Size Constancy in Motion: Neural Signatures of Perception in Dynamic Visual Environments

---

Inaugural-Dissertation
Zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)

An der Justus-Liebig Universität Gießen,
Fachbereich 06: Psychologie und Sportwissenschaften
Otto-Behagel-Straße 10F 35394 Gießen, Deutschland

Vorgelegt von
Elef Schellen

2025

Tag der Disputation: 20.03.2026

Dekan
Prof. Dr. Christiane Hermann

1. Gutachter und Betreuer
Prof. Dr. Karl R. Gegenfurtner

2. Gutachterin
Prof. Dr. Katja Fiehler

To my father, Karel Pieter Schellen

Acknowledgements

I would like to express my gratitude first and foremost to my supervisor, Professor Karl Gegenfurtner, under whose proverbial wings this entire project was able to take shape. I consider it a great privilege to have been able to work as a member of his laboratory, and by extension as a member of the great team of scientists that make up the department of experimental psychology at the Justus-Liebig university in Giessen.

I would further like to express my gratitude to Katja Fiehler, my second supervisor and speaker for the IRTG-1901, which has been a wonderful programme which expanded my academic horizons.

I am further grateful for the great colleagues and peers that have provided their assistance to this project. Specifically: for Robert Ennis and his great tech support all throughout this project, for Florian Bayer and his technical insights, for Natalia Pfening and her extraordinary 3D modeling skills and for Hartmut Studenroth and Sasha Bloch for their expert practical engineering.

There are still many more colleagues who I am grateful to for various reasons, though if I were to mention them all by name, I would turn this acknowledgements section into a thesis chapter unto itself. Let it therefore suffice for me to say that the support and general presence of my many co workers has made my time working in this project as pleasant as it was.

My gratitude, finally, goes out to my family, who have been a source of unwavering support and a foundation on which I could always rely.

Preface

Imagine a warm, summer afternoon. You are lying in a hammock in your backyard, gently rocking back and forth. Meanwhile, some of your friends are kicking a football around, running around the yard, sometimes closer to you, sometimes further away. You would probably imagine this to be a rather pleasant set of circumstances.

However, without the critical factor of perceptual constancy, this scenario would be awfully disorienting and exceedingly unpleasant. If we could not anchor our perceptual inputs to a constant, internal representation of the outside world, with each hammock swing, your body would struggle to adapt to shifting perspectives. The world would appear to sway and pitch uncontrollably, the ground twisting to unnatural angles. Your friends would seem to shrink and grow unnaturally, going from child sized to giant within seconds. If one of your friends kicked the ball in an arc towards you, the approaching ball would seem to grow alarmingly, changing its size in an unsettling and confusing manner.

Yet, amid all of this complex sensory information reaching our brains, a critical stability persists. Your mind is able to anchor all these shifting inputs to a constant version of reality. Size constancy steps in, ensuring the ball maintains a consistent identity, regardless of retinal size or distance. Orientation constancy plays an indispensable role too, assuring us that we’re the ones that are moving, and that the ground remains flat. Despite movements, your perception remains constant. What could be a disorienting, unpleasant experience transforms into a coherent, enjoyable scene. The power of perceptual constancy simplifies the world, making it a straightforward, consistent experience.

This thesis will delve into one key element of perceptual constancy. We will focus on size constancy and the neural processes in the primary visual cortex that underlie it. These inquiries aim to further our understanding of the fundamental mechanisms shaping our perception of reality, and bring us one step closer to understanding why we like lounging in hammocks and playing ball games.

Zusammenfassung

Einleitung

Wahrnehmungskonstanz bezeichnet unsere Fähigkeit, stabile Eigenschaften der Außenwelt korrekt wahrzunehmen, obwohl die entsprechenden sensorischen Signale variieren. Größenkonstanz bezeichnet insbesondere die Fähigkeit unseres visuellen Systems, die Größe eines Objekts trotz Veränderungen des retinalen Bildes korrekt wahrzunehmen, wenn sich unser Abstand zu ihm verändert. Diese Konstanz ermöglicht eine stabile Wahrnehmung der physischen Welt um uns herum und erlaubt es uns, wesentliche Handlungen auszuführen, die wir als selbstverständlich ansehen mögen, wie etwa das Fangen eines Balls, das Abschätzen der Breite einer Türöffnung oder die Beurteilung, ob ein sich näherndes Tier eine Katze oder ein Tiger ist.

Ein Großteil der Literatur zur Größenkonstanz und ihren neuronalen Grundlagen verwendet statische Darstellungen, wodurch eine Lücke in unserem Verständnis dessen bleibt, was in alltäglichen Situationen geschieht, in denen sich die Distanz zwischen einem Objekt und dem Beobachter verändert. Die übergeordnete Frage dieser Dissertation betrifft daher die Neurophysiologie der Größenwahrnehmung unter Bedingungen, in denen sich Objekte bewegen.

Motivation und Forschungsfrage

Die Literatur zu bewegten Reizen ist uneinheitlich. Unter reduzierten Hinweisreizen kann Bewegung die Größen- und Distanzwahrnehmung beeinträchtigen und eine stärkere Abhängigkeit von der retinalen Größe fördern. In reicheren Szenen gibt es Hinweise darauf, dass Bewegung helfen kann, Größenkonstanz aufrechtzuerhalten, jedoch verwenden Studien häufig mehrdeutige Reize oder illusorische Tiefe. Die Mehrdeutigkeit früherer Arbeiten motiviert Untersuchungen in Umgebungen mit reichen Tiefenhinweisen, die reale Betrachtungsbedingungen nachahmen, während gleichzeitig gemessen wird, wie sich neuronale Repräsentationen in frühen visuellen Arealen verhalten, wenn sich Reize in der Tiefe bewegen und dabei entweder die physische Größe konstant oder die retinale Größe konstant gehalten wird.

Die in dieser Dissertation beschriebenen Experimente untersuchen die Neurophysiologie der Größenwahrnehmung unmittelbar vor und nach Bewegung, wenn sich der Reiz zuvor in einer an Tiefenhinweisen reichen Umgebung bewegt hat. Um die Aktivität im frühen visuellen Kortex und insbesondere die Größe von Objektrepräsentationen in diesen

Arealen zu untersuchen, verwenden wir steady-state visuell evozierte Potentiale (SSVEPs), gemessen mithilfe von Elektroenzephalographie (EEG), eine Methode mit nachgewiesener Wirksamkeit. Diese Experimente werden sowohl in Virtual Reality (VR) als auch in einem physischen Aufbau durchgeführt, der dieselbe Geometrie und dieselben Tiefenhinweise aufweist. Die Experimente dieser Dissertation untersuchen, ob der primäre visuelle Kortex weiterhin physische Größe kodiert (wie unter statischen Bedingungen beobachtet) oder ob er bei vorhandener Bewegung zu einer rein retinotopen Repräsentation zurückkehrt.

Konzeptuelle Fortschritte

Über aufeinander abgestimmte Experimente in VR und unter Verwendung eines physischen Aufbaus hinweg finden wir dasselbe neurowissenschaftliche Ergebnis, nämlich dass der frühe visuelle Kortex für Objekte, die sich gerade in der Tiefe bewegt haben, keine größenkonstante Kodierung zeigt. Die SSVEP-Antworten spiegeln nach Bewegung in beiden Versionen der Aufgabe die retinale Größe wider. Demgegenüber zeigt ein Kontrollexperiment mit statischen, aber ansonsten vergleichbaren Reizen partielle Größenkonstanz der Objektrepräsentationen im frühen visuellen Kortex und repliziert damit zuvor veröffentlichte Befunde. Der konzeptuelle Fortschritt besteht in einer klaren Dissoziation zwischen dynamischen und statischen Bedingungen. Nach Bewegung repräsentiert der frühe visuelle Kortex das betrachtete Objekt retinotop, während unter vollständig statischen Bedingungen partielle Größenkonstanz in denselben Hirnregionen auftritt.

Diese Dissoziation stellt einen zentralen konzeptuellen Fortschritt dar. Sie legt nahe, dass Bewegung das Gleichgewicht zwischen feedforward- und feedback-Verarbeitung im frühen visuellen Kortex verändert und es möglicherweise zugunsten von feedforward-Verarbeitung auf Kosten feedback-basierter Einflüsse verschiebt. Das Ergebnis ist, dass dieselbe Kortexregion (V1), die unter statischen Bedingungen Größenkonstanz unterstützt, dies unmittelbar nach Bewegung nicht tut — selbst dann, wenn Verhaltensantworten eine korrekte, größenkonstante Wahrnehmung anzeigen. Das bedeutet, dass größenkonstante Repräsentationen in V1 keine feste Eigenschaft sind, sondern dynamisch durch Betrachtungsbedingungen moduliert werden.

Die Ergebnisse des statischen Kontrollexperiments zeigen außerdem, dass unser SSVEP-Ansatz sensitiv für größenkonstante Signale ist, wenn diese vorhanden sind, und demonstrieren die Replizierbarkeit früherer Befunde dieses Effekts. Unsere Befunde stehen daher mit hoher Wahrscheinlichkeit im Zusammenhang mit Bewegung und sind nicht auf die spezifisch verwendete visuelle Umgebung oder ein Ausbleiben der Replikation früherer Befunde zurückzuführen.

Eine mögliche Erklärung für diese Befunde ist, dass unter dynamischen Betrachtungsbedingungen die feedforward-Verarbeitung im frühen visuellen Kortex gegenüber feedback-Prozessen dominiert, zumindest in den kortikalen Schichten, die am stärksten zum SSVEP beitragen. Unsere Befunde machen deutlich, dass die zeitliche Dynamik und die lamina-spezifische Ausprägung von feedforward- und feedback-Einflüssen im frühen Sehen besser verstanden werden müssen — insbesondere unter ökologisch validen, dynamischen Bedingungen.

Methodische Beiträge

Ein zweiter wesentlicher Beitrag ist methodischer Natur. Diese Arbeit verbindet eine sorgfältig kontrollierte VR-Umgebung mit einer eigens entwickelten Apparatur, die einen physischen Monitor entlang der sagittalen Achse bewegt, und gleicht Reizgeometrie, Timing und Tiefenhinweise zwischen beiden an. Dieses Design ermöglicht es, die ökologische Validität der VR-Umgebung sowie die Datenqualität in VR zu prüfen, ohne das Paradigma wesentlich zu verändern.

Daraus ergeben sich zwei zentrale Einsichten:

- Erstens zeigt sich eine hohe Verhaltensvalidität. Trotz in der Literatur berichteter Unterschiede zwischen VR und realen Bedingungen in Größen- und Distanzurteilen zeigt sich, dass VR ein geeignetes Medium für Experimente ist, sofern die virtuelle Szene reich an Hinweisreizen ist und den physischen Aufbau eng nachbildet. In den Verhaltensaufgaben sind die Unterschiede zwischen VR und dem physischen Aufbau minimal, mit einer durchschnittlichen Differenz von weniger als einem Millimeter.
- Zweitens ist die EEG-Signalqualität in VR deutlich reduziert. Die Interferenz des Headsets mit den Elektroden führt zu einem merklich geringeren Signal-Rausch-Verhältnis. Obwohl VR eine unvergleichliche Kontrolle über komplexe dynamische Szenen bietet, erfolgt dies auf Kosten der EEG-Datenqualität und erfordert sorgfältige Beachtung der Signalintegrität.

Die Dissertation dokumentiert außerdem VR-spezifische methodische Herausforderungen, die für das experimentelle Design relevant sind. Distanzen werden in Szenen mit wenigen Hinweisreizen häufig überschätzt. Um den SSVEP-Kontrast stabil zu halten, können einige Hinweisreize (wie Schattierung) nicht verwendet werden, sodass Szenen sorgfältig gestaltet werden müssen, um ausreichende Tiefeninformation zu gewährleisten. Der Vergleich von VR mit einem physischen Gegenstück hilft zu validieren, dass diese Designentscheidungen die interessierenden Effekte nicht verzerren.

Insgesamt liefert diese Arbeit Unterstützung für den Einsatz von VR in neurowissenschaftlichen Experimenten in den Sehwissenschaften.

