findings should be interpreted as descriptive and hypothesis-generating rather than as evidence for a reliable behavioural effect of digital filtering. The lack of consistent effects in the choice reaction time task further suggests that any potential perceptual benefits of filtering may be diminished when additional cognitive processing is required.

Despite these limitations, the study contributes to the limited literature on Irlen Syndrome and digital stimuli by demonstrating the feasibility of controlled reaction time paradigms for investigating visual stress in digital environments. The results underscore the importance of considering task demands, cognitive load, and individual variability when evaluating the effectiveness of visual interventions. Future research should extend this work to larger participant samples, incorporate more ecologically valid stimuli, and examine the interaction between perceptual and cognitive factors in digital visual processing.

Chapter 10

Concluding Remarks

10.1 Summary

This thesis investigated visual perception across real, digital, and immersive environments by integrating foundational vision science, EEG signal processing, virtual and augmented reality technologies, and behavioural experimentation. Beginning with the basic physics of light and the neurobiology of vision, the early chapters established the theoretical framework necessary to understand how visual information is transduced and processed by the human brain. This foundation was extended through a focused review of visual word processing, highlighting the interaction between low-level sensory mechanisms and higher-order cognitive functions such as reading.

Building on this theoretical background, the thesis developed and evaluated a signal processing framework for analysing EEG data. A comparative examination of time–frequency methods demonstrated the suitability of wavelet-based approaches for non-stationary neural signals, while spatial filtering using Common Spatial Patterns (CSP) was explored for enhancing task-related contrasts. These techniques were integrated into a generalised EEG analysis pipeline, which was subsequently validated using a public motor imagery dataset. The results showed that the pipeline reliably captured physiologically plausible task-related neural activity, including event-related desynchronisation in sensorimotor rhythms, while also highlighting the challenges of detecting robust hemispheric lateralisation at the scalp EEG level.

A central contribution of this thesis lies in the systematic development and validation of an experimental platform for studying visual perception across 2D screens, real-world settings, augmented reality, and virtual reality. Hardware characterisation experiments quantified key confounds associated with these environments, including luminance, latency, and electromagnetic noise. The findings demonstrated that while system latency was minimal and consistent across environments, luminance and device-specific noise varied substantially and represent systematic factors that must be accounted for when interpreting behavioural and EEG results.

Using this validated platform, behavioural and EEG experiments were conducted to examine how stimulus presentation environment influences perceptual processing. Reaction time results suggested a tendency towards slower responses in immersive virtual environments compared to real-world and 2D conditions, although these effects were modest and constrained by limited sample sizes. EEG analyses, necessarily descriptive due to the small number of participants, revealed substantial inter-individual variability and no consistent group-level spatial patterns, underscoring the challenges of drawing robust neural inferences in small-sample immersive EEG studies.

Finally, an exploratory behavioural case study examined the effects of digital colour filtering on reaction time in an individual with Irlen Syndrome. While no statistically significant effects were observed, descriptive trends in simple reaction time tasks suggested that filtering may reduce visual stress under low cognitive load. Although preliminary, this work contributes to a limited body of research on visual stress in digital contexts and motivates further investigation with larger samples and more ecologically valid stimuli.

10.2 Future directions

Despite the breadth of topics addressed, this thesis represents only a small contribution to the broader field of visual neuroscience. Several avenues for future research emerge naturally from the findings and limitations identified throughout this work.

From a signal processing perspective, future studies could enhance the proposed EEG pipeline by incorporating machine learning techniques, such as regularised classifiers, deep learning models, or subject-specific approaches. These methods may improve sensitivity to subtle condition differences and help address the substantial inter-individual variability observed in EEG data. Combining spectral, spatial, and temporal features, or integrating multimodal data, may further improve classification and interpretability.

In terms of immersive technologies, the stimulus generation framework developed in this thesis could be expanded and made more widely accessible, enabling broader adoption and collaborative development. Future experiments could employ more ecologically valid and semantically rich stimuli, such as text, naturalistic scenes, or interactive tasks, to better capture real-world perceptual demands.

The behavioural and EEG experiments conducted in this thesis were limited by small participant samples. Increasing sample sizes would improve statistical power and generalisability and enable more sophisticated statistical modelling, including mixed-effects approaches that explicitly account for individual differences. Larger datasets would also allow for more robust investigation of neural signatures associated with perception in immersive environments.

With respect to Irlen Syndrome, substantial future work is required to move beyond exploratory case studies. Recruiting larger cohorts of individuals with Irlen Syndrome and neurotypical controls, incorporating reading-based and linguistically meaningful stimuli, and combining behavioural and neural measures may help clarify the mechanisms underlying visual stress and the potential efficacy of visual interventions.

Overall, this thesis demonstrates the potential of interdisciplinary approaches that combine neuroscience, signal processing, and immersive technologies to advance the study of visual perception. It is hoped that this work will motivate further research into perception in digital environments and contribute to the development of evidence-based tools and interventions that support individuals in their interaction with modern visual technologies.

BIBLIOGRAPHY

[1] G. Stark, *light*, Nov. 2022.

[2] S. B. McKagan, W. Handley, K. K. Perkins, and C. E. Wieman, "A research-based curriculum for teaching the photoelectric effect," *Am. J. Phys.*, vol. 77, no. 1, pp. 87–94, Jan. 2009.

[3] Y. A. Garbovskiy and A. V. Glushchenko, "Basic laws of geometrical optics: Experimental verification," in *A Practical Guide to Experimental Geometrical Optics.* Cambridge University Press, Dec. 2017, pp. 50–61.

[4] "2.5: Thin lenses," https://phys.libretexts.org/Bookshelves/University_Physics/Book%3A_University_Physics_(OpenStax)/University_Physics_III_-_Optics_and_Modern_Physics_(OpenStax)/02%3A_Geometric_Optics_and_Image_Formation/2.05%3A_Thin_Lenses, Nov. 2016, accessed: 2023-3-6.

[5] R. Snowden, R. J. Snowden, P. Thompson, and T. Troscianko, *Basic Vision: An Introduction to Visual Perception.* OUP Oxford, Feb. 2012.

[6] E. D. Reichle, K. Rayner, and A. Pollatsek, "The E-Z reader model of eye-movement control in reading: comparisons to other models," *Behav. Brain Sci.*, vol. 26, no. 4, pp. 445–76; discussion 477–526, Aug. 2003.

[7] W. Kintsch, "The role of knowledge in discourse comprehension: a construction-integration model," *Psychol. Rev.*, vol. 95, no. 2, pp. 163–182, Apr. 1988.

[8] D. Norris, "Models of visual word recognition," *Trends Cogn. Sci.*, vol. 17, no. 10, pp. 517–524, Oct. 2013.

[9] D. M. Kearns, R. Hancock, F. Hoeft, K. R. Pugh, and S. J. Frost, "The neurobiology of dyslexia," *Teach. Except. Child.*, vol. 51, no. 3, pp. 175–188, Jan. 2019.

[10] A. Bruns, "Fourier-, hilbert- and wavelet-based signal analysis: are they really different approaches?" *J. Neurosci. Methods*, vol. 137, no. 2, pp. 321–332, Aug. 2004.

[11] E. Velarde, Y. Márquez-Bocalandro, and E. Martínez-Montes, "EEG wireless recording and multidimensional analysis," in *Distributed diagnosis and*

Home Healthcare (D2H2), Y. K. Eddie Ng, A. U. Rajendra, and T. Tamura, Eds. American Scientific Publishers, 2012, pp. 1–16.

[12] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, p. 156869, 2011.

[13] M. Jas, D. Engemann, F. Raimondo, Y. Bekhti, and A. Gramfort, "Automated rejection and repair of bad trials in MEG/EEG," in *2016 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*. IEEE, Jun. 2016.

[14] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "EEG motor movement/imagery dataset," 2009.

[15] B. G. Cumming and G. C. DeAngelis, "The physiology of stereopsis," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 203–238, 2001.

[16] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Trans. Inf. Syst.*, vol. E77-D, no. 12, pp. 1321–1329, Dec. 1994.

[17] "Reading music," http://www.irlenny.com/reading-music.html, accessed: 2026-1-25.

[18] "Introduction to perception," in *Sensation and Perception*, 10th ed., E. Goldstein and J. Brockmole, Eds. Mason, OH: CENGAGE Learning Custom Publishing, Feb. 2016.

[19] "Vision: The eye," in *Neuroscience*, 3rd ed., G. J. Augustine, D. Fitzpatrick, and D. Purves, Eds. Sunderland, MA: Sinauer Associates, Jul. 2004, pp. 229–258.

[20] "Central visual pathways," in *Neuroscience*, 3rd ed., G. J. Augustine, D. Fitzpatrick, and D. Purves, Eds. Sunderland, MA: Sinauer Associates, Jul. 2004, pp. 259–282.

[21] S. Bracci and H. P. Op de Beeck, "Understanding human object vision: A picture is worth a thousand representations," *Annu. Rev. Psychol.*, vol. 74, no. 1, pp. 113–135, Jan. 2023.

[22] B. de Haas, D. S. Schwarzkopf, I. Alvarez, R. P. Lawson, L. Henriksson, N. Kriegeskorte, and G. Rees, "Perception and processing of faces in the human brain is tuned to typical feature locations," *J. Neurosci.*, vol. 36, no. 36, pp. 9289–9302, Sep. 2016.

[23] V. Ayzenberg and M. Behrmann, "The dorsal visual pathway represents object-centered spatial relations for object recognition," *J. Neurosci.*, vol. 42, no. 23, pp. 4693–4710, Jun. 2022.

[24] L. M. Padilla, S. H. Creem-Regehr, M. Hegarty, and J. K. Stefanucci, “Decision making with visualizations: a cognitive framework across disciplines,” *Cogn. Res. Princ. Implic.*, vol. 3, no. 1, Dec. 2018.

[25] D. P. McCabe and A. D. Castel, “Seeing is believing: the effect of brain images on judgments of scientific reasoning,” *Cognition*, vol. 107, no. 1, pp. 343–352, Apr. 2008.

[26] P. V. González-Erena, S. Fernández-Guinea, and P. Kourtesis, “Cognitive assessment and training in extended reality: Multimodal systems, clinical utility, and current challenges,” *arXiv [cs.HC]*, 2025.

[27] A. D. Zhdanov, D. D. Zhdanov, N. N. Bogdanov, I. S. Potemin, V. A. Galaktionov, and M. I. Sorokin, “Discomfort of visual perception in virtual and mixed reality systems,” *Program. Comput. Softw.*, vol. 45, no. 4, pp. 147–155, Jul. 2019.

[28] C. Huisingh, E. B. Levitan, M. R. Irvin, P. MacLennan, V. Wadley, and C. Owsley, “Visual sensory and visual-cognitive function and rate of crash and near-crash involvement among older drivers using naturalistic driving data,” *Invest. Ophthalmol. Vis. Sci.*, vol. 58, no. 7, pp. 2959–2967, Jun. 2017.

[29] T. Cvahte Ojsteršek and D. Topolšek, “Influence of drivers’ visual and cognitive attention on their perception of changes in the traffic environment,” *Eur. Transp. Res. Rev.*, vol. 11, no. 1, Dec. 2019.