Weiterreichende Implikationen

Die weiterreichenden Implikationen dieser Befunde bestehen darin, dass neurowissenschaftliche Ergebnisse zur Wahrnehmungskonstanz, die unter statischen Bedingungen gewonnen wurden, nicht ohne Weiteres auf ökologisch validere, dynamische Kontexte übertragen werden können. Während diese Unterscheidung hier speziell für Größenwahrnehmung gezeigt wird, ist zu beachten, dass auch andere Formen der Wahrnehmungskonstanz (wie Form-, Farb- oder Helligkeitskonstanz) häufig unter statischen Bedingungen untersucht werden, obwohl sie in realen Situationen in dynamischen visuellen Umgebungen auftreten.

Obwohl die genauen zeitlichen Eigenschaften dieses Übergangs noch weiter untersucht werden müssen und die vorliegenden Daten nicht direkt belegen, dass dieser Übergang durch eine Verschiebung im Gleichgewicht zwischen feedforward- und feedback-Prozessen verursacht wird, motivieren sie zukünftige Experimente, die empfindlicher für kortikale Schichten und zeitliche Dynamiken sind. Solche Arbeiten könnten ein klareres Bild der neuronalen Grundlagen der Größenwahrnehmung im Besonderen und der visuellen Wahrnehmung im Allgemeinen liefern.

Letztlich legt diese Dissertation ein Umdenken in den Wahrnehmungsneurowissenschaften nahe, das die Komplexität und Dynamik des Sehens in der realen Welt stärker berücksichtigt.

Executive summary

Introduction

Perceptual constancy refers to our ability to correctly perceive stable properties of the outside world, despite the fact that their corresponding sensory signals are in flux. *Size constancy* in particular refers to our visual system's ability to correctly perceive an object's size despite the fact that the retinal image may shrink or grow as our distance to it changes. This constancy affords us a stable impression of the physical world around us, and allows us to perform essential actions that we may take for granted like catching a ball, estimating the width of a doorway and being able to determine whether an approaching animal is a cat or a tiger.

Much of the literature on size constancy and its neural underpinnings makes use of static displays, which leaves a gap in our understanding of what happens during ordinary, everyday episodes where the distance between an object and the observer changes. The overarching question of this thesis therefore relates to the neurophysiology of size perception when things are moving.

Motivation and research question

The literature on moving stimuli is mixed. Under impoverished cues, motion can hinder size and distance perception and encourage reliance on retinal size. In richer scenes there is evidence that movement can help maintain size constancy, but studies often use ambiguous stimuli or illusory depth. The ambiguity in prior work motivates testing in depth cue rich environments that mimic real life viewing conditions, while measuring how neural representations in early visual areas behave when stimuli move in depth with either physical size held constant or retinal size held constant.

The experiments described in this thesis address the neurophysiology of size perception right before and after movement, when the stimulus has recently been in motion in a depth cue rich environment. To probe the activity in the early visual cortex, and more specifically the size of object representations in these areas, we will make use of steady state visual evoked potentials (SSVEP) measured via electroencephalography (EEG), a method with proven effectiveness. These experiments will take place both in Virtual Reality, (VR), and in a physical setup that preserves the same geometry and depth cues. Experiments in this thesis will examine whether the primary visual cortex continues to encode physical size (as it does under static conditions), or whether it reverts to a purely retinotopic representation when motion is involved.

Conceptual Advances

Across matched experiments in VR and using a physical setup, we find the same neuroscientific result, namely that for objects that have just moved in depth, early visual cortex does not show size constant encoding. The SSVEP responses reflect retinal size after motion in both implementations of the task. By contrast, a control experiment using static but otherwise comparable stimuli shows partial size constancy of object representations in early visual cortex, replicating previously published findings. The conceptual advance is a clear dissociation between dynamic and static conditions. Following motion, early visual cortex represents the fixated object in a retinotopic way, whereas in a completely static condition, partial size constancy emerges in these same brain regions.

This dissociation marks a key conceptual advance. It suggests that motion alters the balance between feedforward and feedback processing in early visual cortex, perhaps tipping it in favor of bottom-up input at the expense of perceptual (top-down) tuning. The result is that the same region of cortex (V1) that supports size constancy under static conditions fails to do so immediately after motion—even when behavioural responses indicate correct, size constant perception. This means that size-constant representations in V1 are not a fixed property, but dynamically modulated by viewing conditions.

The static control experiment's results also show that our SSVEP approach is sensitive to size constant signals when they are present, and demonstrates the replicability of earlier findings of this effect. This is to say that our findings are most likely related to motion rather than due to the specific visual environment being used or a failure to replicate earlier findings.

A potential explanation for these findings is that during more dynamic viewing conditions, bottom up processing in early visual cortex dominates over feedback processing, at least in the cortical layers that contribute most to the SSVEP. Our findings highlight a need to understand the temporal dynamics and laminar specificity of feedforward and feedback influences in early vision—especially under ecologically valid, dynamic conditions.

Methodological contributions

A second major contribution is methodological. This work pairs a carefully controlled VR environment with a custom apparatus that moves a physical monitor along the sagittal axis, and equates stimulus geometry, timing and depth cues across the two. This design makes it possible to test the ecological validity of the VR environment and data quality in VR without substantially changing the paradigm.

Two insights follow:

1. Behavioral validity: despite differences between VR and real life in size and distance judgements reported in the literature, we find VR to be an appropriate medium for experimentation, provided that the virtual scene is cue rich and closely mirrors the physical setup. In the behavioural tasks, errors are essentially identical between VR and the physical setup, with an average difference under one millimetre.

2. EEG signal quality: SSVEP data quality is much reduced in VR. Headset interference with electrodes results in remarkably worse signal-to-noise ratios. Although VR offers unmatched control over complex dynamic scenes, this comes at the cost of EEG data quality and necessitates careful attention to signal integrity.

The thesis also catalogues VR specific pitfalls that matter for experimental design. Distance is often overestimated in sparsely cued scenes. To keep SSVEP contrast stable, some cues (like shading) cannot be used, so scenes must be constructed with careful anchoring to preserve depth information. Comparing VR to a physical counterpart helps validate that these design choices do not distort the effects of interest.

All things considered, this work finds support for the use of VR in neuroscientific experiments in vision sciences.

Broader implications

The broader implication of these findings is that we should not assume neuroscientific findings related to perceptual constancy which are collected from experiments using static conditions generalize to more ecologically valid, dynamic contexts. While we demonstrate this distinction for size perception in particular, it should be noted that other forms of perceptual constancy (such as shape, colour or brightness constancy) tend to be studied in static conditions, while in real world viewing, these mechanisms emerge in dynamic visual environments.

Although the precise temporal characteristics of this transition remain to be mapped, and our data do not by themselves prove that this transition is caused by a shift in feedforward versus feedback balance. They do, however, motivate future experiments that are sensitive to more specific cortical layers and temporal dynamics, which could give us a clearer picture of the neural underpinnings of size perception specifically and visual perception more broadly. Ultimately, the thesis invites a rethinking of perceptual neuroscience: one that more fully embraces the complexity and dynamism of real-world vision.

Table of contents

General introduction: perceptual constancy

Like the little story in the preface illustrates, perceptual constancy is one of the most fundamental and useful properties of the sensory system. It allows us to correctly identify constant properties of the outside world despite the ever changing nature of the sensory input that informs us about it. In other words, it allows us to see things for what they are. As a concept, perceptual constancy is not restricted to any one sensory modality. One might even argue that insofar as perception as a whole is a multisensory phenomenon, so is perceptual constancy, as all senses that contribute to our impression of the outside world are in a near-constant state of flux.

Still, many specific examples of perceptual constancy relate to one sensory domain, or at least they are studied with respect to a given sensory domain. Shape constancy, for instance, though theoretically multisensory may be studied visually by presenting different perspectives of a single object (e.g. Lawson, 1999) or haptically, where subjects are asked to identify distinct objects by touch alone (Klatzky, Lederman & Metzger, 1985). Testing perceptual constancy in a domain-specific way allows us to learn more about that sensory system, and likewise in this thesis we will attempt to learn more about the visual system by studying size constancy in a purely visual way.

Generally, perceptual constancies are defined as the ability to perceive a property (colour, shape, size etc) despite *changes* in the sensory input. I would like to focus for a moment on the nature of this change. For methodological reasons, experiments often compare distinct conditions of a stimulus. For instance, McKenzie, Tootell & Day (1980) presented objects of various sizes at various distances to infants in an attempt to probe their size constancy. Infants prefer to look at objects that are new to them, and don’t look as much at objects that have been presented several times already. By studying this looking behaviour, researchers are able to infer whether an infant is able to perceive objects in a size constant manner. If a new object is presented at a distance at which it subtends the same retinal size as a previously habituated object, and the infant looks at it in a manner similar to a new object, we may conclude that the infant is able to extract physical size from retinal size and a distance estimate, and size constancy is achieved. McKenzie et al. (1980) found that 6 and 8 month old infants were able to perceive objects in a size constant manner (at least up to 70 cm distance), but 4 month olds were not.

Indeed, presenting a given object at different distances at different times results in a change in the resultant retinal size of the observing infant during the different viewing opportunities.

However, this change being of such a discrete nature is perhaps not reflective of the type of change that is most commonly encountered in real life. Usually changes in retinal size as a result of changes in distance to an object are continuous. That this difference in the kind of change is important was demonstrated in a later experiment. When Day & McKenzie (1981) presented continuously moving stimuli, 4 month olds *did* demonstrate an ability to perceive objects in a size constant manner. The more common method of comparing discrete conditions can apparently lead to different conclusions than the more realistic continuous presentation of changing stimuli.

Continuous sampling of our environment is an essential element of perception in everyday conditions. This is not to say that discrete observation moments are inherently unrealistic, as we often take quick glimpses in our rearview mirrors when driving and draw conclusions about the distance, speed and size of vehicles behind us. Still, the fact that size perception specifically is affected by the method of stimulus presentation (continuous / dynamic versus discrete / static) is an important thing to take note of. Moreover, the work in this thesis demonstrates that there is a neurophysiological difference in the way size information is processed depending on exactly this distinction.

Size constancy

Theories of human size perception date back to antiquity. Ptolemy, in his *Optics,* describes how the eye receives an image of objects, and that this retinal size is a function of the object's physical size and its distance to the observer (Smith, 1996). One of the earliest discussions on how this retinal size gives rise to our conscious perception of an object's size however, was done by Helmholz (1867). He suggested that depth cues are unconsciously processed, and this estimate of distance is used in conjunction with the retinal size to calculate the physical size of an object. This concept was further explored by Emmert, who demonstrated that the perceived size of an afterimage scales with the distance to the surface that it appears to be projected upon (1881). This relation between size and distance, later called Emmert's law, still makes for an interesting demonstration of the phenomenon, and can be experienced by staring at a bright light for some time, ideally in a dimly lit room, and then looking around at various surfaces. The way the afterimage seems to shrink and grow when looking at a near and far surface is striking. Experiencing this yourself, it is hard to deny that an interplay between the retinal size and the perceived distance to the object plays a huge role in size perception.

A formalisation of the basic principle first described by Ptolemy is the size-distance invariance hypothesis (Epstein, Park & Casey, 1961) , which is still commonly referred to in size perception research, and which posits that perceived size is a function of the retinal size and the perceived distance to the target. It further posits that perceived distance and perceived size must co-vary.

This is, of course, not the full story, and the size-distance invariance hypothesis is not an infallible rule that explains every aspect of human size perception. Many other factors besides just perceived distance influence the perceived size of objects and the success of size constancy in our perception. Size constancy, in naturalistic viewing conditions works quite well, though with a tendency towards overestimations (Wagner, 2012). A size matching task between near objects and far objects, revealed that at larger distances (~30 meters and over) *overconstancy* (overestimation of the physical size of the object) started to take over (Gilinsky, 1955), at least when the instructions are to match the physical size of the target object. In an attempt to see where and how size constancy breaks down, Holway and Boring (1941) gradually restricted vision. Whereas stimuli viewed binocularly, with free head movement and in a feature-rich environment were perceived in a size constant manner, once all these conditions were restricted (monocular viewing through a peephole with the hallway draped in black cloth) observers reverted to reporting the retinal size of the stimuli, displaying *underconstancy* (underestimation of the physical size of the object). Subsequent research that reduced available depth information even more thoroughly reinforced the finding that under restricted depth cue conditions, underconstancy is common (Lichten & Lurie, 1950).

The variables that influence the degree of size constancy are legion, and range from stimulus properties to depth cue salience and even experimental instructions. Modern understandings of size constancy emphasize that various cognitive factors play a role in size perception, making the final percept a function of bottom up sensory information (weighted according to reliability) and top down cognitive factors (Kersten, Mamassian & Yuille, 2004; Sperandio & Chouinard, 2015). A particularly important top down fact is that of known size (Ittelson, 1961). We generally encounter objects we are familiar with, or at least objects from categories that we are familiar with (even if we see a previously unknown cat, we have a good estimate of the average size of cats). In the absence of any other depth cues, known size can be relied on to provide an estimation of object size, and by extension, distance. It is important to note that, especially since the experiment on size constancy in this thesis were mostly performed in virtual reality (VR), known size as a cue is especially heavily weighted in VR (Rzepka, Hussey, Maltz, Babin, Wilcox & Culham, 2022).

Size constancy in dynamic conditions

Like stated before, much of our everyday visual experience involves continuous observation of the visual world around us, and often involves motion of either ourselves or the things around us. It appears as though our visual system is primed to deal with motion. Johansson (1964) presented observers with expanding and contracting luminous squares, presented in visually impoverished conditions. This stimulus was ambiguous, in that it could be interpreted as growing and shrinking, or moving towards and away. Observers generally interpreted the stimulus as moving, indicating a strong prior to perceive motion of a static object instead of changes in size. This finding was found again when using an even simpler, 3-dot formation that expanded and contracted (Borjesson & von Hofsten, 1973).

Whereas the above experiments used an artificial stimulus that did not physically move, others used actual movement. Early work using real motion produced conflicting results. Brosgole, McNichol, Doyle & Neylon (1976) report that when a stimulus is moving, observers rely more on retinal size, leading to underconstancy. It should be noted that the stimulus used here was still quite ambiguous, as it consisted of two illuminated ping pong balls presented in an otherwise dark environment. Some years later, Noguchi, & Taya (1981) report the opposite, claiming that size constancy is maintained when viewing moving stimuli, even in impoverished conditions where one would expect size constancy to be worse, indicating that motion can serve as a useful cue to size and distance. Notably, Noguchi & Taya (1981) do not refer to Brosgole et al. 's (1976) earlier work, nor does later work by Brosgole (1993) refer to Nougichi & Taya (1981).

A more theoretical approach to the concept of size constancy in dynamic conditions was taken by Herschenson (1992). He posited that size constancy in dynamic conditions is fundamentally different from that during static conditions. He argued for the existence of a ‘rigidity constraint’, whereby objects are perceived as rigid, having a constant size, and thereby interpreting any change in retinal size as a change in the distance to the object, maintaining a constant representation of the object’s size. This is different from the classic size-distance invariance hypothesis. Whereas the size-distance invariance hypothesis posits we infer physical size from retinal size and perceived distance, Herschenson’s kinetic size-distance invariance hypothesis claims distance is inferred from a given, rigid size and a change in retinal size. He further claims that any kind of size perception during static conditions is essentially a momentary snapshot of the principles at work during dynamic size perception, and that dynamic size perception is the default context for our visual system.

His own experiments also showed that the kinetic size-distance invariance hypothesis does not always apply, as although ambiguous stimuli would often be interpreted as moving, sometimes they would be interpreted as changing size, or moving and changing size concurrently (Herschenson, 1992, 1993). In conclusion, although Herschenson described the kinetic size-distance invariance hypothesis as an automatic assumption of our perceptual system, it is clear that we are perfectly capable of perceiving objects whose retinal size, physical size and distance are changing concurrently veridically (think of watching a deflating balloon flying through the air). It is therefore more useful to interpret the kinetic size-distance invariance hypothesis as referring to a strong prior that our visual system has, and by no means prescriptive of our perception. Several of the experiments in this thesis demonstrate our ability to discern concurrent size change from movement, although not always perfectly.

A different theory was proposed by Gogel & Da Silva (1987a); they propose a two-process theory of size and distance perception. The two processes involved are the primary one, dealing with bottom-up sensory information and informed by the size-distance invariance hypothesis, and the secondary one dealing with more cognitive factors like known size. These two processes are thought to interact, and one might be more prominent than the other depending on the availability of the sensory information available (Gogel & Da Silva, 1987b). Though the authors do not explicitly mention Bayesian weighting of cue reliability, the same logic is implied in their theory. When it comes to dynamic size perception, Gogel notes that size and distance estimates must co-vary with estimates of motion, and that there is a tradeoff between judging an object as moving, judging it as changing size or combination of these two things (Gogel, 1998). Our findings in experiment 3 of this thesis support this theory.

All of the above described experiments that study motion use stimulus motion, and not motion of the observer. That this is an important distinction was demonstrated by Combe & Wexler (2010). They showed that size constancy is better when the change in distance to the stimulus is due to observer motion as opposed to motion of the stimulus. This makes sense, because self initiated movement has the added benefit of informing the observer about the degree of motion via proprioceptive means. This finding further highlights that motion can be a very valuable cue to size constancy in more naturalistic circumstances.

Optical illusions

Recent work by Ryan Mruczek has emphasized the role of motion in well established visual illusions. If we agree with Herschenson’s claim that static size constancy is really just a special case of the more important, overarching dynamic size constancy, we might be more interested in how motion affects these long standing illusions than in their static counterparts. The Ebbinghaus illusion, for instance, may be greatly enhanced by motion of either the stimulus or the observer's eyes (Mruczek, Blair, Strothers & Caplovitz, 2015). The greater illusory effect, and thereby less veridical perception, is attributed to the increased uncertainty that the concurrently moving stimulus and inducers create. When applied to the corridor illusion, however, motion decreases illusion magnitude (Mruczek, Blair, Cullen & Caplovitz, 2020). In this version, the stimulus travels up and down the corridor while retaining its retinal size. It appears that the combination of changing target and changing context greatly enhances illusory effects, whereas only changing targets or only changing context do not, or even reduce the illusion magnitude (Mruczek, Fanelli, Kelly & Caplovitz, 2022).

Looming and receding

Not all movement is created equal, and there is behavioural and neuropsychological evidence that movement along the sagittal axis is interpreted differently depending on its direction. Movement towards the observer, also known as looming motion, captures attention more so than receding motion (Franconeri & Simons, 2003). Further, Vagnoni, Lourenco & Longo (2015) demonstrated that the perceived threat and speed of approach modulate ERP components, with an interaction existing between these two factors. Given that the motion in our experiments is always along the sagittal axis, this is an important concept to keep in mind going forward.

In the size perception related experiments reported in this thesis, motion will be a constant topic of interest. Our findings will support the notion that size perception in dynamic conditions, even when the stimulus actively changes size, does not significantly improve or impede accuracy.

In conclusion, motion has the capacity to improve or impede size perception. If motion enhances uncertainty like in Mruczek's illusions or occurs in an ambiguous, simulated way, it can impede size constant perception. However, if motion occurs in a way that is naturalistic, it can act as a powerful cue that enhances accurate, size constant perception. When it comes to scientific significance, I would claim that size constancy and its neural mechanisms studied in dynamic conditions are at least as important as those studied in static conditions, if not more so, due to the nearly ever present nature of motion in naturalistic, everyday vision.

The neuroscience of size perception

Our (mostly) accurate, size constant perception of the outside world is the result of a dynamic process in our brains that involves a number of critical regions interacting with each other in a complex way. Although the focus of our work will be on the primary visual cortex, this is only one node in the greater network that gives rise to size constant visual perception. This network includes both the ventral and dorsal visual processing streams, each providing their own unique contributions. The primary visual cortex both sends and receives information from these streams, and is an essential part of the size perception network.

One might think that the primary visual cortex is an unlikely candidate for size constant perception. After all, V1 has long since been shown to have a retinotopic organisation (Inouye, 1909), displaying size in accordance with the retinal impression of an object multiplied by the cortical magnification of the relevant portion of the visual field (Engel, Glover, & Wandell, 1997; Adams & Horton, 2003; Duncan & Boynton, 2003). Meanwhile, many processes indirectly involved in size perception, especially those relating to pictorial cues, occurs in extrastriate areas (Ban & Welchman, 2015; Cumming & Parker, 2000; Parker, 2007). However, a pivotal experiment using fMRI showed that object representations in V1 scaled with perceived physical size (Murray, Boyaci, & Kersten, 2006). Using a Ponzo / hallway illusion, where two retinally identical stimuli are presented at different locations in an image of a hallway, these authors showed that the further stimulus which is perceptually larger also elicits a larger representation in V1. Other neuroimaging studies using illusions have reinforced this finding, with the neural representation of the larger line in the Müller-Lyer illusion also appearing larger in V1 (Ho & Schwarzkopf, 2022). Additionally, the earlier described Emmert's law also affects V1 activity (Sperandio, Chouinard & Goodale, 2012).

Afterimages that appear physically larger when projected upon further surfaces also coincide with larger representations in V1, hinting at a coupling between perceived size and early visual cortex activity.

Illusion magnitudes of size and distance related illusions further seem to be closely tied to the anatomical makeup of the primary visual cortex. Individuals with smaller cortical surface areas of V1 experience greater illusion magnitudes when viewing Ponzo and Ebbinghaus illusions (Schwarzkopf, Song & Rees, 2011). Also in the absence of size and distance illusions, and even in the absence of any contextual cue whatsoever, individual differences in V1 receptive fields predict differences in size perception (Moutsiana et al., 2016). This enforces the idea that the very anatomy of V1 is fundamental to size perception.

Still, V1 does not produce size constant representations on its own. Depth information is processed and extracted in brain regions beyond V1, and are then fed back to V1 where the size representation is updated. Chen, Sperandio, Henry & Goodale (2019) demonstrated the time course of size constancy in the visual cortex. The distinction between retinally equal, physically different stimuli becomes observable in ERPs at ~150ms, indicating that later feedback signals from higher order visual areas are required for size constancy to be established.These higher order visual areas include the lateral occipital sulcus (LOS) and the caudal inferior temporal gyrus (caudal ITG)(Grill-Spector, Kourtzi & Kanwisher, 2001). These areas are important in extracting shape from 3D information, and thereby extract important depth cues from the visual environment, which are consequently used to scale representations of objects (Georgieva, Todd, Peeters & Orban, 2008). Another one of these higher order visual areas that is important for size perception is the lateral-occipital cortex (LOC). Weidner & Fink (2007) showed that the strength of the Müller-Lyer illusion, where a horizontal line can appear larger or smaller based on the direction of surrounding arrows, was correlated with LOC activity. Further, neurophysiological evidence from experiments using monkeys shows that their LOC-analogous area shows firing rates that correspond to physical size, regardless of current retinal size (Vighneshvel & Arun, 2015). This makes the LOC a likely candidate for a size-invariant region, which represents objects in a size constant way. Further still down the ventral stream, representations of object categories can be distinguished by physical size (Konkle & Oliva, 2012). These authors showed, in an fMRI study, that there is a medial-to-lateral gradient in the ventral temporal cortex along which objects are represented on a large-to-small scale; bathtubs and tables elicit more medial activity, whereas paperclips and strawberries elicit more lateral activity.

The LOC provides feedback information via cortico-cortical connections to V1 (Pennartz, Dora, Muckli & Lorteije, 2019), and consequent V1 representations are linked to conscious visual perception (Chen, Weidner, Zeng, Fink, Müller & Conci, 2021). A study using TMS highlighted this order of events. Disruption of the LOC by TMS stimulation at an earlier time (around 150-200ms after stimulus onset) caused disruption of size perception, and disruption of V1 at a later time (200-250ms) caused disruption of size perception (Zeng, Fink & Weidner, 2020). These findings suggest that information processing essential for size perception occurs in LOC, and is then sent back to V1. Additionally, it demonstrates that this size constant representation in V1 is no epiphenomenon, and is directly linked to how we perceive the size of objects. More support for the notion that there is a direct link between V1 activity and conscious visual size perception comes from Wang, Chen & Jiang (2021). These authors demonstrated that anodal transcranial direct current stimulation (tDCS) to early visual areas enhances the illusory size effect of the Ebbinghaus illusion.