[30] L. Broeker, M. Haeger, O. Bock, B. Kretschmann, H. Ewolds, S. Künzell, and M. Raab, “How visual information influences dual-task driving and tracking,” *Exp. Brain Res.*, vol. 238, no. 3, pp. 675–687, Mar. 2020.

[31] M. Gori and G. Sandini, “Multisensory perception and action in painting: science, creativity, and technology,” *Front. Psychol.*, vol. 16, no. 1630107, p. 1630107, Aug. 2025.

[32] J. Misersky, D. Peeters, and M. Flecken, “The potential of immersive virtual reality for the study of event perception,” *Front. Virtual Real.*, vol. 3, no. 697934, Jul. 2022.

[33] D. M. Levi, “Applications and implications for extended reality to improve binocular vision and stereopsis,” *J. Vis.*, vol. 23, no. 1, p. 14, Jan. 2023.

[34] H. Shin, J. Lee, G. Lee, J. Park, and S. Lee, “Effect of vibrotactile feedback and perception induced by transcranial magnetic stimulation over the primary motor cortex in virtual hand illusion,” *Virtual Real.*, vol. 28, no. 4, Sep. 2024.

[35] W. Wu, Z. Zhao, A. Du, and J. Lin, “Effects of multisensory integration through spherical video-based immersive virtual reality on students’ learning performances in a landscape architecture conservation course,” *Sustainability*, vol. 14, no. 24, p. 16891, Dec. 2022.

[36] D. Burke, H. Crompton, and C. Nickel, “The use of extended reality (XR) in higher education: A systematic review,” *TechTrends*, vol. 69, no. 5, pp. 998–1011, Sep. 2025.

[37] M. Vasarainen, S. Paavola, and L. Vetoshkina, “A systematic literature review on extended reality: Virtual, augmented and mixed reality in working life,” *International Journal of Virtual Reality*, vol. 21, no. 2, pp. 1–28, Oct. 2021.

[38] H. U. Khan, Y. Ali, F. Khan, and M. A. Al-Antari, “A comprehensive study on unraveling the advances of immersive technologies (VR/AR/MR/XR) in the healthcare sector during the COVID-19: Challenges and solutions,” *Heliyon*, vol. 10, no. 15, p. e35037, Aug. 2024.

[39] C.-H. Lai, C.-C. Chen, and S.-M. Wu, “Analysis of key factors for XR extended reality immersive art experience,” *International Journal of Social Sciences and Artistic Innovations*, vol. 3, no. 1, pp. 24–36, Mar. 2023.

[40] M. Perdices and D. A. Cooper, “Simple and choice reaction time in patients with human immunodeficiency virus infection,” *Ann. Neurol.*, vol. 25, no. 5, pp. 460–467, May 1989.

[41] A. D. Goswami, J. Roy, A. Naskar, and S. Chakraborty, “A laboratory based study on the effect of CCT change of LED light sources on reaction time and visibility level for object recognition,” *Optik*, vol. 264, p. 169353, Aug. 2022.

[42] C. Perry and H. Long, “What is going on with visual attention in reading and dyslexia? a critical review of recent studies,” *Brain Sci.*, vol. 12, no. 1, p. 87, Jan. 2022.

[43] Kristjánsson and H. M. Sigurdardottir, “The role of visual factors in dyslexia,” *J. Cogn.*, vol. 6, no. 1, p. 31, Jun. 2023.

[44] M. J. Farah, *Visual Agnosia, Second Edition*, 2nd ed., ser. Bradford Books. Bradford Book, May 2014.

[45] F. A. Lygo, B. Richard, A. R. Wade, A. B. Morland, and D. H. Baker, “Neural markers of suppression in impaired binocular vision,” *Neuroimage*, vol. 230, no. 117780, p. 117780, Apr. 2021.

[46] H. M. Sigurdardottir, H. R. Omarsdottir, and A. S. Valgeirsdottir, “Reading problems and their connection with visual search and attention,” *Dyslexia*, vol. 30, no. 2, p. e1764, May 2024.

[47] M. A. Alanazi, S. A. Alanazi, and U. L. Osuagwu, “Evaluation of visual stress symptoms in age-matched dyslexic, meares-irlen syndrome and normal adults,” *Int. J. Ophthalmol.*, vol. 9, no. 4, pp. 617–624, Apr. 2016.

[48] E. Niedermeyer and F. H. Lopes Da Silva, Eds., *Electroencephalography*, 5th ed. Philadelphia, PA: Lippincott Williams and Wilkins, Nov. 2004.

[49] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional magnetic resonance imaging*, 3rd ed. New York, NY: Oxford University Press, Sep. 2014.

[50] M. Murugappan, “Combining spatial filtering and wavelet transform for classifying human emotions using EEG signals,” *J. Med. Biol. Eng.*, vol. 31, no. 1, pp. 45–51, 2011.

[51] S. I. Newton, “Opticks. or a treatise of the reflections, refractions, inflections and of light,” *Optom. Vis. Sci.*, vol. 10, no. 5, p. 190, May 1933.

[52] A. Einstein, “Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt,” *Ann. Phys.*, vol. 322, no. 6, pp. 132–148, Jan. 1905.

[53] H. Hertz, “Ueber einen einfluss des ultravioletten lichtes auf die electrische entladung,” *Ann. Phys.*, vol. 267, no. 8, pp. 983–1000, Jan. 1887.

[54] T. V. Chirila and Y. Hong, “Chapter C2 the vitreous humor,” in *Handbook of Biomaterial Properties*, W. Murphy, J. Black, and G. Hastings, Eds. New York, NY: Springer New York, 2016, pp. 125–134.

[55] M. Ghodrati, S.-M. Khaligh-Razavi, and S. R. Lehky, “Towards building a more complex view of the lateral geniculate nucleus: Recent advances in understanding its role,” *Prog. Neurobiol.*, vol. 156, pp. 214–255, Sep. 2017.

[56] P. H. Schiller, “Central connections of the retinal ON and OFF pathways,” *Nature*, vol. 297, no. 5867, pp. 580–583, Jun. 1982.

[57] D. H. Hubel and T. N. Wiesel, “Receptive fields and functional architecture of monkey striate cortex,” *J. Physiol.*, vol. 195, no. 1, pp. 215–243, Mar. 1968.

[58] R. B. Tootell, N. K. Hadjikhani, W. Vanduffel, A. K. Liu, J. D. Mendola, M. I. Sereno, and A. M. Dale, “Functional analysis of primary visual cortex (V1) in humans,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 3, pp. 811–817, Feb. 1998.

[59] V. A. Lamme, H. Supèr, R. Landman, P. R. Roelfsema, and H. Spekreijse, "The role of primary visual cortex (V1) in visual awareness," *Vision Res.*, vol. 40, no. 10-12, pp. 1507–1521, 2000.

[60] A. Anzai, X. Peng, and D. C. Van Essen, "Neurons in monkey visual area V2 encode combinations of orientations," *Nat. Neurosci.*, vol. 10, no. 10, pp. 1313–1321, Oct. 2007.

[61] T. Huff, N. Mahabadi, and P. Tadi, "Neuroanatomy, visual cortex," in *StatPearls*. Treasure Island (FL): StatPearls Publishing, Aug. 2023.

[62] M. G. P. Rosa and P. R. Manger, "Clarifying homologies in the mammalian cerebral cortex: the case of the third visual area (V3)," *Clin. Exp. Pharmacol. Physiol.*, vol. 32, no. 5-6, pp. 327–339, 2005.

[63] H. Ikei, H. Jo, and Y. Miyazaki, "Physiological effects of visual stimulation by a japanese low wooden table: A crossover field experiment," *Int. J. Environ. Res. Public Health*, vol. 20, no. 14, Jul. 2023.

[64] A. J. Levi, Y. Zhao, I. M. Park, and A. C. Huk, "Sensory and choice responses in MT distinct from motion encoding," *J. Neurosci.*, vol. 43, no. 12, pp. 2090–2103, Mar. 2023.

[65] P. P. Thakral and S. D. Slotnick, "Disruption of MT impairs motion processing," *Neurosci. Lett.*, vol. 490, no. 3, pp. 226–230, Mar. 2011.

[66] L. D. Liu and C. C. Pack, "The contribution of area MT to visual motion perception depends on training," *Neuron*, vol. 95, no. 2, pp. 436–446.e3, Jul. 2017.

[67] A. Tosoni, R. Guidotti, C. Del Gratta, G. Committeri, and C. Sestieri, "Preferential coding of eye/hand motor actions in the human ventral occipito-temporal cortex," *Neuropsychologia*, vol. 93, no. Pt A, pp. 116–127, Dec. 2016.

[68] D. J. McKeefry, M. P. Burton, C. Vakrou, B. T. Barrett, and A. B. Morland, "Induced deficits in speed perception by transcranial magnetic stimulation of human cortical areas V5/MT+ and V3A," *J. Neurosci.*, vol. 28, no. 27, pp. 6848–6857, Jul. 2008.

[69] S. Pitzalis, M. I. Sereno, G. Committeri, P. Fattori, G. Galati, F. Patria, and C. Galletti, "Human v6: the medial motion area," *Cereb. Cortex*, vol. 20, no. 2, pp. 411–424, Feb. 2010.

[70] D. L. Woods, J. M. Wyma, E. W. Yund, T. J. Herron, and B. Reed, "Factors influencing the latency of simple reaction time," *Front. Hum. Neurosci.*, vol. 9, p. 131, Mar. 2015.

[71] “Types of eye movements and their functions,” in *Neuroscience*, D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, and S. Mark Williams, Eds. Sunderland, MA: Sinauer Associates, Jan. 2001.

[72] S. Genc, E. P. Raven, M. Drakesmith, S.-J. Blakemore, and D. K. Jones, “Novel insights into axon diameter and myelin content in late childhood and adolescence,” *Cereb. Cortex*, vol. 33, no. 10, pp. 6435–6448, May 2023.

[73] P. H. Lu, G. J. Lee, E. P. Raven, K. Tingus, T. Khoo, P. M. Thompson, and G. Bartzokis, “Age-related slowing in cognitive processing speed is associated with myelin integrity in a very healthy elderly sample,” *J. Clin. Exp. Neuropsychol.*, vol. 33, no. 10, pp. 1059–1068, Dec. 2011.

[74] E. Sarrias-Arrabal, S. Eichau, A. Galvao-Carmona, E. Domínguez, G. Izquierdo, and M. Vázquez-Marrufo, “Deficits in early sensory and cognitive processing are related to phase and nonphase EEG activity in multiple sclerosis patients,” *Brain Sci.*, vol. 11, no. 5, p. 629, May 2021.

[75] A. J. Hughes, D. R. Denney, and S. G. Lynch, “Reaction time and rapid serial processing measures of information processing speed in multiple sclerosis: complexity, compounding, and augmentation,” *J. Int. Neuropsychol. Soc.*, vol. 17, no. 6, pp. 1113–1121, Nov. 2011.