Although the dorsal stream may be thought to be more related to perception for action (Goodale & Milner, 1992; Cavina-Pratesi, Goodale & Culham, 2007), it does play an important role in purely visual size perception. Recent work by Chen, Wu, Yu & Sperandio (2024) indicated that increased feedback activity from the precuneus, in the parietal lobe, to V1 was associated with illusory effects of the Ebbinghaus illusion; increased functional coupling from the precuneus as well as decreased inhibitory activity within the occipital lobe were associated with larger illusion magnitudes. This again reinforces the notion that dynamic feedback processes from higher order visual areas affect size perception in early visual areas, and highlights the many regions that are involved in achieving size constant perception.

Disruptions in the brain regions mentioned above may be expected to lead to inappropriate size perception. Alice in Wonderland Syndrome (AIWS) is a perceptual disorder that results in a disruption of normal size perception, with the patients own body or objects around them seeming strangely large or small. This syndrome is relatively rare, with a meta-analysis only finding 169 reported cases in the literature (Blom, 2016). Lacking any neuroimaging data for many of these cases, physicians have previously speculated as to the possible causes of the perceptual symptoms. Recently, however, a study was able to collect 37 cases of patients with lesions suffering from AIWS for which neuroimaging data was available (Friedrich et al., 2024). The authors found that although the exact location of the lesions was relatively heterogenous, in 97% of cases the lesions fell within a network of brain regions related to size perception.

A strong, negative connection existed in this network between the locus of the lesion and the inferior parietal lobule, an area which is linked to object size perception and size constancy (Harvey, Fracasso, Petridou & Dumoulin, 2015; Kristensen, Fracasso, Dumoulin, Almeida & Harvey, 2021).

The focus of the neurophysiological inquiries in this thesis will be on the primary visual cortex. Like the studies discussed above demonstrate, V1 plays a very important role in the network that facilitates size constant perception. Being the first cortical site of visual information processing, It both transmits information in a feedforward manner for further processing and receives information in a feedback manner, which then causes representations in V1 to occur in a size constant manner. However,V1 is not necessarily the first link in this long chain of feed-forward and feedback size- and distance processing. Information reaching V1 from earlier in the visual processing stream can modulate V1 activity. Vergence information has been shown to alter retinotopic organisation of V1 (Cottereau, Durand, Vayssiere & Trotter, 2014). Further, V1 is not a mere relay station for depth information, but processes size and distance information by encoding binocular disparity (Cumming & DeAngelis, 2001). In all, although earlier theories might have given the impression that V1 is a mere projection screen of the retina, a more modern understanding describes it as a central locus in the brain network that produces size constancy, and an area that plays a direct, measurable role in our subjective size perception.

The experiments in this thesis aim to further probe the role of V1 in size perception, focussing on dynamic conditions, and using stimuli that move, shrink and grow.

Introduction to the SSVEP

Purkinje, in 1818, published “Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht”, in which he makes note of many interesting visual phenomena. Mostly acting as his own experimental subject and using introspection to come to his conclusions, he reports on things like phosphenes experienced upon pinching his own eyeballs, and afterimages experienced after brief exposures to a lit candle in a dark room. He also discusses his experience of gazing at a spinning wheel, inspired by Newton’s colour disk but with black and white segments instead. Gazing at this spinning disk, Purkinje describes the perception of secondary, illusory shapes which are not present on the wheel itself. He concludes about this experience that: “The general requirement [...] is a rapid alternation of light and darkness within the visual field; the more sharply they alternate, the more vivid the resulting phenomenon.” This visual flickering between light and darkness forms the basis of the modern steady state visually evoked potential (SSVEP) technique. Indeed, a little less than two centuries after Purkinje’s publication, Herrmann (2001) described similar illusory shapes evoked by rapidly flickering stimuli. Though this time instead of these shapes being curiosities described in a qualitative way, the perceived shapes were able to be accurately matched to predictions made by a neuroscientific model of the early visual cortex (Tass, 1995). Even though Purkinje made no claims about the neuroscientific basis of this particular finding, we can apply his axiom just as well to the SSVEP, in that the general requirement is a rapid alternation of light and darkness, and that it increases with more sharp alterations; that’s to say higher contrast (Spekreijse, 1966).

The first experiment that made use of what we would now consider proper SSVEP methodology was performed by Adrian and Mathews (1934). Inspired by the description of alpha activity by the inventor of EEG, Hans Berger, they set out to better understand the origin of what was then known as the Berger rhythm (Berger, 1929). Their now classic findings demonstrated that the brain would show oscillatory activity in synchrony with a flickering stimulus, overriding the alpha activity that would dominate the EEG signal while at rest. Their findings also demonstrated that the alpha rhythm described by Berger was indeed of cortical origin, as opposed to having muscular origins or being purely artifactual, as was still suspected at the time. The next step in narrowing down the source of SSVEPs was made by Walter, Dovey, and Shipton (1946), who determined the point of origin was the visual cortex. This finding has stood the test of time, though more recent research has nuanced it somewhat. The source of SSVEPs is indeed thought to be the early visual cortex, with the primary visual cortex V1 being the main contributor, and the SSVEP thereby being suited as a measure to study V1 activity (Di Russo et al., 2007; Wittevrongel et al., 2018; Mohr, Geuzebroek & Kelly, 2024). There are, however, extrastriate contributors to the SSVEP which include V2, V3, V5/MT and V4/V8 (Di Russo et al., 2007). Source localisation studies combining high density EEG with individual retinotopic maps gathered with fMRI confirmed that the order of activation evoked by SSVEPs mirrors those of more regular, transient visual effects. Di Russo et al. (2007), using this method, found SSVEP activation traveling from V1 to V5/MT, V3A and V4/V8, which suggests a similar hierarchical processing to other visual stimuli.

After Walter et al. (1946), the next big step in SSVEP research came in the ‘60s. In fact, the term SSVEP first makes an appearance in this period, since the phenomenon thus far was referred to as photic driving. Van der Tweel & Verduyn Lunel (1965) published a landmark study in which they describe several important aspects of the SSVEP. For one, they reported a frequency sensitivity function that showed the strongest SSVEP amplitude to occur at frequencies between 9 and 15 Hz. At higher frequencies, SSVEP amplitude decreased, and at lower frequencies, they noted distortions in the SSVEP. More specifically, they noted occasions where the second harmonic (the frequency of twice the stimulation frequency) would have a higher amplitude than the stimulation frequency itself, indicating nonlinearities in the way the brain processed the stimulus. Van der Tweel & Verduyn Lunel (1965) also reported that SSVEPs are not strictly correlated with conscious perception, the brain was able to generate SSVEPs even at higher frequencies where the stimulus was perceived to no longer be flickering, and was instead seen as a continuously bright light (the flicker fusion threshold). A final point of interest was the fact that two flickering stimuli presented in counterphase to each eye caused a marked reduction in SSVEP amplitude, suggesting that the neural generators, at least within certain frequency bounds, summed linearly.

The frequency sensitivity function reported by van der Tweel & Verduyn Lunel (1965) was later expanded upon, with multiple peaks at which the SSVEP was strongest being identified at 10 Hz (coinciding with a common peak in alpha range), another around 18 Hz, and a third in the higher beta range, between 40 and 60 Hz (Regan, 1975). These values were always reported as the sheer amplitude of the SSVEP, and were not corrected for endogenous activity present in these frequency bands. This type of correction was first introduced by Meigen & Bach (1999), and provides a better indication of SSVEP strength. The SSVEP as reported in the work presented in this thesis always refers to this kind of corrected value. Earlier reported frequency sensitivity functions, therefore, were likely affected by endogenous frequency peaks, like that at 10 Hz in the alpha band. This means that these peaks do not necessarily reflect frequencies at which the brain is maximally receptive to repetitive stimulation. More recent studies, especially those using brain-computer interfaces (BCI) have instead reported signal to noise ratios (SNR) to describe frequencies where SSVEPs are strongest, though these have largely agreed with older findings on what frequencies work best (Wu & Su, 2014; Duart, Quiles, Suay, Chio, Garcia & Morant, 2020).

Prior to the ‘60s, the SSVEP was still of limited practical value, though this would change in subsequent years. Although Tweel & Verduyn Lunel (1965) demonstrated that the SSVEP was not a perfect corollary of conscious visual perception, with flicker frequencies that were too high to subjectively perceive still appearing as SSVEPs, some corollaries between SSVEPs and subjective perception do exist. There appeared to be a linear relationship between psychophysical measures of spatial frequency sensitivity and SSVEPs elicited by stimuli of varying spatial resolutions (Campbell & Gubisch, 1966; Campbell & Maffei, 1970). Building upon these principles, Tyler, Apkarian, Levi & Nakayama (1979) developed the sweep-VEP. This technique consisted of a flickering stimulus, kept at a constant stimulation frequency and contrast, and ramped up in spatial resolution until a point where the SSVEP would disappear against the background EEG activity. This point would be taken as a measure of maximum visual acuity, and because it required no active input from the observer, could be used to assess infant visual acuity. This seminal finding highlighted a major benefit of SSVEPs, in that it allows testing of populations that cannot participate in behavioural testing, such as pre-verbal children and various clinical populations (Gittinger & Sokol, 1982; Celesia, Meredith & Pluff, 1983; Taylor & McCulloch, 1992). This principle has also expanded beyond clinical testing of the visual system. SSVEP-based BCIs now enable individuals who can no longer communicate through speech, gestures, or writing to communicate through computer programs by selectively attending to SSVEP-inducing stimuli (Middendorf, McMillan, Calhoun, & Jones, 2000). The efficiency of these BCIs is also continually increasing; whereas the BCIs examined by Middendorf et al. (2000) allows the user to make binary choices every two seconds, Chen, Wang, Nakanishi, Gao, Jung & Gao (2015) developed an SSVEP based BCI speller that allows the user to spell words at a rate of ~12 per minute.

Another important aspect of the SSVEP discovered in the 90s is its relation to attention. Morgan, Hansen & Hillyard (1996) found that when there are two dot patterns flickering at two different frequencies, covert attention to one or the other pattern enhances the SSVEP of the corresponding flicker frequency. This attention-based enhancement of SSVEP has since been expanded upon, with studies showing that feature, location and object based attention is able to enhance the SSVEP (Pei, Pettet & Norcia, 2002; Anderson, Fuchs & Müller, 2011; Adamian, Anderson & Hillyard, 2020). These phenomena are thought to reflect enhanced firing rates of neurons in early visual cortex in response to top-down attentional modulation. These findings have further expanded the utility of the SSVEP from sensory physiology to a versatile metric to use in various cognitive neuroscience experiments.

In the work presented here, we will use the SSVEP as a metric of size representations of objects in early visual cortex. As mentioned before, the primary visual cortex (V1) is the main source of SSVEP signals. Given V1s retinotopic organisation (Dougherty, Koch, Brewer, Fischer, Modersizki & Wandell, 2003; Adams & Horton, 2003), (retinally) larger stimuli will excite more neurons in V1, resulting in a stronger SSVEP. Early work by Spekreijse and Reits (1982) showed that larger stimuli were indeed associated with stronger SSVEPs and a larger signal-to-noise ratio. More recently, Duszyk et al. (2014) investigated the role of stimulus size on SSVEP amplitude, finding a linear relationship between the two using stimuli ranging between .57° and 3.72° visual angle. This is surprising, given the anything but linear relationship between the visual field and the proportion of V1 dedicated to it. For instance, cortical magnification is such that 8 to 10% of the striate cortex is dedicated to foveal vision, which in itself only covers .01% of the retinal surface (Azzopardi & Cowey, 1993). Still, this linearity between stimulus size and SSVEP amplitude makes the SSVEP all the more suited as a metric for stimulus size.

SSVEPs’ use as a metric for size extends beyond just retinal size. Recent work was able to show partial size constancy in early visual cortex using SSVEPs by disambiguating retinal size and physical size (Chen, McManus, Valsecchi, Harris & Gegenfurtner, 2019). This finding highlights the value of the SSVEP as a way of probing cortical representations of objects, and is a great inspiration to the work presented here.

Our current understanding of the SSVEP is that it is a complex phenomenon, and its amplitude is affected by a plethora of factors. Although this gives it many possible use cases, it also means that the stimulus parameters need to be strictly controlled, making it a potentially finicky metric. Especially comparisons between experiments must be made with the consideration that any difference in the above named factors may affect the outcomes. In the conclusions of this thesis, we will delve further into which of these factors affected our results and how this affected our results.

Preliminary experiments

In the first experiment of this research project, we delved into the effects of constantly moving stimuli on size perception in the early visual cortex. The consensus in the literature indicated that, in an environment sufficiently rich in depth cues, objects would be represented in early visual cortex in a partially size constant way. The degree to which this size constancy would be present in SSVEPs is known to vary. Studying size constancy requires an experimental design that allows for disambiguation between retinal size and physical size. Using static stimuli, this can be achieved by designing conditions in such a way that they contain sets of stimuli that are identical in retinal size, but different in physical size, and vice versa. Using dynamic stimuli, this requires stimuli that are constant in physical size, as well as stimuli that are constant in retinal size as they move. The former is not particularly challenging, as this is the natural way in which most objects tend to behave. However, the laws of optics dictate that in order for an object to retain a constant angular size as the distance to the observer grows, the object must grow as well.

Although having objects grow and shrink in a systematic and controlled manner is often difficult in real life, virtual reality greatly eases this problem. For this purpose, we created a hallway in virtual reality using the unreal engine (Epic Games, 2019) in which we presented our stimuli (figure 1). The hallway was intentionally rich in depth cues, having textures walls and a textured floor, since pictorial depth cues are known to enhance size constancy (Rock, Shallo & Schwartz, 1970). The black background on which the stimulus was presented served to increase contrast, a major determinant of SSVEP strength (Allen, Norcia & Tyler, 1986). It also served as another depth cue, as a floating and otherwise featureless square could be ambiguous in its distance to the observer. With the background reaching the surrounding walls and floor, the distance to the stimulus was hopefully clear to the observer.

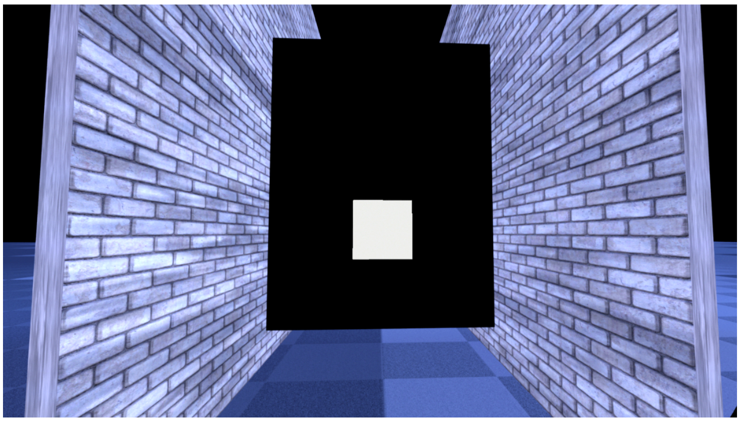


Figure 1: Virtual reality environment with stimulus for illustration purposes

Because we are interested in size perception during dynamic conditions, we needed a method to track how the relative strength of the SSVEP changed as the inducing stimulus moved and changed size. Although several methods exist, we opted for a moving-window analysis which involves taking short windows of EEG data in the time domain and extracting the relevant frequencies from these consecutive windows. For more details on this method, as well as our considerations for what parameters to use, please see Appendix A.

We hypothesized that stimulus motion could affect the representation of objects in early visual cortex in one of two ways. It could either increase the degree of size constancy present, making the representation more stable over the course of the movement, or it could decrease the degree of size constancy, coding objects according to their retinal size (figure 2).

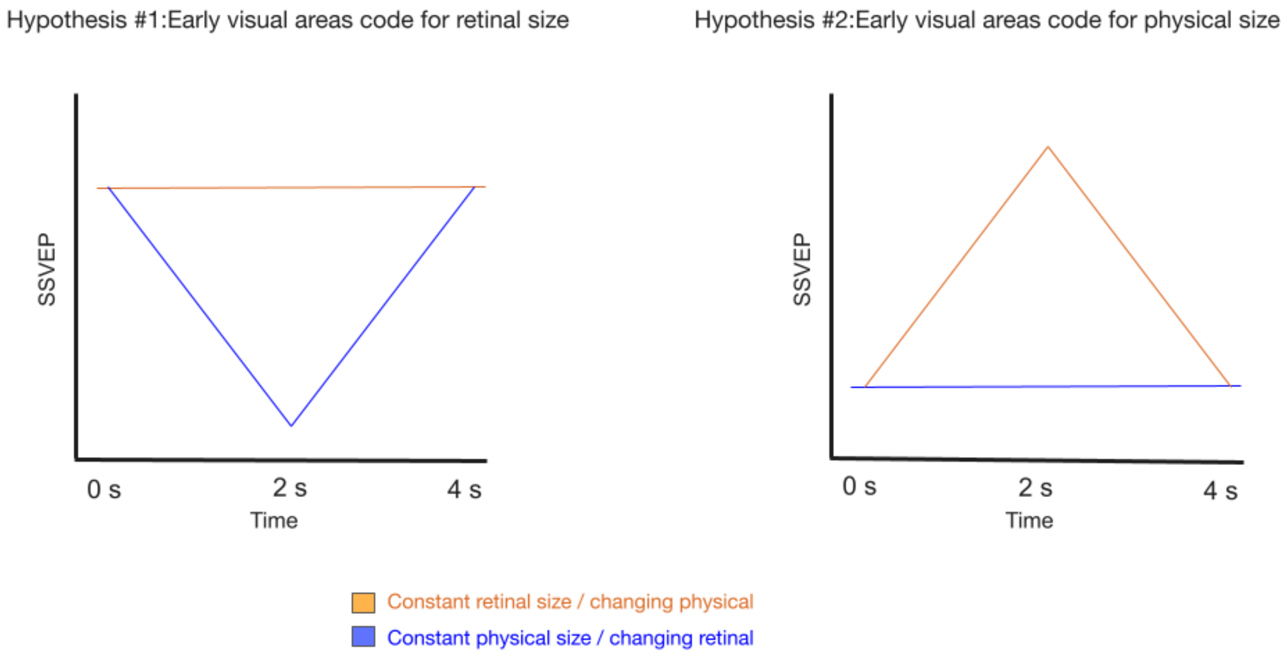


Figure 2: expected SSVEP outcomes for competing hypotheses of early visual cortex size representation

Being an exploratory study, data was collected from one participant. 40 trials were presented of squares flickering with a frequency of 5Hz and a duty cycle of .5 (Herrmann, 2001; Cecotti, 2010). Stimuli appeared at 150cm distance from the observer, moved away to a distance of 350cm, then moved back to the starting position. This continuous oscillation lasted for 10 cycles, or 40 seconds before the trial ended. Trials were randomly presented, half maintained their physical size, half maintained their retinal size. Stimuli across all trials, at the near distance, subtended 23.8° visual angle per side. Physically constant stimuli subtended 10° visual angle per side at the far distance. Retinally constant stimuli, of course, subtended 23.8° visual angle at the far distance, having physically grown.

EEG data was collected from 13 electrodes (Fz, F3,T7, Pz,P3, P7, O1, Oz, O2, P4, P8, T8, F4) with a sampling rate of 500hz using a Brainvision LiveAmp wireless amplifier (Brain Products GmbH, Germany) and using Brainvision Recorder recording software (Brain Products GmbH, 2024. Electrodes were laid out according to the international 10-20 system. An active reference was used at the FCz location and the ground electrode was placed at AFz.

EEG preprocessing was done with a combination of EEGLAB (Delorme & Makeig, 2004) and custom MATLAB scripts. EEG data was filtered with a high pass filter at .5Hz (full width at half maximum of 1Hz) and a low pass filter at 45Hz (FWHM of 11,25Hz). To extract the SSVEP magnitude and plot it over time, we used 400ms window widths and a 2ms step size. SSVEPs plotted in figure 3 are the summation of the stimulus frequency and its first two harmonics. All SSVEP values were normalized before averaging, causing them to be expressed in z-values.

Contrary to our hypotheses, the data did not conform to either predicted outcome. Although SSVEPs corresponding to the physically constant were lower when furthest away from the observer than compared to surrounding time bins, indicating a more retinally driven coding, the same occurred for retinally constant stimuli. Further, the SSVEP elicited by the physically constant stimulus being relatively low when nearest the observer was not predicted by either hypothesis, since this is when it projects the largest retinal image. The overall patterns of this data indicated that the SSVEP reflected more than only size coding, to the extent that the size information present in the SSVEP could not be reliably extracted when using the present paradigm. Continuous motion appeared to introduce additional factors into the SSVEP, which reduced its usefulness as a marker for size perception.

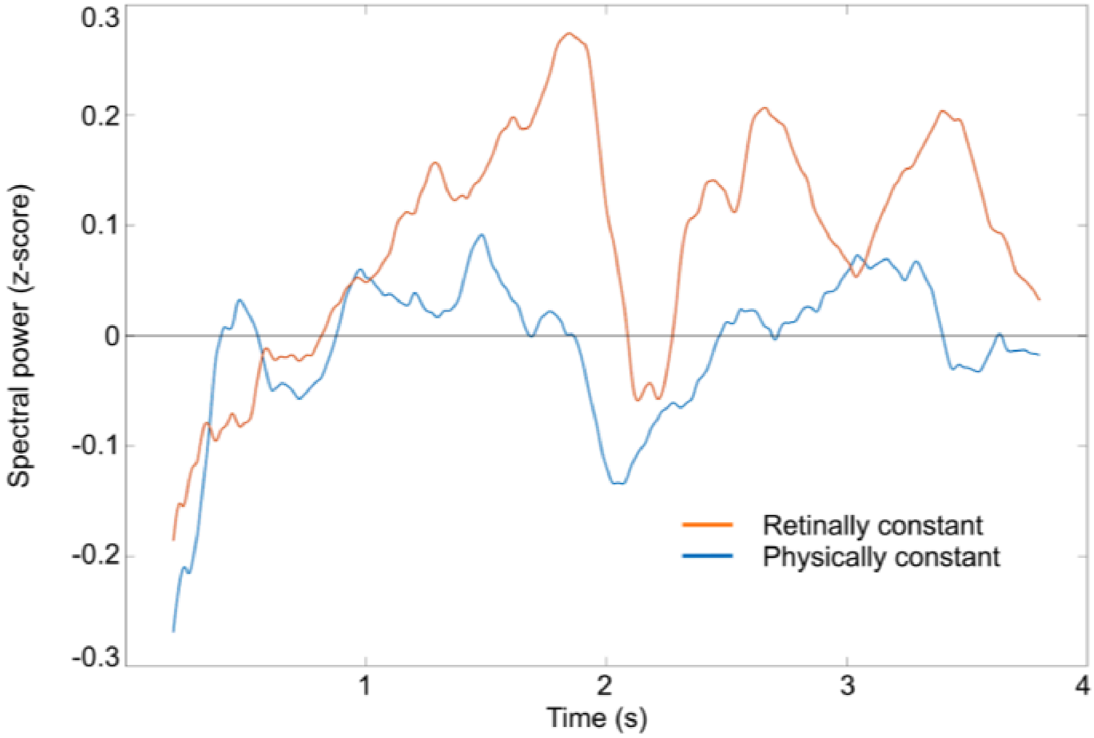


Figure 3: SSVEP magnitude over the course of one stimulus cycle. Stimuli at 0s are closest to the observer, furthest away at 2s and closest again at 4s. The stimulus represented by the blue line is physically constant, orange represents the retinally constant stimulus.

As a next step, we decided to remove continuous motion in our experiment, and instead bookended motion with periods where the stimulus would remain stationary. This was meant to simplify the experiment, and allow for a more straightforward analysis of the effects of movement on SSVEP strength. Further, trials would only contain a single period of motion, as opposed to multiple cycles of back and forth movement like in the previous experiment. A trial would now consist of a stimulus appearing, either at the far or near location, remaining stationary for one second, moving to the opposite location and remaining stationary there for another second before disappearing. There was no distinction between retinally or physically constant stimuli, instead a single stimulus subtending 23° visual angle when close, and 10° visual angle when far was used. Data was collected using the same apparatus and EEG methodology as in the previous experiment. 400 trials were presented in each condition (movement towards observer and movement away from observer), resulting in roughly 53 minutes of usable EEG data from a single observer.