[76] I. Rudvin, A. Valberg, and B. E. Kilavik, “Visual evoked potentials and magnocellular and parvocellular segregation,” *Vis. Neurosci.*, vol. 17, no. 4, pp. 579–590, Jul. 2000.

[77] A. J. Ries and J. B. Hopfinger, “Magnocellular and parvocellular influences on reflexive attention,” *Vision Res.*, vol. 51, no. 16, pp. 1820–1828, Aug. 2011.

[78] M. T. Schmolesky, Y. Wang, D. P. Hanes, K. G. Thompson, S. Leutgeb, J. D. Schall, and A. G. Leventhal, “Signal timing across the macaque visual system,” *J. Neurophysiol.*, vol. 79, no. 6, pp. 3272–3278, Jun. 1998.

[79] M. W. Self and P. R. Roelfsema, “Scene perception in early vision: Figure-ground organization in the lateral geniculate nucleus,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 112, no. 22, pp. 6784–6785, Jun. 2015.

[80] P. Blanpied and H. Oksendahl, “Reaction times and electromechanical delay in reactions of increasing and decreasing force,” *Percept. Mot. Skills*, vol. 103, no. 3, pp. 743–754, Dec. 2006.

[81] N. A. Maffiuletti, P. Aagaard, A. J. Blazevich, J. Folland, N. Tillin, and J. Duchateau, “Rate of force development: physiological and methodologi-

cal considerations," *Eur. J. Appl. Physiol.*, vol. 116, no. 6, pp. 1091–1116, Jun. 2016.

[82] I. Bautmans, S. Vantieghem, E. Gorus, Y.-R. Grazzini, Y. Fierens, A. Pool-Goudzwaard, and T. Mets, "Age-related differences in pre-movement antagonist muscle co-activation and reaction-time performance," *Exp. Gerontol.*, vol. 46, no. 8, pp. 637–642, Aug. 2011.

[83] A. Szpala and A. Rutkowska-Kucharska, "Electromechanical response times in the knee muscles in young and old women," *Muscle Nerve*, vol. 56, no. 6, pp. E147–E153, Dec. 2017.

[84] J. D. Jiménez-García, A. Martínez-Amat, F. Hita-Contreras, R. Fábrega-Cuadros, F. Álvarez Salvago, and A. Aibar-Almazán, "Muscle strength and physical performance are associated with reaction time performance in older people," *Int. J. Environ. Res. Public Health*, vol. 18, no. 11, p. 5893, May 2021.

[85] L. van Maanen, O. Portoles, and J. P. Borst, "The discovery and interpretation of evidence accumulation stages," *Comput. Brain Behav.*, vol. 4, no. 4, pp. 395–415, Dec. 2021.

[86] D. Rangelov, J. Fellrath, and J. B. Mattingley, "Integrated perceptual decisions rely on parallel evidence accumulation," *J. Neurosci.*, vol. 44, no. 33, p. e2368232024, Aug. 2024.

[87] A. Gupta, R. Bansal, H. Alashwal, A. S. Kacar, F. Balci, and A. A. Moustafa, "Neural substrates of the drift-diffusion model in brain disorders," *Front. Comput. Neurosci.*, vol. 15, p. 678232, 2021.

[88] R. M. Yerkes and J. D. Dodson, "The relation of strength of stimulus to rapidity of habit-formation," *J. Comp. Neurol. Psychol.*, vol. 18, no. 5, pp. 459–482, Nov. 1908.

[89] R. Ulrich and S. Mattes, "Does immediate arousal enhance response force in simple reaction time?" *Q. J. Exp. Psychol. A*, vol. 49, no. 4, pp. 972–990, Nov. 1996.

[90] C. W. N. Saville, H. M. de Morree, N. M. Dundon, S. M. Marcora, and C. Klein, "Effects of caffeine on reaction time are mediated by attentional rather than motor processes," *Psychopharmacology (Berl.)*, vol. 235, no. 3, pp. 749–759, Mar. 2018.

[91] T. P. Doyle, R. S. Lutz, J. K. Pellegrino, D. J. Sanders, and S. M. Arent, "The effects of caffeine on arousal, response time, accuracy, and performance in division I collegiate fencers," *J. Strength Cond. Res.*, vol. 30, no. 11, pp. 3228–3235, Nov. 2016.

[92] C. Sanchis, E. Blasco, F. G. Luna, and J. Lupiáñez, “Effects of caffeine intake and exercise intensity on executive and arousal vigilance,” *Sci. Rep.*, vol. 10, no. 1, p. 8393, May 2020.

[93] A. J. Wawrzyniak, M. Hamer, A. Steptoe, and R. Endrighi, “Decreased reaction time variability is associated with greater cardiovascular responses to acute stress,” *Psychophysiology*, vol. 53, no. 5, pp. 739–748, May 2016.

[94] C. Stone, L. Ney, K. Felmingham, D. Nichols, and A. Matthews, “The effects of acute stress on attentional networks and working memory in females,” *Physiol. Behav.*, vol. 242, no. 113602, p. 113602, Dec. 2021.

[95] I. Lorenzo, J. Ramos, C. Arce, M. A. Guevara, and M. Corsi-Cabrera, “Effect of total sleep deprivation on reaction time and waking EEG activity in man,” *Sleep*, vol. 18, no. 5, pp. 346–354, Jun. 1995.

[96] M. Taheri and E. Arabameri, “The effect of sleep deprivation on choice reaction time and anaerobic power of college student athletes,” *Asian J. Sports Med.*, vol. 3, no. 1, pp. 15–20, Mar. 2012.

[97] J. van den Berg and G. Neely, “Performance on a simple reaction time task while sleep deprived,” *Percept. Mot. Skills*, vol. 102, no. 2, pp. 589–599, Apr. 2006.

[98] G. Rohenkohl, I. C. Gould, J. Pessoa, and A. C. Nobre, “Combining spatial and temporal expectations to improve visual perception,” *J. Vis.*, vol. 14, no. 4, Apr. 2014.

[99] T. Ghuntla, H. Mehta, P. Gokhale, and C. Shah, “Influence of practice on visual reaction time,” *J. Mahatma Gandhi Inst. Med. Sci.*, vol. 19, no. 2, p. 119, 2014.

[100] G. Del Rossi, A. Malaguti, and S. Del Rossi, “Practice effects associated with repeated assessment of a clinical test of reaction time,” *J. Athl. Train.*, vol. 49, no. 3, pp. 356–359, May 2014.

[101] A. Reinhartz, T. Strobach, T. Jacobsen, and C. C. von Bastian, “Mechanisms of training-related change in processing speed: A drift-diffusion model approach,” *J. Cogn.*, vol. 6, no. 1, p. 46, Aug. 2023.

[102] A. Voormann and J. Miller, “Sequential effects on reaction time distributions: Commonalities and differences across paradigms,” *J. Cogn.*, vol. 7, no. 1, p. 68, Sep. 2024.

[103] B. E. Jones, “Modulation of cortical activation and behavioral arousal by cholinergic and orexinergic systems,” *Ann. N. Y. Acad. Sci.*, vol. 1129, no. 1, pp. 26–34, 2008.

[104] R. Datta and E. A. DeYoe, “I know where you are secretly attending! the topography of human visual attention revealed with fMRI,” *Vision Res.*, vol. 49, no. 10, pp. 1037–1044, Jun. 2009.

[105] I. Taylor and M. M. Taylor, *Psychology of reading.* San Diego, CA: Academic Press, Nov. 1983.

[106] K. Rayner, “Eye movements in reading and information processing,” *Psychol. Bull.*, vol. 85, no. 3, pp. 618–660, May 1978.

[107] ——, “Eye movements in reading and information processing: 20 years of research,” *Psychol. Bull.*, vol. 124, no. 3, pp. 372–422, 1998.

[108] R. E. Morrison, “Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades,” *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 10, no. 5, pp. 667–682, 1984.

[109] K. Apel, V. S. Henbest, and J. Masterson, “Orthographic knowledge: clarifications, challenges, and future directions,” *Read. Writ.*, vol. 32, no. 4, pp. 873–889, Apr. 2019.

[110] M. van den Boer and P. F. de Jong, “Parallel and serial reading processes in children’s word and nonword reading,” *J. Educ. Psychol.*, vol. 107, no. 1, pp. 141–151, Feb. 2015.

[111] C. A. Becker, “Semantic context and word frequency effects in visual word recognition,” *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 5, no. 2, pp. 252–259, May 1979.

[112] J. Davies-Thompson, S. Johnston, Y. Tashakkor, R. Pancaroglu, and J. J. S. Barton, “The relationship between visual word and face processing lateralization in the fusiform gyri: A cross-sectional study,” *Brain Res.*, vol. 1644, pp. 88–97, Aug. 2016.

[113] A. Buchweitz, R. A. Mason, L. M. B. Tomitch, and M. A. Just, “Brain activation for reading and listening comprehension: An fMRI study of modality effects and individual differences in language comprehension,” *Psychol. Neurosci.*, vol. 2, no. 2, pp. 111–123, 2009.

[114] R. Staples and W. W. Graves, “Neural components of reading revealed by distributed and symbolic computational models,” *Neurobiol. Lang. (Camb.)*, vol. 1, no. 4, pp. 381–401, Oct. 2020.

[115] P. Solana and J. Santiago, “Does the involvement of motor cortex in embodied language comprehension stand on solid ground? a p-curve analysis and test for excess significance of the TMS and tDCS evidence,” *Neurosci. Biobehav. Rev.*, vol. 141, no. 104834, p. 104834, Oct. 2022.

[116] J. V. Dickens, M. E. Fama, A. T. DeMarco, E. H. Lacey, R. B. Friedman, and P. E. Turkeltaub, “Localization of phonological and semantic contributions to reading,” *J. Neurosci.*, vol. 39, no. 27, pp. 5361–5368, Jul. 2019.

[117] J. Glazzard, “The impact of dyslexia on pupils’ self-esteem,” *Support Learn.*, vol. 25, no. 2, pp. 63–69, May 2010.

[118] British Dyslexia Association, “What is dyslexia?” https://www.bdadyslexia.org.uk/dyslexia/about-dyslexia/what-is-dyslexia, accessed: 2024-11-15.

[119] S. Heim, J. Tschierse, K. Amunts, M. Wilms, S. Vossel, K. Willmes, A. Grabowska, and W. Huber, “Cognitive subtypes of dyslexia,” *Acta Neurobiol. Exp. (Wars.)*, vol. 68, no. 1, pp. 73–82, 2008.

[120] P. M. Paz-Alonso, M. Oliver, G. Lerma-Usabiaga, C. Caballero-Gaudes, I. Quiñones, P. Suárez-Coalla, J. A. Duñabeitia, F. Cuetos, and M. Carreiras, “Neural correlates of phonological, orthographic and semantic reading processing in dyslexia,” *NeuroImage Clin.*, vol. 20, pp. 433–447, Aug. 2018.

[121] S. A. Ali, N. A. Fadzil, F. Reza, F. Mustafar, and T. Begum, “A mini review: Visual and auditory perception in dyslexia,” *Int. Med. J. Malays.*, vol. 20, no. 4, Sep. 2021.

[122] B. D. Chouinard, C. I. Zhou, S. Hrybouski, E. S. Kim, and J. Cummine, “A functional neuroimaging case study of meares-irlen syndrome/visual stress (MISViS),” *Brain Topogr.*, vol. 25, no. 3, pp. 293–307, Jul. 2012.

[123] A. Martin, M. Schurz, M. Kronbichler, and F. Richlan, “Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies,” *Hum. Brain Mapp.*, vol. 36, no. 5, pp. 1963–1981, May 2015.

[124] F. Richlan, M. Kronbichler, and H. Wimmer, “Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies,” *Hum. Brain Mapp.*, vol. 34, no. 11, pp. 3055–3065, Nov. 2013.

[125] R. Hancock, F. Richlan, and F. Hoeft, “Possible roles for fronto-striatal circuits in reading disorder,” *Neurosci. Biobehav. Rev.*, vol. 72, pp. 243–260, Jan. 2017.

[126] F. Richlan, M. Kronbichler, and H. Wimmer, “Meta-analyzing brain dysfunctions in dyslexic children and adults,” *Neuroimage*, vol. 56, no. 3, pp. 1735–1742, Jun. 2011.

[127] A. Meyler, T. A. Keller, V. L. Cherkassky, J. D. E. Gabrieli, and M. A. Just, "Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: a longitudinal study of neuroplasticity," *Neuropsychologia*, vol. 46, no. 10, pp. 2580–2592, Aug. 2008.

[128] I. Kriss and B. J. W. Evans, "The relationship between dyslexia and Meares-Irlen syndrome," *J. Res. Read.*, vol. 28, no. 3, pp. 350–364, Aug. 2005.

[129] P. G. Griffiths, R. H. Taylor, L. M. Henderson, and B. T. Barrett, "The effect of coloured overlays and lenses on reading: a systematic review of the literature," *Ophthalmic Physiol. Opt.*, vol. 36, no. 5, pp. 519–544, Sep. 2016.

[130] J. D. S. Miyasaka, R. V. G. Vieira, E. S. Novalo-Goto, E. Montagna, and R. Wajnsztejn, "Irlen syndrome: systematic review and level of evidence analysis," *Arq. Neuropsiquiatr.*, vol. 77, no. 3, pp. 194–207, Mar. 2019.

[131] D. Gabor, "Theory of communication," *J. Inst. Electr. Eng. - Gen.*, vol. 94, no. 73, pp. 58–58, Jan. 1947.

[132] L. Cohen, *Time-frequency analysis.* Old Tappan, NJ: Prentice Hall, 1995.

[133] H.-H. Wu and S. Wu, "Various proofs of the cauchy-schwarz inequality," 2009.

[134] "What is a spectrogram?" https://pnsn.org/spectrograms/what-is-a-spectrogram, accessed: 2024-10-10.

[135] C. Tallon-Baudry and O. Bertrand, "Oscillatory gamma activity in humans and its role in object representation," *Trends Cogn. Sci.*, vol. 3, no. 4, pp. 151–162, 1999.

[136] M. X. Cohen, "A better way to define and describe morlet wavelets for time-frequency analysis," *Neuroimage*, vol. 199, pp. 81–86, Oct. 2019.

[137] T. Guo, T. Zhang, E. Lim, M. Lopez-Benitez, F. Ma, and L. Yu, "A review of wavelet analysis and its applications: Challenges and opportunities," *IEEE Access*, vol. 10, pp. 58 869–58 903, 2022.

[138] J. Morlet, G. Arens, E. Fourgeau, and D. Glard, "Wave propagation and sampling theory—part I: Complex signal and scattering in multilayered media," *Geophysics*, vol. 47, no. 2, pp. 203–221, Feb. 1982.

[139] E. W. Weisstein, "Convolution," https://mathworld.wolfram.com/Convolution.html, accessed: 2024-8-16.

[140] A. Guillet and F. Argoul, "Uncertainty and information in physiological signals: Explicit physical trade-off with log-normal wavelets," *J. Franklin Inst.*, vol. 361, no. 18, p. 107201, Dec. 2024.

[141] P. M. R. Reis, F. Hebenstreit, F. Gabsteiger, V. von Tscharner, and M. Lochmann, "Methodological aspects of EEG and body dynamics measurements during motion," *Front. Hum. Neurosci.*, vol. 8, p. 156, Mar. 2014.

[142] L. Wei, Y. Yu, Y. Qin, and S. Zhang, "A survey of EEG-based approaches to classroom attention assessment in education," *Information (Basel)*, vol. 16, no. 10, p. 860, Oct. 2025.

[143] W. Klonowski, E. Olejarczyk, and R. Stepien, "'epileptic seizures' in economic organism," *Physica A*, vol. 342, no. 3-4, pp. 701–707, Nov. 2004.

[144] J.-L. Tan, Z.-F. Liang, R. Zhang, Y.-Q. Dong, G.-H. Li, M. Zhang, H. Wang, and N. Xu, "Suppressing of power line artifact from electroencephalogram measurements using sparsity in frequency domain," *Front. Neurosci.*, vol. 15, p. 780373, Oct. 2021.

[145] M. Gyurkovics, G. M. Clements, K. A. Low, M. Fabiani, and G. Gratton, "The impact of 1/f activity and baseline correction on the results and interpretation of time-frequency analyses of EEG/MEG data: A cautionary tale," *Neuroimage*, vol. 237, no. 118192, p. 118192, Aug. 2021.

[146] H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Regularized common spatial patterns with generic learning for EEG signal classification," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2009, pp. 6599–6602, 2009.

[147] F. Costa, G. Patow, and J. Azorín, "Common spatial pattern for the classification of imagined geometric objects," in *Proceedings of the 3rd International Conference on Computer-Human Interaction Research and Applications.* SCITEPRESS - Science and Technology Publications, 2019.

[148] C. Zhang and A. Eskandarian, "A computationally efficient multiclass time-frequency common spatial pattern analysis on EEG motor imagery," *arXiv [eess.SP]*, 2020.

[149] O. Falzon, K. P. Camilleri, and J. Muscat, "The analytic common spatial patterns method for EEG-based BCI data," *J. Neural Eng.*, vol. 9, no. 4, p. 045009, Aug. 2012.

[150] X. Yong, R. K. Ward, and G. E. Birch, "Robust common spatial patterns for EEG signal preprocessing," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2008, pp. 2087–2090, 2008.

[151] A. M. Bastos and J.-M. Schoffelen, "A tutorial review of functional connectivity analysis methods and their interpretational pitfalls," *Front. Syst. Neurosci.*, vol. 9, p. 175, 2015.

[152] G. Van Hoey, R. Van de Walle, B. Vanrumste, M. D'Hav, I. Lemahieu, and P. Boon, "Beamforming techniques applied in EEG source analysis," https://raggedsleeverecords.com/wp-content/uploads/2016/03/beamforming-techniques-eeg-analysis1.pdf, accessed: 2024-12-16.

[153] J. C. Eliassen, E. L. Boespflug, M. Lamy, J. Allendorfer, W.-J. Chu, and J. P. Szaflarski, "Brain-mapping techniques for evaluating poststroke recovery and rehabilitation: a review," *Top. Stroke Rehabil.*, vol. 15, no. 5, pp. 427–450, 2008.

[154] R. P. N. Rao, *Brain-Computer Interfacing.* Cambridge University Press, Sep. 2013.

[155] S. Finnigan and M. J. A. M. van Putten, "EEG in ischaemic stroke: quantitative EEG can uniquely inform (sub-)acute prognoses and clinical management," *Clin. Neurophysiol.*, vol. 124, no. 1, pp. 10–19, Jan. 2013.

[156] Y. He, D. Eguren, J. M. Azorín, R. G. Grossman, T. P. Luu, and J. L. Contreras-Vidal, "Brain-machine interfaces for controlling lower-limb powered robotic systems," *J. Neural Eng.*, vol. 15, no. 2, p. 021004, Apr. 2018.

[157] J. Mellinger, G. Schalk, C. Braun, H. Preissl, W. Rosenstiel, N. Birbaumer, and A. Kübler, "An MEG-based brain-computer interface (BCI)," *Neuroimage*, vol. 36, no. 3, pp. 581–593, Jul. 2007.

[158] A. Y. Paek, A. Kilicarslan, B. Korenko, V. Gerginov, S. Knappe, and J. L. Contreras-Vidal, "Towards a portable magnetoencephalography based brain computer interface with optically-pumped magnetometers," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2020, pp. 3420–3423, Jul. 2020.

[159] S. Mellor, T. M. Tierney, R. A. Seymour, R. C. Timms, G. C. O'Neill, N. Alexander, M. E. Spedden, H. Payne, and G. R. Barnes, "Real-time, model-based magnetic field correction for moving, wearable MEG," *Neuroimage*, vol. 278, no. 120252, p. 120252, Sep. 2023.

[160] M. Rea, M. Rana, N. Lugato, P. Terekhin, L. Gizzi, D. Brötz, A. Fallgatter, N. Birbaumer, R. Sitaram, and A. Caria, "Lower limb movement preparation in chronic stroke: A pilot study toward an fNIRS-BCI for gait rehabilitation," *Neurorehabil. Neural Repair*, vol. 28, no. 6, pp. 564–575, Jul. 2014.

[161] P.-Y. Lin, J.-J. J. Chen, and S.-I. Lin, "The cortical control of cycling exercise in stroke patients: an fNIRS study," *Hum. Brain Mapp.*, vol. 34, no. 10, pp. 2381–2390, Oct. 2013.

[162] K. M. Arun, K. A. Smitha, P. N. Sylaja, and C. Kesavadas, "Identifying resting-state functional connectivity changes in the motor cortex using

fNIRS during recovery from stroke," *Brain Topogr.*, vol. 33, no. 6, pp. 710–719, Nov. 2020.

[163] W.-L. Chen, J. Wagner, N. Heugel, J. Sugar, Y.-W. Lee, L. Conant, M. Malloy, J. Heffernan, B. Quirk, A. Zinos, S. A. Beardsley, R. Prost, and H. T. Whelan, "Functional near-infrared spectroscopy and its clinical application in the field of neuroscience: Advances and future directions," *Front. Neurosci.*, vol. 14, p. 724, Jul. 2020.

[164] Y.-H. Kim, S. H. Jang, W. M. Byun, B. S. Han, K.-H. Lee, and S. H. Ahn, "Ipsilateral motor pathway confirmed by combined brain mapping of a patient with hemiparetic stroke: a case report," *Arch. Phys. Med. Rehabil.*, vol. 85, no. 8, pp. 1351–1353, Aug. 2004.

[165] A. Jaillard, C. D. Martin, K. Garambois, J. F. Lebas, and M. Hommel, "Vicarious function within the human primary motor cortex? a longitudinal fMRI stroke study," *Brain*, vol. 128, no. Pt 5, pp. 1122–1138, May 2005.

[166] S. S. Yeo and S. H. Jang, "Ipsilateral motor pathway without contralateral motor pathway in a stroke patient," *NeuroRehabilitation*, vol. 30, no. 4, pp. 303–306, 2012.