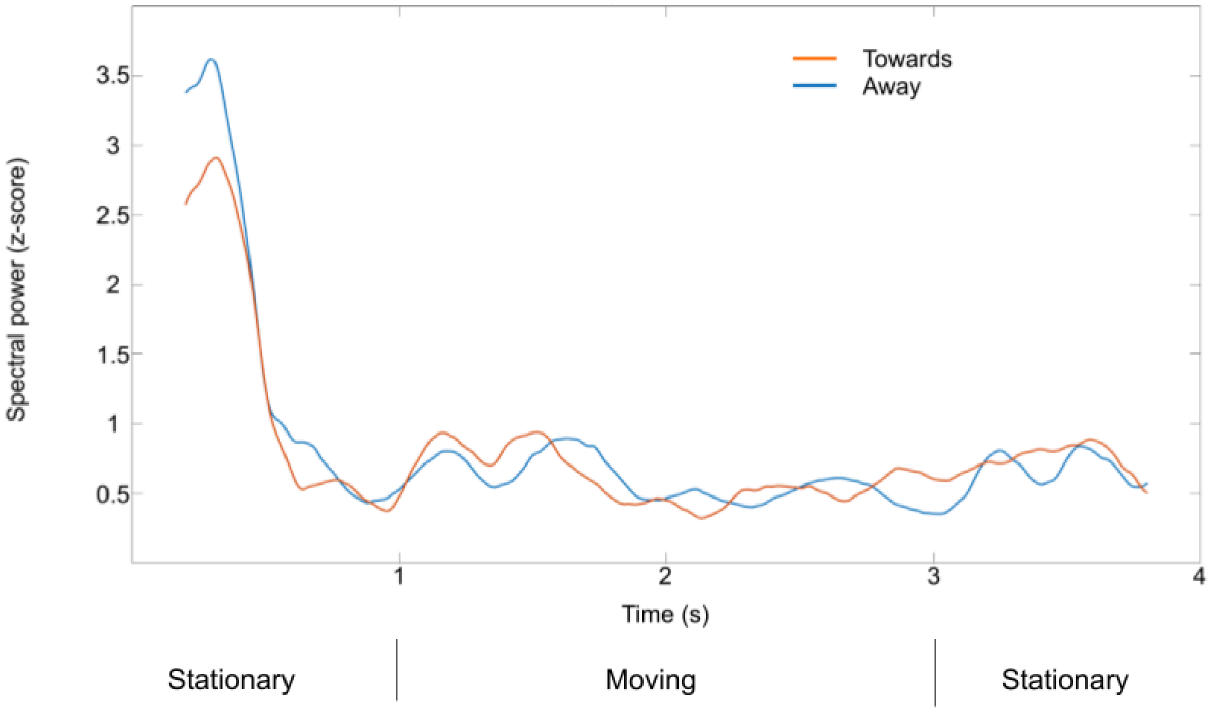


Figure 4: SSVEP magnitude over the course of one stimulus cycle. Stimuli begin moving at 1s, and stop moving at 3s. Orange trace represents stimuli that moved towards the observer, blue trace represents stimuli that moved away from the observer.

Results showed a strong initial increase in SSVEP amplitude, peaking at roughly 300ms after stimulus onset, which normalized around 600ms after stimulus onset. A further set of peaks of lesser magnitude was observed immediately after movement onset (figure 4). In analysing the data, a window width of 400ms and a step size of 2ms was used. Due to this set of parameters, the first data point visible in figure 4 contains the extracted SSVEP from EEG data between 0 and 400ms after stimulus onset. Due to the temporal smearing effect caused by wider window widths described earlier, the actual peak SSVEP may occur earlier than depicted here. Regardless, the nature of this peak was valuable information, as SSVEPs in their initial stages are often described as increasing at first before reaching a steady state (Regan, 1966). Wacker et al. (2011) describe a rapid increase over the first 400ms, and a plateau being reached at around 700ms post stimulus onset, Radtke, Martens & Gruber (2021), describe SSVEPs as increasing over the first 500ms, and reaching a plateau thereafter. Our findings do not match up with this characterisation per se, though they certainly support the same conclusion that it is best to use data for SSVEP analyses starting at 500ms post stimulus onset.

The increased activation observed at around 1000ms could be indicative of an attentional effect elicited by the onset of stimulus movement, or it could be indicative of sensory adaptation to the movement of the stimulus. A similar peak, though to a lesser extent, may be observed after the motion ceases at 3000ms. Attending stimuli, either covertly or overtly, is reported to enhance resultant SSVEPs (Ding, Sperling, & Srinivasan, 2006), and sudden motion onset can capture attention in a bottom-up way (Smith & Abrams, 2018). Even though the stimulus is already attended before motion onset occurs, a transient boost in the SSVEP due to a short term boost in attention is not implausible. Alternatively, the boost in SSVEP magnitude could be more bottom-up and sensory in nature. Activity in neurons of the magnocellular pathway is generally speaking more transient, and is more sensitive to motion, and the increase in SSVEP magnitude could be indicative of an increase in M-pathway activity (McKeefry, Russell, Murray, & Kulikowski, 1996).

In order to verify the timing of the above described phenomenon, an additional experiment was conducted with a very similar methodology. The main difference was in the timing of movement onset, where instead of remaining stationary for 1 second, the stimulus would remain stationary for 2 seconds. 3 observers completed the experiment, which consisted of 40 trials. Trials were split by direction of movement, and presented randomly. Analysis parameters were identical to the previous experiment. Results are displayed in figure 5, and mostly support the conclusions from the previous experiment. Again, a large initial peak in the SSVEP can be observed, which normalizes around 500ms, A secondary peak following motion onset, this time at 2000ms may again be observed, though only in one movement condition. A peak at the cessation of movement, too, can again be observed albeit in one condition. The duration of these peaks, too, is comparable with that observed in the previous experiment.

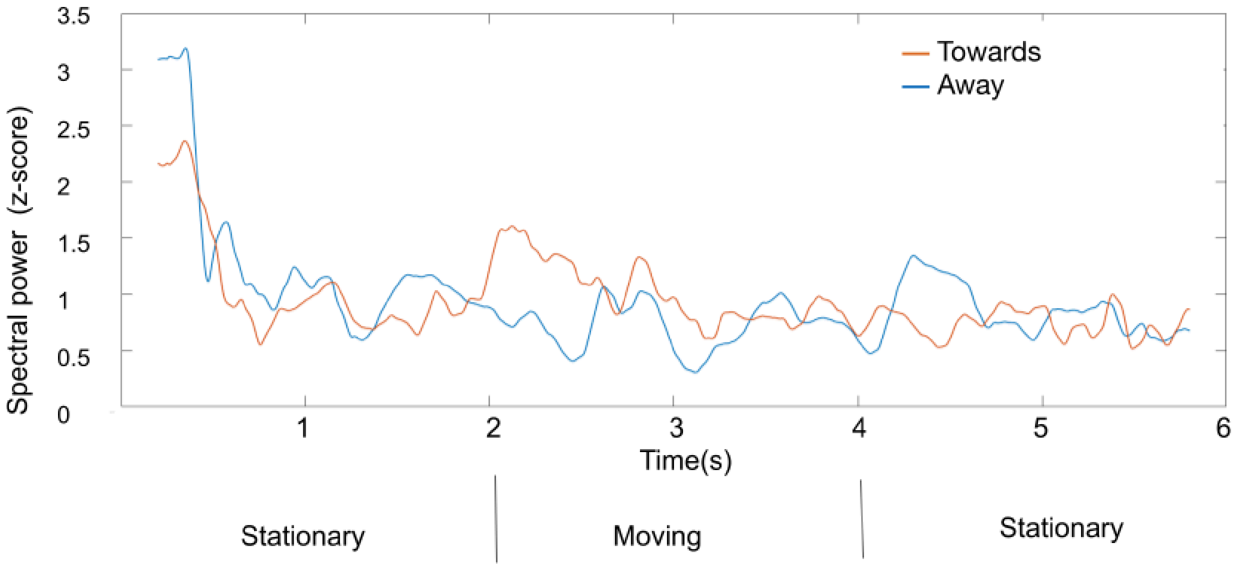


Figure 5: SSVEP magnitude over the course of one stimulus cycle. Stimuli begin moving at 2s, and stop moving at 4s. Orange trace represents stimuli that moved towards the observer, blue trace represents stimuli that moved away from the observer.

These experiments provide valuable information about what parts of the data to select when using the SSVEP as a metric for size perception. The first 500ms after stimulus onset are best avoided due to the large initial spike in SSVEP amplitude. Traces in this initial phase, however, do seem to vary with (retinal) stimulus size. Note that in figures 4 and 5 stimuli moving away from the observer initially appear at a close distance, and subtend a larger visual angle than those moving towards the observer. This larger visual angle is apparent in the larger SSVEP magnitude elicited by this stimulus. This difference between close (and retinally large) and far (and retinally small) stimuli, if anything, becomes less apparent as the SSVEP normalizes.

The question of why retinal size can be distinguished in the SSVEP at all this early on is worth addressing. Given our window width of 400ms, the first data point visible in figures 4 and 5 contains spectral values extracted from a fourier transformation of the first 400ms of EEG data. That size information is present in EEG data relatively early on has been demonstrated previously (Busch, Debener, Kranczioch, Engel & Herrmann, 2004). The P1/N1 complex, occurring before 200ms after stimulus onset is known to vary in amplitude with (retinal) size of stimuli (Pfabigan, Sailer & Lamm, 2015). If size is a task-relevant stimulus property, this variation is even more pronounced. The P1 and N1 are further associated with the alpha and theta bands, with Klimesh, Schack, Schabus, Doppermayr, Gruber & Sauseng (2004) demonstrating the relation between alpha and theta band phase locking and the P1/N1 components. Increased amplitude of waves in the theta band (4-7Hz) as part of the N1/P1 complex, then, are likely to be observed when extracting the 5Hz frequency from the first 400ms after stimulus onset, explaining the larger trace for the retinally larger stimulus. It is important to remember that the SSVEP is likely still gaining in strength in this period, and that even though effects observed in the data might relate to stimulus size, they might not do so via the same mechanisms as the SSVEP proper.

Therefore, the cautious approach is to leave the data from the first 500ms after stimulus presentation out of further analysis that is meant to study size perception. In order to still have sufficient data to extract SSVEPs with a high degree of frequency resolution, we can simply extend the period during which the stimulus is stationary.

An additional experiment was performed to investigate the nature of the transient, roughly 500ms lasting peaks observed in these experiments. Given the sparsity of SSVEP research that analyses the data in a fine temporal aspect, we cannot directly link our observations to any well established phenomenon. Given that these peaks seem to coincide with stimulus appearance and sudden changes in stimulus behavior, it was intuitive to link them to attention. Still, having only observed them in relation to sudden appearance and a change in movement, it could not be excluded that these were phenomena specific to the conditions in which they were observed.

To this end, we designed an experiment in which the SSVEP inducing stimulus changed in colour, while remaining constant in its size and distance. The study of colour perception using SSVEPs has been extensively described in the literature, and this experiment is not meant to contribute to it in any meaningful way. It has been established that colour based attentional effects on SSVEP magnitude are large, even larger than those elicited by luminance, orientation or spatial effects (Adamian & Andersen, 2024). If the short, transient effects observed in earlier data are indeed attentional in nature, colour changes ought to induce them in the same, if not a higher magnitude. Further, specific colours tend to elicit different SSVEP magnitudes. Red and green stimuli tend to induce stronger SSVEPs than blue stimuli. Presumably this is due to the stronger parvocellular activation elicited by L and M cones, which is more sensitive to fast flickering, compared to weaker koniocellular activation by S cones, which is less sensitive to flickering stimuli (Duart, Quiles, Suay, Chio, García &

Morant, 2020). These overall differences should not matter to our analyses, however, since we will be focussing on the transient effects surrounding the change of colour as opposed to overall SSVEP levels during any one colour. We hypothesized that transient increases, much like those observed during motion onset, would occur when stimulus colour suddenly changed.

The experiment was designed in the same virtual hallway as described above. Stimuli would remain stationary at a close distance, 150cm from the observer, subtending 23° visual angle. Stimuli would change colour every 3 seconds, with stimuli starting off green, before turning blue, then red before disappearing after a total presentation time of 9 seconds. EEG collection and preprocessing methods were identical to those described above. Data was collected from one participant. 75 trials were performed.