[167] A. K. Rehme, L. J. Volz, D.-L. Feis, S. B. Eickhoff, G. R. Fink, and C. Grefkes, "Individual prediction of chronic motor outcome in the acute post-stroke stage: Behavioral parameters versus functional imaging," *Hum. Brain Mapp.*, vol. 36, no. 11, pp. 4553–4565, Nov. 2015.

[168] X. Ru, K. He, B. Lyu, D. Li, W. Xu, W. Gu, X. Ma, J. Liu, C. Li, T. Li, F. Zheng, X. Yan, Y. Yin, H. Duan, S. Na, S. Wan, J. Qin, J. Sheng, and J.-H. Gao, "Multimodal neuroimaging with optically pumped magnetometers: A simultaneous MEG-EEG-fNIRS acquisition system," *Neuroimage*, vol. 259, no. 119420, p. 119420, Oct. 2022.

[169] C. Yen, C.-L. Lin, and M.-C. Chiang, "Exploring the frontiers of neuroimaging: A review of recent advances in understanding brain functioning and disorders," *Life (Basel)*, vol. 13, no. 7, p. 1472, Jun. 2023.

[170] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, "The CAVE: audio visual experience automatic virtual environment," *Commun. ACM*, vol. 35, no. 6, pp. 64–72, Jun. 1992.

[171] A. Pinegger, S. C. Wriessnegger, J. Faller, and G. R. Müller-Putz, "Evaluation of different EEG acquisition systems concerning their suitability for building a brain-computer interface: Case studies," *Front. Neurosci.*, vol. 10, p. 441, Sep. 2016.

[172] J. Xu, S. Mitra, C. Van Hoof, R. F. Yazicioglu, and K. A. A. Makinwa, "Active electrodes for wearable EEG acquisition: Review and electronics design methodology," *IEEE Rev. Biomed. Eng.*, vol. 10, pp. 187–198, Jan. 2017.

[173] J. J. Shih, D. J. Krusienski, and J. R. Wolpaw, "Brain-computer interfaces in medicine," *Mayo Clin. Proc.*, vol. 87, no. 3, pp. 268–279, Mar. 2012.

[174] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak, "Brain-computer interface spellers: A review," *Brain Sci*, vol. 8, no. 4, Mar. 2018.

[175] J. R. Wolpaw, "Chapter 6 - brain–computer interfaces," in *Handbook of Clinical Neurology*, M. P. Barnes and D. C. Good, Eds. Elsevier, Jan. 2013, vol. 110, pp. 67–74.

[176] P. Wierzgała, D. Zapała, G. M. Wojcik, and J. Masiak, "Most popular signal processing methods in motor-imagery BCI: A review and meta-analysis," *Front. Neuroinform.*, vol. 12, p. 78, Nov. 2018.

[177] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, Jan. 2012.

[178] A. Ortiz-Rosario and H. Adeli, "Brain-computer interface technologies: from signal to action," *Rev. Neurosci.*, vol. 24, no. 5, pp. 537–552, 2013.

[179] J. Decety, "The neurophysiological basis of motor imagery," *Behav. Brain Res.*, vol. 77, no. 1-2, pp. 45–52, May 1996.

[180] M. Lotze and U. Halsband, "Motor imagery," *J. Physiol. Paris*, vol. 99, no. 4-6, pp. 386–395, Jun. 2006.

[181] L. P. McAvinue and I. H. Robertson, "Measuring motor imagery ability: A review," *Eur. J. Cogn. Psychol.*, vol. 20, no. 2, pp. 232–251, Mar. 2008.

[182] E. Thomas, M. Dyson, and M. Clerc, "An analysis of performance evaluation for motor-imagery based BCI," *J. Neural Eng.*, vol. 10, no. 3, p. 031001, Jun. 2013.

[183] M. Hamedi, S.-H. Salleh, and A. M. Noor, "Electroencephalographic motor imagery brain connectivity analysis for BCI: A review," *Neural Comput.*, vol. 28, no. 6, pp. 999–1041, Jun. 2016.

[184] W.-Y. Hsu and Y.-N. Sun, "EEG-based motor imagery analysis using weighted wavelet transform features," *J. Neurosci. Methods*, vol. 176, no. 2, pp. 310–318, Jan. 2009.

[185] H. Yu, S. Ba, Y. Guo, L. Guo, and G. Xu, "Effects of motor imagery tasks on brain functional networks based on EEG mu/beta rhythm," *Brain Sci.*, vol. 12, no. 2, p. 194, Jan. 2022.

[186] S. Saha and M. Baumert, "Intra- and inter-subject variability in EEG-based sensorimotor brain computer interface: A review," *Front. Comput. Neurosci.*, vol. 13, Jan. 2020.

[187] A. Destexhe, M. Rudolph, and D. Paré, "The high-conductance state of neocortical neurons in vivo," *Nat. Rev. Neurosci.*, vol. 4, no. 9, pp. 739–751, Sep. 2003.

[188] W. van Drongelen, *Signal Processing for Neuroscientists*, 2nd ed. San Diego, CA: Academic Press, May 2018.

[189] S. Sanei and J. A. Chambers, *EEG Signal Processing.* Newy York: Wiley-Interscience, May 2013.

[190] C. Q. Lai, H. Ibrahim, M. Z. Abdullah, J. M. Abdullah, S. A. Suandi, and A. Azman, "Artifacts and noise removal for electroencephalogram (EEG): A literature review," in *2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE).* IEEE, Apr. 2018.

[191] S. Makeig and A. J. Bell, "Independent component analysis of electroencephalographic data," https://proceedings.neurips.cc/paper_files/paper/1995/file/754dda4b1ba34c6fa89716b85d68532b-Paper.pdf, accessed: 2024-12-30.

[192] J.-F. Cardoso, "Blind signal separation: statistical principles," *Proc. IEEE Inst. Electr. Electron. Eng.*, vol. 86, no. 10, pp. 2009–2025, 1998.

[193] J. Iriarte, E. Urrestarazu, M. Valencia, M. Alegre, A. Malanda, C. Viteri, and J. Artieda, "Independent component analysis as a tool to eliminate artifacts in EEG: a quantitative study," *J. Clin. Neurophysiol.*, vol. 20, no. 4, pp. 249–257, Jul. 2003.

[194] E. Oja, A. Hyvarinen, and J. Karhunen, *Independent Component Analysis*, 1st ed., ser. Adaptive and Cognitive Dynamic Systems: Signal Processing, Learning, Communications and Control. Nashville, TN: John Wiley & Sons, Mar. 2004.

[195] I. Nouira, A. Ben Abdallah, and M. H. Bedoui, "EEG potential mapping by 3D interpolation methods," in *2014 International Conference on Multimedia Computing and Systems (ICMCS).* IEEE, Apr. 2014.

[196] F. Perrin, J. Pernier, O. Bertrand, and J. F. Echallier, “Spherical splines for scalp potential and current density mapping,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 72, no. 2, pp. 184–187, Feb. 1989.

[197] N. Yamawaki, C. Wilke, Z. Liu, and B. He, “An enhanced time-frequency-spatial approach for motor imagery classification,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, no. 2, pp. 250–254, Jun. 2006.

[198] T. Wang, J. Deng, and B. He, “Classifying EEG-based motor imagery tasks by means of time-frequency synthesized spatial patterns,” *Clin. Neurophysiol.*, vol. 115, no. 12, pp. 2744–2753, Dec. 2004.

[199] T. Wang and B. He, “An efficient rhythmic component expression and weighting synthesis strategy for classifying motor imagery EEG in a brain-computer interface,” *J. Neural Eng.*, vol. 1, no. 1, pp. 1–7, Mar. 2004.

[200] A. Ferrante, C. Gavriel, and A. Faisal, “Data-efficient hand motor imagery decoding in EEG-BCI by using morlet wavelets & common spatial pattern algorithms,” in *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, Apr. 2015.

[201] N. M. Jadeja, *How to read an EEG*. Cambridge, England: Cambridge University Press, Jun. 2021.

[202] P. Delgado, P. Passos, F. Tavares, and H. A. Ferreira, “Electroencephalographic and cognitive task analysis of working memory and attention in athletes: A systematic review,” *Psychol. Sport Exerc.*, vol. 80, no. 102930, p. 102930, Sep. 2025.

[203] M. Degirmenci, Y. K. Yuce, M. Perc, and Y. Isler, “Statistically significant features improve binary and multiple motor imagery task predictions from EEGs,” *Front. Hum. Neurosci.*, vol. 17, p. 1223307, Jul. 2023.

[204] D. M. Groppe, T. P. Urbach, and M. Kutas, “Mass univariate analysis of event-related brain potentials/fields II: Simulation studies,” *Psychophysiology*, vol. 48, no. 12, pp. 1726–1737, Dec. 2011.

[205] V. Saccá, M. Campolo, D. Mirarchi, A. Gambardella, P. Veltri, and F. C. Morabito, “On the classification of EEG signal by using an SVM based algorithm,” in *Multidisciplinary Approaches to Neural Computing*, ser. Smart innovation, systems and technologies. Cham: Springer International Publishing, 2018, pp. 271–278.

[206] X. Zhao, H. Zhang, G. Zhu, F. You, S. Kuang, and L. Sun, “A multi-branch 3D convolutional neural network for EEG-based motor imagery classification,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 10, pp. 2164–2177, Oct. 2019.

[207] T. Liu and D. Yang, "A three-branch 3D convolutional neural network for EEG-based different hand movement stages classification," *Sci. Rep.*, vol. 11, no. 1, p. 10758, May 2021.

[208] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "LSTM-based EEG classification in motor imagery tasks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2086–2095, Nov. 2018.

[209] S. Choo, H. Park, J.-Y. Jung, K. Flores, and C. S. Nam, "Improving classification performance of motor imagery BCI through EEG data augmentation with conditional generative adversarial networks," *Neural Netw.*, vol. 180, no. 106665, p. 106665, Dec. 2024.

[210] M. Saqib, Y. Zhu, M. D. Wang, and B. Beaulieu-Jones, "Regularization of deep neural networks for EEG seizure detection to mitigate overfitting," *Proc. COMPSAC*, vol. 2020, pp. 664–673, Jul. 2020.

[211] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. E215–20, Jun. 2000.

[212] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.

[213] H. Yuan, T. Liu, R. Szarkowski, C. Rios, J. Ashe, and B. He, "Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: an EEG and fMRI study of motor imagery and movements," *Neuroimage*, vol. 49, no. 3, pp. 2596–2606, Feb. 2010.

[214] T. Zhan, K. Yin, J. Xiong, Z. He, and S.-T. Wu, "Augmented reality and virtual reality displays: Perspectives and challenges," *iScience*, vol. 23, no. 8, p. 101397, Aug. 2020.

[215] R. Perla and R. Hebbalaguppe, "Google cardboard dates augmented reality : Issues, challenges and future opportunities," *arXiv [cs.HC]*, Jun. 2017.

[216] P. Tranton, *Samsung gear vr: An easy guide for beginners.* Morrisville, NC: Lulu.com, Jan. 2016.

[217] P. A. Rauschnabel, A. Rossmann, and M. C. tom Dieck, "An adoption framework for mobile augmented reality games: The case of pokémon go," *Comput. Human Behav.*, vol. 76, pp. 276–286, Nov. 2017.

[218] Z. Oufqir, A. El Abderrahmani, and K. Satori, “ARKit and ARCore in serve to augmented reality,” in *2020 International Conference on Intelligent Systems and Computer Vision (ISCV).* IEEE, Jun. 2020.

[219] S. H. Ko and J. Rogers, “Functional materials and devices for XR (VR/AR/MR) applications,” *Adv. Funct. Mater.*, vol. 31, no. 39, p. 2106546, Sep. 2021.

[220] C. Ferguson, P. M. Davidson, P. J. Scott, D. Jackson, and L. D. Hickman, “Augmented reality, virtual reality and gaming: an integral part of nursing,” *Contemp. Nurse*, vol. 51, no. 1, pp. 1–4, 2015.

[221] F. Bonetti, G. Warnaby, and L. Quinn, “Augmented reality and virtual reality in physical and online retailing: A review, synthesis and research agenda,” in *Augmented Reality and Virtual Reality: Empowering Human, Place and Business*, T. Jung and M. C. tom Dieck, Eds. Cham: Springer International Publishing, 2018, pp. 119–132.

[222] K.-T. Huang, C. Ball, J. Francis, R. Ratan, J. Boumis, and J. Fordham, “Augmented versus virtual reality in education: An exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications,” *Cyberpsychol. Behav. Soc. Netw.*, vol. 22, no. 2, pp. 105–110, Feb. 2019.

[223] L. V. Eshuis, M. J. van Gelderen, M. van Zuiden, M. J. Nijdam, E. Vermetten, M. Olff, and A. Bakker, “Efficacy of immersive PTSD treatments: A systematic review of virtual and augmented reality exposure therapy and a meta-analysis of virtual reality exposure therapy,” *J. Psychiatr. Res.*, vol. 143, pp. 516–527, Nov. 2021.

[224] S. E. Kober, J. Kurzmann, and C. Neuper, “Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: an EEG study,” *Int. J. Psychophysiol.*, vol. 83, no. 3, pp. 365–374, Mar. 2012.

[225] J. Steuer, “Defining virtual reality: Dimensions determining telepresence,” *J. Commun.*, vol. 42, no. 4, pp. 73–93, Dec. 1992.

[226] A. Krugliak and A. Clarke, “Towards real-world neuroscience using mobile EEG and augmented reality,” *Sci. Rep.*, vol. 12, no. 1, p. 2291, Feb. 2022.

[227] A. A. De Luna, *Introduction to virtual reality.* Oakville, ON, Canada: Arcler Press, Jan. 2022.

[228] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, “A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda,” *Comput. Educ.*, vol. 147, no. 103778, p. 103778, Apr. 2020.

[229] C. Morse, “Gaming engines: Unity, unreal, and interactive 3D spaces,” *Technol.—Archit. Des.*, vol. 5, no. 2, pp. 246–249, Jul. 2021.

[230] lolambean, markingmyname, AMollis, Sean-Kerawala, rafek, alexbuckgit, DhurataJ, typride, TimShererWithAquent, vtieto, qianw, nivnar, hferrone, qmberli, and DCtheGeek, “What is mixed reality?” https://learn.microsoft.com/en-us/windows/mixed-reality/discover/mixed-reality, Jan. 2023, accessed: 2024-12-18.

[231] G. Zari, S. Condino, F. Cutolo, and V. Ferrari, “Magic leap 1 versus microsoft HoloLens 2 for the visualization of 3D content obtained from radiological images,” *Sensors (Basel)*, vol. 23, no. 6, p. 3040, Mar. 2023.

[232] S. Rokhsaritalemi, A. Sadeghi-Niaraki, and S.-M. Choi, “A review on mixed reality: Current trends, challenges and prospects,” *Appl. Sci. (Basel)*, vol. 10, no. 2, p. 636, Jan. 2020.

[233] “Processor speed: What is processor speed,” https://www.lenovo.com/gb/en/glossary/what-is-processor-speed/, accessed: 2026-1-20.

[234] J. Peirce, J. R. Gray, S. Simpson, M. MacAskill, R. Höchenberger, H. Sogo, E. Kastman, and J. K. Lindeløv, “PsychoPy2: Experiments in behavior made easy,” *Behav. Res. Methods*, vol. 51, no. 1, pp. 195–203, Feb. 2019.

[235] C. Bartels, M. Wegrzyn, A. Wiedl, V. Ackermann, and H. Ehrenreich, “Practice effects in healthy adults: a longitudinal study on frequent repetitive cognitive testing,” *BMC Neurosci.*, vol. 11, no. 1, p. 118, Sep. 2010.

[236] B. M. O’Donell, J. F. Barraza, and E. M. Colombo, “The effect of chromatic and luminance information on reaction times,” *Vis. Neurosci.*, vol. 27, no. 3-4, pp. 119–129, Jul. 2010.

[237] A. V. Cardello, “Reaction time and visual brightness: within-subject correlations,” *Percept. Mot. Skills*, vol. 48, no. 1, pp. 107–115, Feb. 1979.

[238] R. Knight, C. Mazzi, and S. Savazzi, “Assessing the effects of physical and perceived luminance contrast on RT and TMS-induced percepts,” *Exp. Brain Res.*, vol. 233, no. 12, pp. 3527–3534, Dec. 2015.

[239] D. Bridges, A. Pitiot, M. R. MacAskill, and J. W. Peirce, “The timing mega-study: comparing a range of experiment generators, both lab-based and online,” *PeerJ*, vol. 8, p. e9414, Jul. 2020.

[240] K. Raaen and I. Kjellmo, “Measuring latency in virtual reality systems,” in *Entertainment Computing - ICEC 2015.* Springer International Publishing, 2015, pp. 457–462.

[241] M. Faraday, "XXXII. on static electrical inductive action," *Lond. Edinb. Philos. Mag. J. Sci.*, vol. 22, no. 144, pp. 200–204, Mar. 1843.

[242] J. J. Gibson and L. Carmichael, *The senses considered as perceptual systems.* Houghton Mifflin Boston, 1966, vol. 2.

[243] R. L. Gregory, "The intelligent eye," 1970.

[244] M. Mishkin, L. G. Ungerleider, and K. A. Macko, "Object vision and spatial vision: two cortical pathways," *Trends Neurosci.*, vol. 6, pp. 414–417, Jan. 1983.

[245] L. Shen, X. Hu, E. Yacoub, and K. Ugurbil, "Neural correlates of visual form and visual spatial processing," *Hum. Brain Mapp.*, vol. 8, no. 1, pp. 60–71, 1999.

[246] A. E. Richardson, D. R. Montello, and M. Hegarty, "Spatial knowledge acquisition from maps and from navigation in real and virtual environments," *Mem. Cognit.*, vol. 27, no. 4, pp. 741–750, Jul. 1999.

[247] S. Babilon and T. Q. Khanh, "Color appearance rating of familiar real objects under immersive viewing conditions," *Color Res. Appl.*, vol. 43, no. 4, pp. 551–568, Aug. 2018.

[248] F. Díaz-Barrancas, H. Cwierz, P. J. Pardo, L. Pérez, and M. I. Suero, "Spectral color management in virtual reality scenes," *Sensors (Basel)*, vol. 20, no. 19, p. 5658, Oct. 2020.

[249] G. Y. Yildiz, R. Skarbez, I. Sperandio, S. J. Chen, I. J. Mulder, and P. A. Chouinard, "Linear perspective cues have a greater effect on the perceptual rescaling of distant stimuli than textures in the virtual environment," *Atten. Percept. Psychophys.*, vol. 86, no. 2, pp. 653–665, Feb. 2024.

[250] F. El Jamiy and R. Marsh, "Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality," *IET Image Process.*, vol. 13, no. 5, pp. 707–712, Apr. 2019.

[251] J. Y. F. d. L. Antão, J. Y. F. de Lima Antão, L. C. de Abreu, R. T. de Almeida Barbosa, T. B. Crocetta, R. Guarnieri, T. Massetti, T. P. C. Antunes, J. Tonks, and C. B. de Mello Monteiro, "Use of augmented reality with a motion-controlled game utilizing alphabet letters and numbers to improve performance and reaction time skills for people with autism spectrum disorder," pp. 16–22, 2020.

[252] T. Demiralp, Z. Bayraktaroglu, D. Lenz, S. Junge, N. A. Busch, B. Maess, M. Ergen, and C. S. Herrmann, "Gamma amplitudes are coupled to theta

phase in human EEG during visual perception," *Int. J. Psychophysiol.*, vol. 64, no. 1, pp. 24–30, Apr. 2007.

[253] N. A. Busch, J. Dubois, and R. VanRullen, "The phase of ongoing EEG oscillations predicts visual perception," *J. Neurosci.*, vol. 29, no. 24, pp. 7869–7876, Jun. 2009.

[254] H. Piéron, "II. recherches sur les lois de variation des temps de latence sensorielle en fonction des intensités excitatrices," *Annee Psychol.*, vol. 20, no. 1, pp. 17–96, 1913.

[255] R. Ratcliff, "Methods for dealing with reaction time outliers," *Psychol. Bull.*, vol. 114, no. 3, pp. 510–532, 1993.

[256] R. Ulrich and J. Miller, "Effects of truncation on reaction time analysis," *J. Exp. Psychol. Gen.*, vol. 123, no. 1, pp. 34–80, Mar. 1994.

[257] D. J. Lewkowicz, "The concept of ecological validity: What are its limitations and is it bad to be invalid?" *Infancy*, vol. 2, no. 4, pp. 437–450, Oct. 2001.

[258] A. Gronowski, D. C. Arness, J. Ng, Z. Qu, C. W. Lau, D. Catchpoole, and Q. V. Nguyen, "The impact of virtual and augmented reality on presence, user experience and performance of information visualisation," *Virtual Real.*, vol. 28, no. 3, Jul. 2024.

[259] X. Zhou, A. U. Batmaz, A. S. Williams, D. Schreiber, and F. R. Ortega, "I did not notice: A comparison of immersive analytics with augmented and virtual reality," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems.* New York, NY, USA: ACM, May 2024.

[260] L. F. H. Basile, R. Anghinah, P. Ribeiro, R. T. Ramos, R. Piedade, G. Ballester, and E. P. Brunetti, "Interindividual variability in EEG correlates of attention and limits of functional mapping," *Int. J. Psychophysiol.*, vol. 65, no. 3, pp. 238–251, Sep. 2007.

[261] M. B. Sacoman, "Irlen syndrome: Manifestations and the intervention process," https://ecronicon.net/assets/ecne/pdf/ECNE-12-00640.pdf, accessed: 2024-11-11.

[262] J. H. Kim, H.-J. Seo, S.-G. Ha, and S.-H. Kim, "Functional magnetic resonance imaging findings in meares-irlen syndrome: a pilot sudy," *Korean J. Ophthalmol.*, vol. 29, no. 2, pp. 121–125, Apr. 2015.