Results confirmed our hypothesis, with the extracted SSVEP showing marked, transient increases upon stimulus colour change (figure 6). These peaks were roughly similar in duration, lasting around 400 to 500ms. Additionally, the initial peak upon stimulus appearance was similar to that observed in previous experiments. These findings support the notion that attentional and/or sensory effects elicited by stimulus appearance, motion onset / motion change and colour change appear in the SSVEP as quick, transient peaks. Attempting to use SSVEPs as a measure for size perception, then, requires stationary stimuli. This is not to say that size perception in dynamic circumstances cannot be reliably studied using SSVEPs, rather it requires methodology that includes periods of non-movement. These periods would preferably be several seconds long, since the time-frequency trade off dictates that a lengthier signal provides more precise data concerning the frequency spectrum of the collected EEG data.

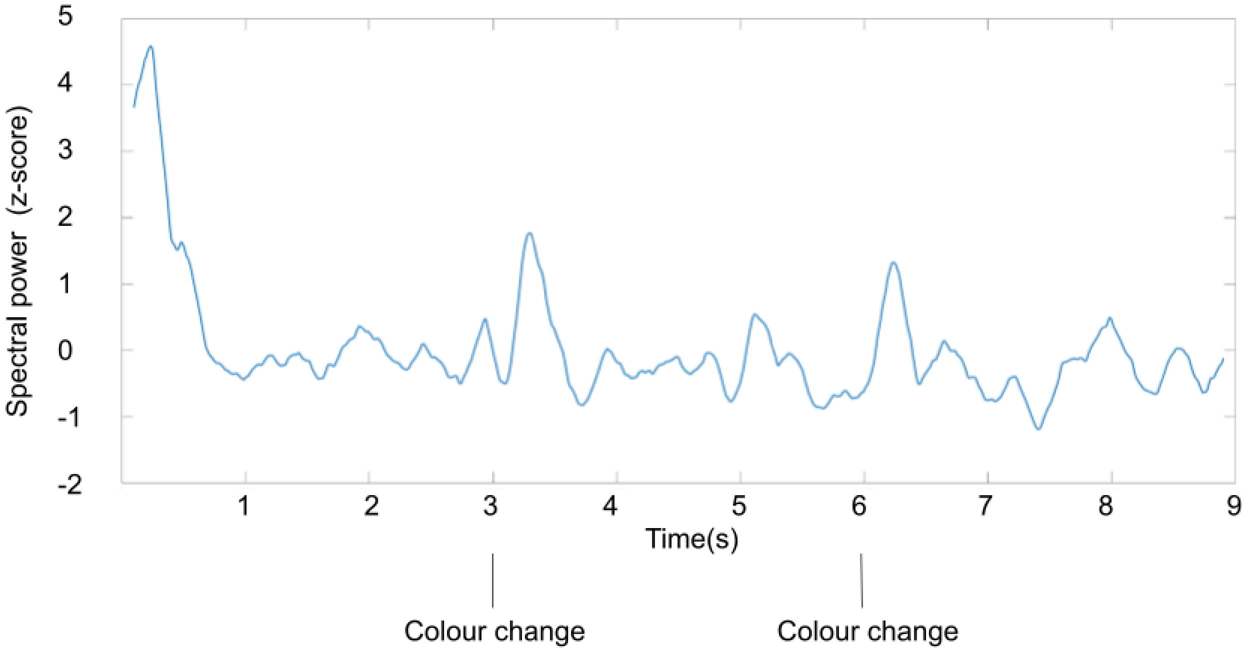


Figure 6: SSVEP magnitude over the course of one stimulus cycle. Stimuli change colour every 3 seconds before disappearing. Colour order was green, blue, red.

Despite not being the object of investigation, the transient peaks in the SSVEP data are interesting enough to warrant some contemplation. SSVEPs are traditionally studied over a longer period of time, and short term fluctuations are not often considered. Recently there appears to have been an increased interest in the time course of SSVEPs, with researchers mapping out SSVEP amplitude development over the course of periods of time roughly 1 minute in length (Prado-Gutiérrez et al., 2019; Radtke et al. 2020; Labecki, Nowicka & Suffczynski, 2019; Labecki, Nowicka, Wróbel & Suffczynski, 2024; Zhang, Valsecchi, Gegenfurtner & Chen, 2023). These experiments seem to study the SSVEP for its own sake, rather than using it as a metric for other phenomena. Even though there is some degree of discrepancy between the results found by the various authors, the common conclusion is that SSVEPs are not, as the name would imply, in a continuous steady state, and that the fluctuations of these SSVEPs are worth studying.

The transient effects we found in the above described experiments are shorter in duration than the effect described in this recent wave of research. Our findings do not allow us to characterize these effects as exclusively bottom-up / perceptual or top-down / attentional. An earlier suggestion that the peaks are due to transient M-pathway activation seems less likely given that they can also be observed in response to changes in colour, since the M-pathway is insensitive to colour (Gegenfurtner, 2003). Of course, this assumes that the transient peaks observed in relation to movement onset are the same in nature as those observed in relation to colour change. This might not be true at all, and these two effects may have different neural underpinnings altogether. A more focussed investigation into the nature of these transient peaks might be fruitful in learning more about the neural mechanisms that drive them. For our purposes, it suffices that stimulus appearance, movement onset and colour change induce transient increases in SSVEP amplitude that do not correlate to object size, and in using SSVEPs as a marker for perceived size, these phenomena ought to be avoided.

Experiment 1: Motion, interrupted

Based on the findings of the above described experiments, we refined our methodology to better isolate the contribution of size to the overall SSVEP (figure 7). The revised experimental design involved a structured sequence of visual stimulus presentations. The structure of a single trial would consist of the following components:

- Initial Appearance: Each stimulus first appeared and remained stationary for a period of 4 seconds. This phase allowed us to capture the SSVEP response to the stimulus at its starting position.
- Motion Phase: Following the stationary period, the stimulus then moved for 3 seconds. During this phase, the stimulus could either move away from or towards the observer. The stimuli could also shrink or grow, depending on whether its size was physically or retinally constant.
- Final Stationary Phase: After the motion, the stimulus returned to a stationary state for another 4 seconds. This allowed us to observe the SSVEP response post-movement.

The stimulus, similarly to previous experiments, would be a flickering white square presented on a black background in virtual reality.

For the benefit of data quality, and to increase the likelihood that the SSVEP data being studied was reflective of stimulus size instead of any attentional or otherwise transient effect, we implemented a crucial analytical step. We used the last 3.5 seconds of each 4-second stationary segment (both first and second) for our analysis, deliberately excluding the initial 0.5 seconds. This approach was taken to remove any artefacts that might be caused by the sudden appearance of the stimulus or the effect of movement, which are described in the previous chapter.

This methodology became the standard for our subsequent experiments. When interpreting the following results, especially the graphical representations, it’s important to note that any graph showing two data points along the x-axis corresponds to these specific time segments: the first and last 3.5 seconds of the stationary phases. This approach allowed us to have a cleaner and more focused analysis, concentrating on the SSVEP responses that were less likely to be influenced by immediate stimulus appearance or movement initiation.

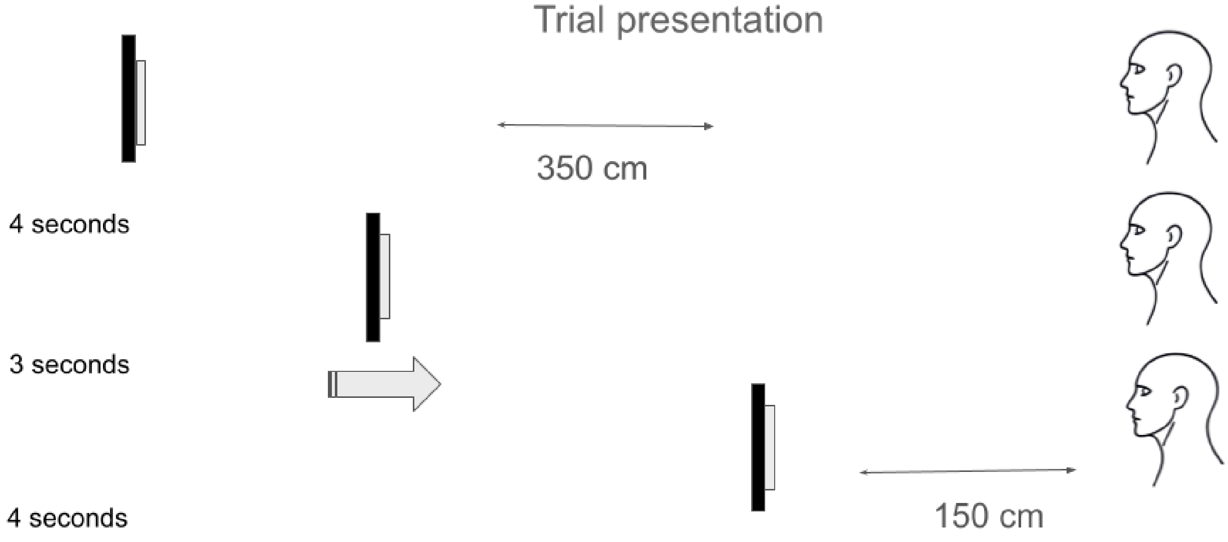


Figure 7: Top to bottom: progression of a single trial. Stimuli would appear, remain stationary for 4 seconds, move over the course of 3 seconds and remain stationary for another 4 seconds before disappearing. Although this figure depicts the screen moving towards the observer, the screen could both move towards or away from observers.

The first experiment we conducted using this new methodology had the same basic research question as the first experiment described in the previous question. Namely, how is the size of moving stimuli represented in early visual cortex. We again focussed on movement in depth, with stimuli moving either away from or towards the observer. Stimuli could again behave in one of two ways; they could either maintain a constant physical size throughout the trial, or they could shrink or grow in such a way as to retain a constant retinal size throughout the trial. Stimuli in this retinally constant condition would be assigned a retinal size, and they would maintain this retinal size relative to the observer, which in virtual reality meant the location of the camera in the virtual environment. Similarly to the previous experiment, two competing hypotheses were put forward concerning the outcomes. If early visual areas code stimulus size in a (partially) size constant manner, SSVEPs are expected to be the same before and after movement in the physically constant condition. Conversely, SSVEPs are expected to vary with the physical stimulus size in the retinally constant condition. If early visual areas instead encode objects in a purely retinotopic manner, SSVEPs are expected to display the inverse effect.

6 participants took part in the experiment. All participants had normal or corrected-to-normal vision. The experiment consisted of 80 trials, split by stimulus size, movement direction and type of constancy, resulting in 10 trials per unique combination of these factors. Trials were split into blocks of 10, and trial order was fully randomized. Stimuli could appear in one of two locations, either at 150cm or at 350cm in front of the observer. The experiment took place in virtual reality, in the same hallway environment as described earlier. Stimuli would remain stationary for 4 seconds before moving to the opposite location; stimuli that appeared at 150cm would move away from the observer to a distance of 350cm, and vice versa.

Stimuli were squares flickering at 15 Hz according to a square wave pattern with a duty cycle of 0.5. Stimuli could have one of three sizes; small (26.25cm per side), medium (61.26cm per side) or large (142.9cm per side). Small stimuli would appear at the close distance, where they would subtend 10° visual angle, the same angular size as a medium stimulus at the far distance. Large stimuli would appear at the far distance, where they would subtend 23.08° visual angle, the same as medium stimuli at the close distance. With medium stimuli appearing at both distances before moving, this gives us the necessary data to measure the size constancy present in the SSVEP data before any movement occurs. Stimuli sizes, distances and angular sizes are displayed in table 1.

	Large (142.9 cm)	Medium (61.26 cm)	Small (26.25 cm)
Far (350 cm)	23.08°	10°	4.30°
Near (150 cm)	50.94°	23.08°	10°

Table 1: Sizes, distances and angular sizes of stimuli used. Stimuli were squares with the physical and retinal sizes listed above referring to the length of a side of this square.

After the disappearance of the stimulus, a reference square would appear at an intermediate distance of 250 cm in front of the participant. Participants were tasked with adjusting the size of this square to match that of the stimulus. In case of retinally constant stimuli, participants were asked to replicate the physical size of the stimuli as they most recently saw it.