APPENDICES

A Matlab code

A.1 Fourier transform

We start with $e^{-i\omega t}$. This equation can be generated using:

```
c1 = 0.2;

G = exp(2*pi*c1*1i*t)
```

Next, equation B.1 can be plotted using:

```
figure;

plot(G,"-")

axis equal

grid on

xlabel("Re")

ylabel("Im")

axis([-2.5 2.5 -2.5 2.5])

ax = gca;

ax.XAxisLocation = ’origin’;

ax.YAxisLocation = ’origin’;
```

To generate the signal for equation B.2, we can use:

```
t = 0:0.01:100;
```

```
f = 5;

S = cos(2*pi*f*t) + 1.1;
```

Next, equation B.2 can be plotted using:

```
figure;

hold on;

plot(F,"-")

plot(COM, ’rx’, ’MarkerSize’, 10, ’LineWidth’, 2)

axis equal

grid on

xlabel("Re")

ylabel("Im")

axis([-2.5 2.5 -2.5 2.5])

ax = gca;

ax.XAxisLocation = ’origin’;

ax.YAxisLocation = ’origin’;

hold off;
```

Equation B.3 can then be plotted by setting $c1 = f$, then running the above code again to generate the new plot.

To get all the averages of all points on the complex domain plot, we use:

```
c_values = 0:0.01:20;

COMarr = [];


for c = c_values

    G = exp(2*pi*c*1i*t);

    F = S.*G;

    COM = sum(F)/length(t);

    COMarr = [COMarr; COM];

end
```

We then plot equation B.4 using:

```
figure;
plot(c_values, COMarr);
xlabel('Frequency (Hz)');
ylabel('Magnitude');
grid on;
```

A.2 STFT and wavelet transform

We start with the surrogate signal equation 4.12, shown in figure 4.2. This signal can be generated using:

```
f_t = zeros(size(t));
% Signal for t = 0 to t = 10 with 4 Hz and 10 Hz components
t1 = t(t >= 0 & t <= 3);
f_t(t >= 0 & t <= 3) = sin(2*pi*4*t1) + sin(2*pi*10*t1);
% Signal for t = 10.01 to t = 20 with 6 Hz and 12 Hz components
t2 = t(t > 3 & t <= 6);
% Signal for t = 20.01 to t = 30 with 8 Hz and 14 Hz components
f_t(t > 3 & t <= 6) = sin(2*pi*6*t2) + sin(2*pi*12*t2);
t3 = t(t > 6 & t <= 9);
f_t(t > 6 & t <= 9) = sin(2*pi*8*t3) + sin(2*pi*14*t3);
```

We can try applying a Fourier transform, as shown in figure 4.3. This can be done using:

```
Fs = 1 / (t(2) - t(1));  % Sampling frequency
N = length(f_t);          % Number of samples
f = linspace(0, Fs/2, floor(N/2)+1);  % Frequency vector for plotting
F_f_t = fft(f_t); % Compute FFT
P_f_t = abs(F_f_t) / N;   % Normalize the magnitude
```

```
P_f_t = P_f_t(1:length(f)) * 2;  % Take the positive half of the spectrum
freq_range = (f >= 1 & f <= 20);  % Logical index for desired frequency range
f_plot = f(freq_range);           % Frequencies in the desired range
P_f_t_plot = P_f_t(freq_range);   % Magnitudes in the desired range
```

Figure 4.2 and figure 4.3 can then be plotted together using:

```
figure;
subplot(2,1,1);
plot(t, f_t);
title('Time-Domain Signal');
xlabel('Time (s)');
ylabel('Amplitude');

subplot(2,1,2);
plot(f_plot, P_f_t_plot);
title('Magnitude of Fourier Transform');
xlabel('Frequency (Hz)');
ylabel('Magnitude');
```

We then define the parameters for STFT

```
segment_duration = 0.5;  % Duration of each segment in seconds
segment_samples = segment_duration * Fs;  % Number of samples in each segment
overlap_samples = 0;  % Overlap between segments (can be set to half or more fo
n_segments = floor(length(f_t) / segment_samples); % Number of segments
```

Figure 4.4 can then be plotted using:

```
figure;
spectrogram(f_t, segment_samples, overlap_samples, [], Fs, 'yaxis');
title('Spectrogram of Signal');
xlabel('Time (s)');
```

```
ylabel('Frequency (Hz)');
ylim([0 35]);  % Set y-axis limits to 0 to 30 Hz
colorbar;
```

The wavelet transform using Morlet wavelet can be easily applied using one line of code:

```
[wt, frequencies] = cwt(f_t, 'amor', Fs);
```

Figure 4.6 can then be plotted using:

```
figure;
surface(t, frequencies, abs(wt));
shading interp;
axis tight;
xlabel('Time (s)');
ylabel('Frequency (Hz)');
title('Time-Frequency Spectrogram using CWT');
colorbar;
```

The Hilbert-Huang transform can be applied using:

```
emd(f_t);
imf = emd(f_t,'Display',1);
hht(imf,Fs,'FrequencyLimits',[0 35])
```

These functions automatically generate the figures for the IMFs and the HHT spectrogram.

B Basic mathematics

B.1 Basics of Fourier transform

The Fourier Transform is a mathematical operation that decomposes a time-domain signal into its constituent frequencies, transforming it from the time domain to the frequency domain (cite). Mathematically, the Fourier Transform of the signal $H(t)$ can be represented mathematically using equation B.1.

$$F(\omega) = \int_{-\infty}^{+\infty} H(t)\mathrm{e}^{-i\omega t} dt \tag{B.1}$$

where $F(\omega)$ represents the fourier transform of the signal, ω is the angular frequency, and $\mathrm{e}^{-i\omega t}$ (a complex exponential) captures both the amplitude and phase of each frequency component. $|F(\omega)^2|$ represents the frequency spectrum of the signal.

To understand how Fourier transform works, we can start with a surrogate signal $H(t)$ defined by B.2:

$$H(t) = cos(2\pi t) + 1.1 \tag{B.2}$$

The purpose of adding 1.1 is to move the whole cosine signal above the x-axis and all positive, for easier analysis. Next, we plot a complex domain plot of $\mathrm{e}^{-i\omega t}$, as shown in figure B.1:

Notice how $\mathrm{e}^{-i\omega t}$ draws a circle on the complex domain plot. This makes $\mathrm{e}^{-i\omega t}$ such a useful equation to represent rotation, or revolution, in mathematics.

The magnitude value on each time point in the raw signal is multiplied with the corresponding value on the $\mathrm{e}^{-i\omega t}$ complex domain plot. In this case, we start

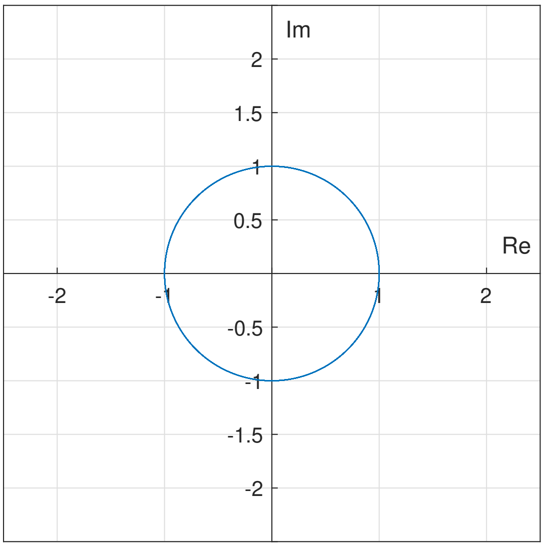


Figure B.1: Plot of $\mathrm{e}^{-i\omega t}$ on a complex domain plot

with $f = 0.2$, or $\omega = 0.4\pi$. We get the resulting plot as shown in in figure B.2:

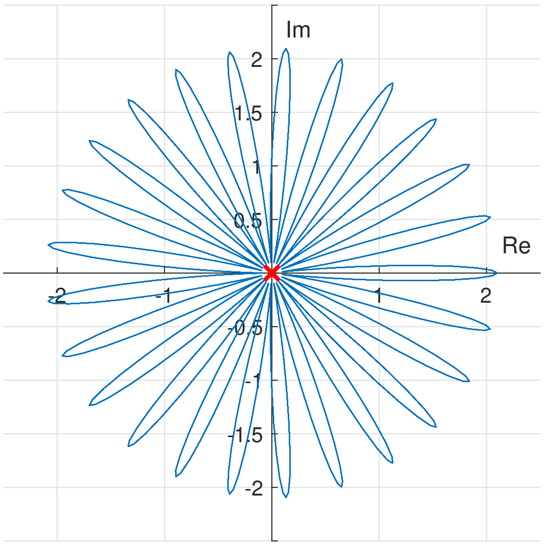


Figure B.2: Multiplying $\mathrm{e}^{-i\omega t}$ with $H(t)$ on a complex domain plot, where $f = 0.2$, showing an average close to 0

We then find the average of all points on the resulting complex domain plot, represented by the red cross on figure B.2. Next, we repeat the step with $f = 5$, or $\omega = 10\pi$, the same frequency as that of the signal, as shown in figure B.3.

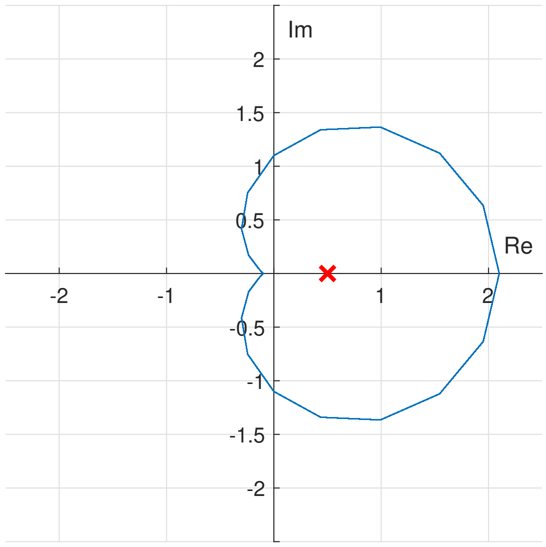


Figure B.3: Multiplying $\mathrm{e}^{-i\omega t}$ with $H(t)$ on a complex domain plot, where $f = 5$, showing a jump in average to 0.5

Notice how when the frequency of $\mathrm{e}^{-i\omega t}$ does not match the frequency of the signal, the resulting average is close to 0. However, when the frequency of $\mathrm{e}^{-i\omega t}$ is equals to the frequency of the signal, the resulting average spikes up, in this case to 0.5. To illustrate further, we plot the resulting averages against the frequency of $\mathrm{e}^{-i\omega t}$ in figure B.4.

B.2 Basics of linear algebra

In order to understand how the CSP algorithm works geometrically, one must have a solid foundation in the understanding basic linear algebra. Many textbooks on linear algebra have already discussed this topic at length, so a summary of concepts required to understand the CSP algorithm will be discussed here.

A discussion on linear algebra usually begins with vectors, which are the alphabets of linear algebra. Simply put, vectors are mathematical objects with both magnitude and direction. For example, let's say we want to express the following vector on the cartesian plane.

We can write this mathematically as a vertical matrix, with the x and y

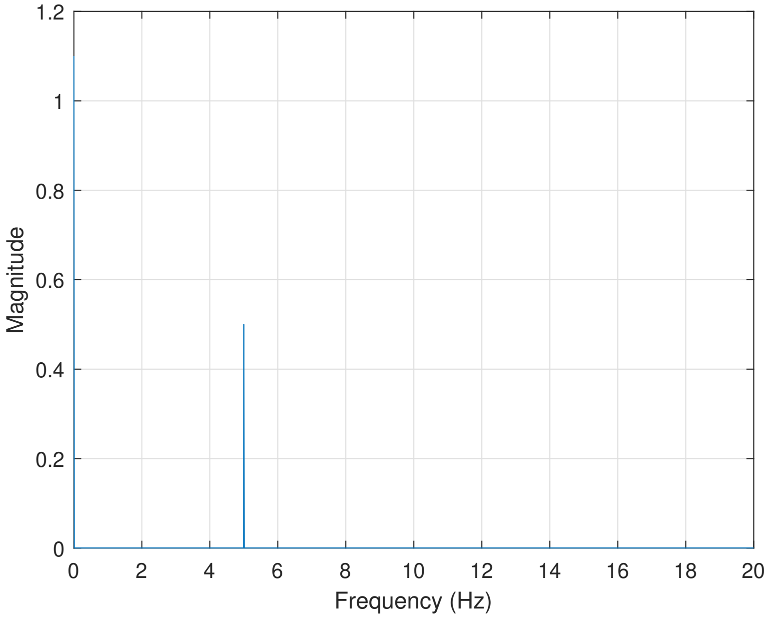


Figure B.4: Discrete fourier transform, where all the resulting averages are plotted against the frequency of $e^{-i\omega t}$ on a complex domain plot

coordinates of the point the vector ends on, as the first and second numbers respectively, like so: $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$. This allows us to define a vector that lands on point (2,3), assuming that it starts from the origin.

B.2.1 Matrixes

EEG signals are usually presented in a graph form, with a line for each channel. This allows researchers to visually inspect each channel before the signals are further processed.

However, in order for a computer to process these signals, they would have to be represented in numbers form. This is where matrixes come into play. The channels can be represented as rows, and the magnitude each signal points can be represented as a single element in the matrix.

B.2.2 Linear transformation

Linear transformation is the mapping of a single vector onto another vector space. This is defined by identifying the basis vectors in the new vector space, assuming

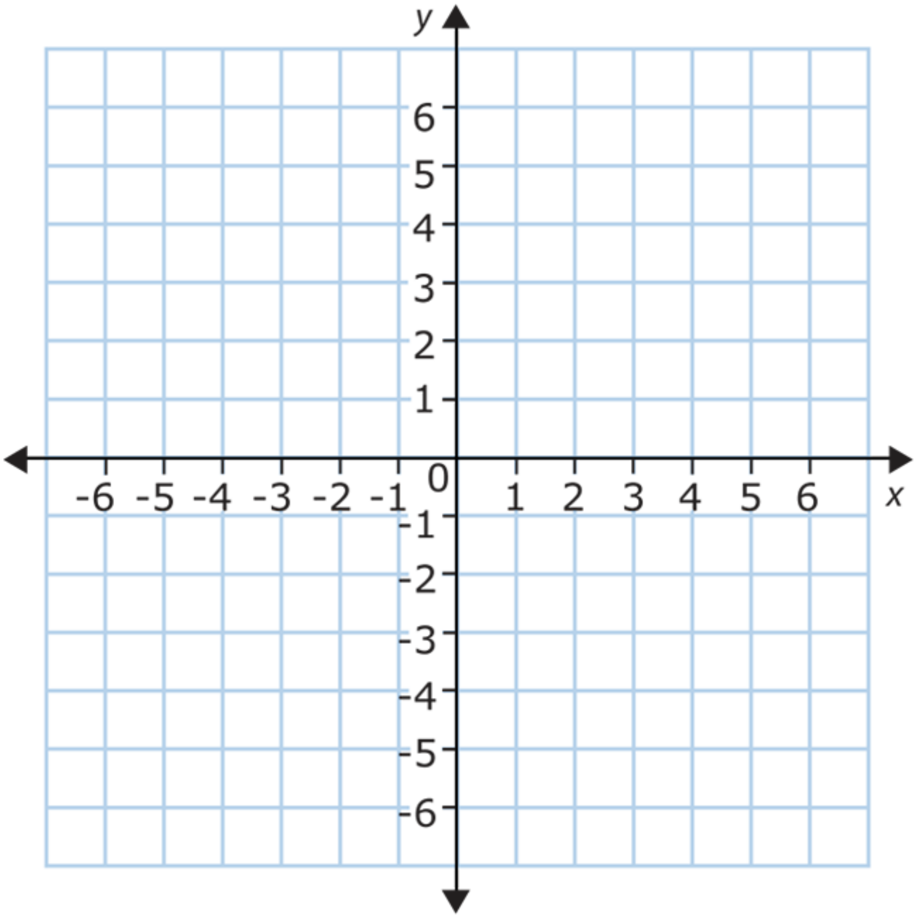


Figure B.5: A general cartisian plane.

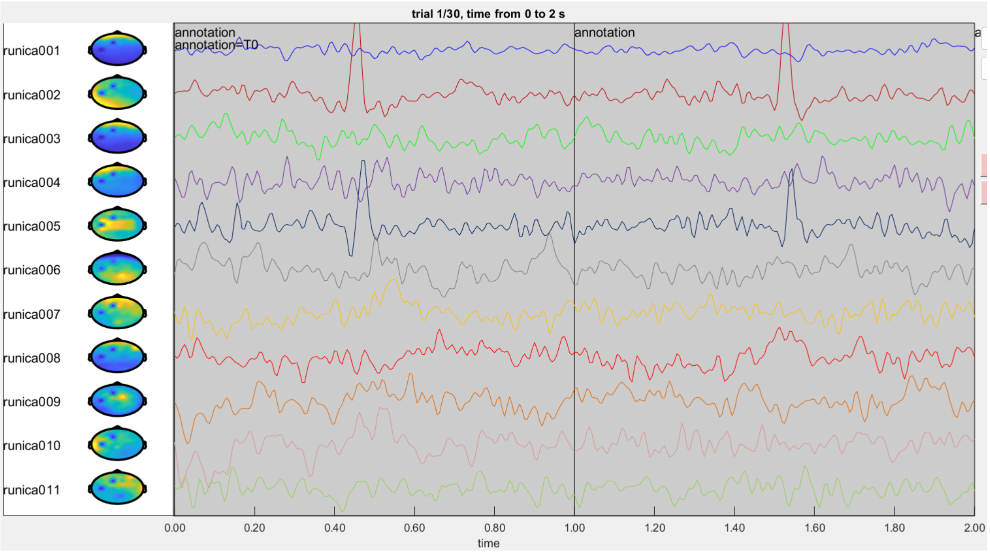


Figure B.6: EEG signal captured from a subject who is at rest.

that the basis vector in the original vector space is i = [1,0] and j = [0,1].

Visual inspection shows that by taking the eigenvectors of the covariance matrix as the transformation matrix. Mathematically prove that the eigenvectors of the covariance matrix is a rotation matrix, and scales the matrix by [1,0;0,-1].

Suppose we have:

$$T = R \times S$$

Where T is the transformation matrix, R is the rotation matrix and S is the scaling matrix. We also have:

$$R = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Let:

$$S = \begin{bmatrix} cos(a) & -sin(b) \\ sin(c) & cos(d) \end{bmatrix}$$

Linear transformation equations:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x \begin{bmatrix} a \\ c \end{bmatrix} + y \begin{bmatrix} b \\ d \end{bmatrix}$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

So:

$$T = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \times \begin{bmatrix} \cos(a) & -\sin(b) \\ \sin(c) & \cos(d) \end{bmatrix} = \begin{bmatrix} \cos(a) + 0 & -sin(b) + 0 \\ 0 - sin(c) & 0 - cos(d) \end{bmatrix} = \begin{bmatrix} cos(a) & sin(b) \\ sin(c) & cos(d) \end{bmatrix}$$

Where a, b, c and d are arbitrary constants. Prove that if T = eigenvectors of covariance matrix, $a = b = c = d$.

Given:

$$cov = \begin{bmatrix} var(x) & cov(x,y) \\ cov(x,y) & var(y) \end{bmatrix} = \begin{bmatrix} \frac{\sum_1^n (x_i - \mu_x)^2}{n} & \frac{\sum_1^n (x_i - \mu_x)(y_i - \mu_y)}{n} \\ \frac{\sum_1^n (x_i - \mu_x)(y_i - \mu_y)}{n} & \frac{\sum_1^n (y_i - \mu_y)^2}{n} \end{bmatrix} = \frac{1}{n} \left[\begin{matrix} \sum_1^n (x_i - \mu_x)^2 \\ \sum_1^n (x_i - \mu_x)(y_i - \end{matrix} \right.$$

And:

$$cov \times \overrightarrow{v} = \lambda \times \overrightarrow{v}$$

$$cov = \overrightarrow{v} \times \lambda \times \overrightarrow{v}^{-1}$$

"Positive definite" refers to a real matrix A having the property that xTAx is positive for all non-zero vectors x.

B.2.3 Fourier transform

The most fundamental way break a signal down into it's frequency components is through the Fourier Transform. Fourier transform allows for the decomposition of complex, time-domain EEG signals into their frequency components, making it easier to analyze brain activity in specific frequency bands. Fourier transform works by transforming the EEG signal from the time domain to the frequency domain, which results in a spectrum showing the amplitude or power of each frequency component within the signal. A tutorial on the mathematics is detailed in the appendix. The frequency-based representation of the EEG signal allows for the identification of brain rhythm at specific frequencies, which are associated with cognitive functions or disorders. Additionally, it is widely used to filter noise, isolate specific frequency bands, and apply spectral analysis techniques.

There are two main types of Fourier Transforms: the Discrete Fourier Transform (DFT) and the Continuous Fourier Transform (CFT), each suited to different types of signals.

The Continuous Fourier Transform is used for continuous, analog signals, where time or space can be measured at infinitely small intervals. The CFT transforms a continuous-time signal $f(t)$ into its frequency-domain representation $F(\omega)$. It is given by the integral:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-i\omega t}dt \tag{B.3}$$

Here, ω represents the frequency, and i is the imaginary unit. On the other hand, the Discrete Fourier Transform is used for discrete, sampled signals. The DFT transforms a discrete signal $x[n]$, where n is the index of samples, into its frequency-domain representation $X[k]$ through the sum:

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-i(2\pi kn/N)} \tag{B.4}$$

where N is the number of samples, and k represents discrete frequency bins. As digital systems do not have the ability to store analog data, any analog signals recorded, such as EEG signal, are converted to discrete digital data before being stored. Hence, DFT is typically used to convert a digital signal into its frequency-domain representation.

As this chapter is focused on the mathematics of signal processing, the continuous equations, such as CFT will be shown. However, it should be noted that the surrogate signals used in this chapter are still discrete digital signals, so discrete methods are needed to process them. In the next chapter, where signal processing techinques and their applications are discussed, discrete equations, such as DFT will be shown.