EEG data was collected from 13 electrodes (Fz, F3,T7, Pz,P3, P7, O1, Oz, O2, P4, P8, T8, F4) with a sampling rate of 500hz using a Brainvision LiveAmp wireless amplifier (Brain

Products GmbH, Germany) and using Brainvision Recorder recording software (Brain Products GmbH, 2024. Electrodes were laid out according to the international 10-20 system. An active reference was used at the FCz location and the ground electrode was placed at AFz. Impedances were kept below 5kΩ.

EEG preprocessing was done with a combination of EEGLAB (Delorme & Makeig, 2004) and custom MATLAB scripts. EEG data was filtered with a high pass filter at .5Hz (full width at half maximum of 1Hz) and a low pass filter at 45Hz (FWHM of 11,25Hz). Since using a Laplacian reference is shown to be optimal for analysing SSVEP data (Zhang, Valsecchi, Gegenfurtner & Chen, 2023), the data was re-referenced to a Laplacian reference using the CSD toolbox (Kayser & Tenke, 2006). Data was detrended, and a Tukey window with an alpha of 0.25 was applied to it. Data was then Fourrier transformed. Per trial, frequency spectra from three occipital electrodes; O1, Oz and O2 were averaged into one spectrum. From these averaged spectra, the stimulation frequency and the first two harmonics were extracted. These peak values were corrected for the spectral power of their surrounding frequencies, subtracting the average of the nearest eight frequencies from them. SSVEP peak values were then summed to form a single SSVEP value, and standardised to z-scores for each participant.

Figure 8 shows resultant SSVEPs from physically constant stimuli. Despite the sample size not being large enough to result in significant differences in SSVEP strength of the stimuli during the first segment of the data, on an ordinal scale the results imply partial size constancy in the signal. The large stimulus which first appears far from the observer (purple in fig 8) elicits a larger SSVEP than the retinally equally sized medium stimulus that is presented close to the observer (orange in fig 8). Similarly, the medium sized stimulus at the far distance (yellow in figure 8) elicits a larger SSVEP than a small stimulus which is close (blue in figure 8). Comparing this to SSVEPs in the second segment, after the movement has occurred, further shows that size constancy is not perfect. Rather, stimuli that move towards the observer elicit greater SSVEPs than before to go along with their increased retinal size, and the opposite occurs for stimuli moving away from the observer. For a summary of the retinal and physical sizes of stimuli depicted in figure 8, see table 2.

	Small, moving away	Medium, moving away	Medium, moving towards	Large, moving towards
First segment				
Retinal size	10°	23.08°	10°	23.08°
Physical size	26.25 cm	61.26 cm	61.26 cm	142.9 cm
Distance	150 cm	150 cm	350 cm	350 cm
Second segment				
Retinal size	4.30°	10°	23.08°	50.94°
Physical size	26.25 cm	61.26 cm	61.26 cm	142.9 cm
Distance	350 cm	350 cm	150 cm	150 cm

Table 2: Sizes, distances and angular sizes of stimuli during the first and second segments in the physically constant condition. Stimuli were squares with the physical and retinal sizes listed above referring to the length of a side of this square.

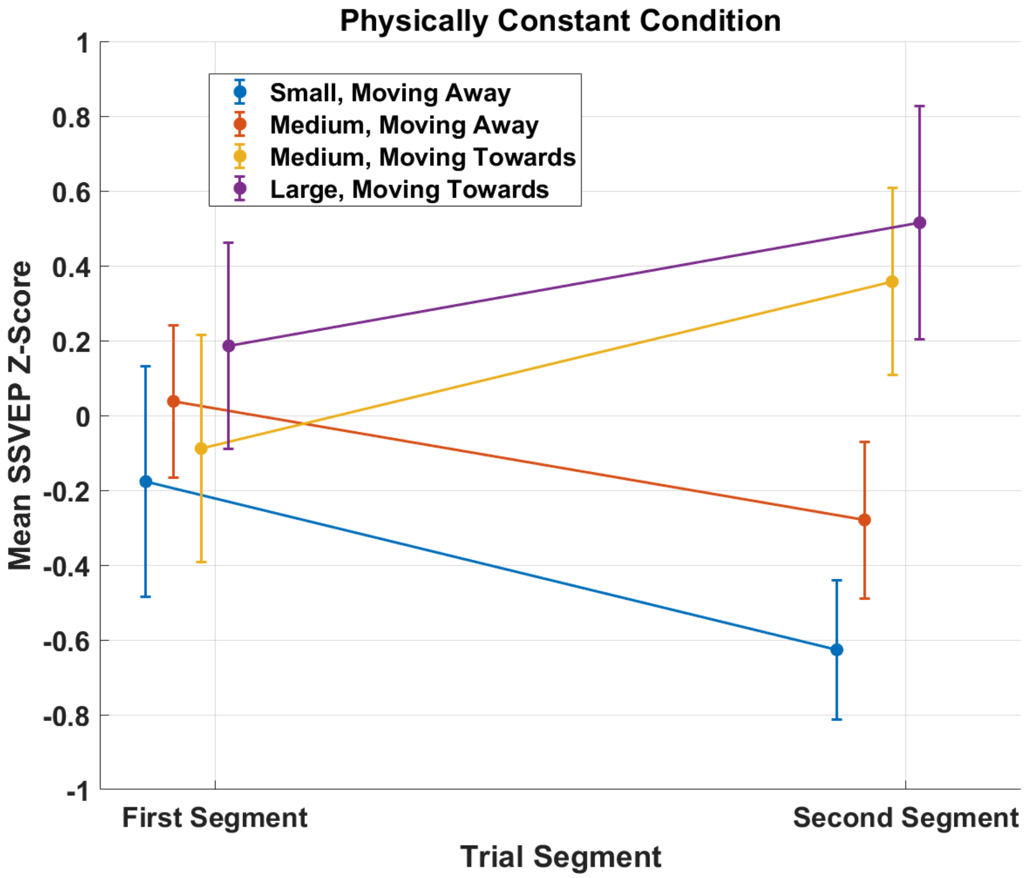


Figure 8: SSVEP strength for stimuli in the physically constant condition. Error bars represent +/- 2 standard errors.

Any degree of variation of SSVEP strength in this condition is indicative of the effect of retinal size on the SSVEP, since the physical size of stimuli did not vary. The proportionality of retinal size influence versus physical size influence is hard to deduce, because of how SSVEPs scale with stimulus size. At smaller sizes, SSVEPs may scale linearly with stimulus size (Duszyk et al., 2014), though this linearity disappears at larger (>5°) sizes (Ng & Kian, 2012; Chen, McManus, Valsecchi, & Gegenfurtner, 2019). Even given the assumption that SSVEPs are purely reflective of the amount of retinotopic tissue stimulated by an object, an increase in retinal size will be proportional to the retinotopic organisation of these visual areas, which largely overrepresent the foveal part of the visual field, and which quickly falls off toward the periphery (Schira, Tyler, Breakspear & Spehar, 2009). Since V1 anatomy also differs on an individual basis, which in turn affects SSVEP responses, more information about the neuroanatomy of each participant would have to be known to conclusively link SSVEP strength to a corresponding size representation in V1 in (Vanegas, Blangero & Kelly, 2013).

On the other hand, in the retinally constant condition results are less intuitive. Surprisingly, retinally constant stimuli that move away from the observer elicit weaker SSVEPs after movement, despite the fact that their retinal size remained constant and their physical size increased. Similarly unexpected is that stimuli moving towards the observer show increasingly strong SSVEPs despite retaining their retinal size, and shrinking physically. Neither of the above mentioned hypotheses predicted this outcome, nor is it easily explained. For a summary of the retinal and physical sizes of stimuli depicted in figure 9, see table 3.

	Small, moving away	Medium, moving away	Medium, moving towards	Large, moving towards
First segment				
Retinal size	10°	23.08°	10°	23.08°
Physical size	26.25 cm	61.26 cm	61.26 cm	142.9 cm
Distance	150 cm	150 cm	350 cm	350 cm
Second segment				
Retinal size	10°	23.08°	10°	23.08°
Physical size	61.26 cm	142.9 cm	26.25 cm	61.26 cm
Distance	350 cm	350 cm	150 cm	150 cm

Table 3: Sizes, distances and angular sizes of stimuli during the first and second segments in the retinally constant condition. Stimuli were squares with the physical and retinal sizes listed above referring to the length of a side of this square.

When we compare SSVEPs from both conditions, the counterintuitive findings from the retinally constant condition are highlighted (figure 10). Stimuli in the first segment appear to elicit SSVEPs in line with their retinal size, and although there is no strong support for partial size constancy here, these results are still in line with what can reasonably be expected. In the second segment, however, the SSVEPs make less sense. The physically larger stimulus at a further distance elicits a smaller SSVEP compared with two medium sized stimuli at a closer distance. This not only indicates a lack of size constancy, but if the difference were significant, a kind of opposite size constancy where physically larger stimuli elicit a smaller neural response in early visual areas. Clearly, the combination of movement and change of size elicits neural activity that is not readily explained by size constant theories of early visual processing.

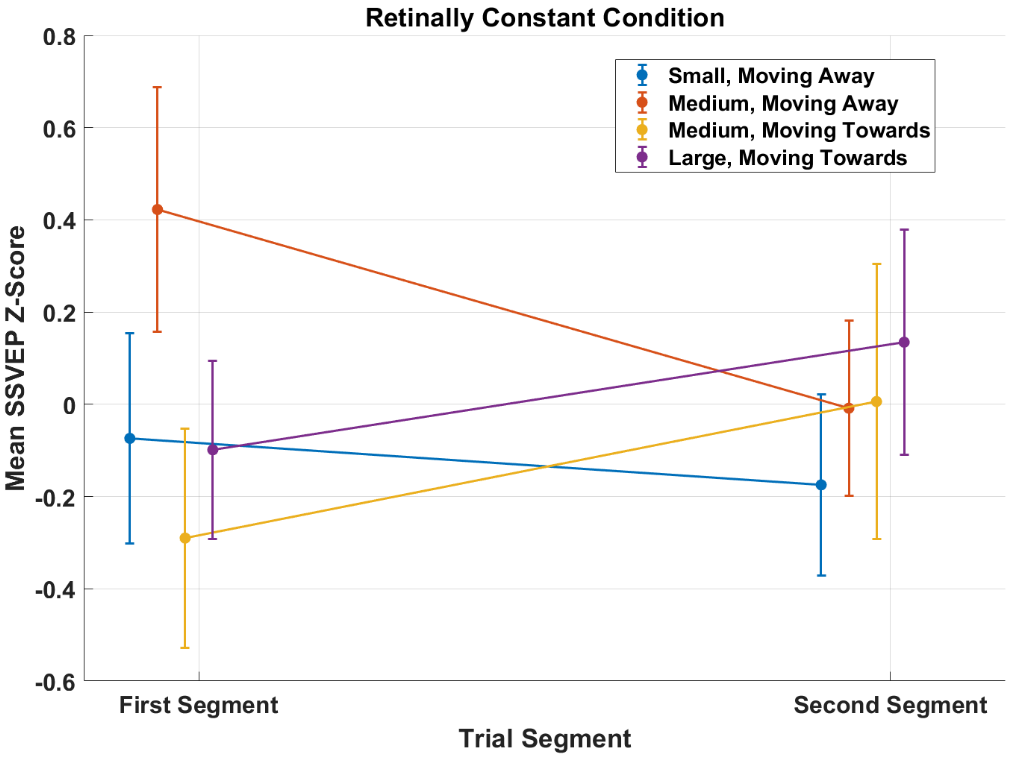


Figure 9: SSVEP strength for stimuli in the retinally constant condition. Error bars represent +/- 2 standard errors.

There are two possible avenues of thought along which these findings can be explained. One assumes that these observed SSVEPs accurately depict size perception in early visual cortex, and another that assumes that they are not. The latter is instantly more likely due to the distribution of data points in the first segment, preceding any movement. Whereas in the physically constant condition the SSVEPs in the first data segment reflect partial size constancy in accordance with previous research, stimuli in the retinally constant condition do not. Notably the SSVEP induced by the large stimulus, first appearing at the far distance, appears particularly out of order. This does little to assuage concerns about the counterintuitive change in SSVEP that occurs after stimulus movement, rather it might indicate that these results are driven by something other than neurological size coding.

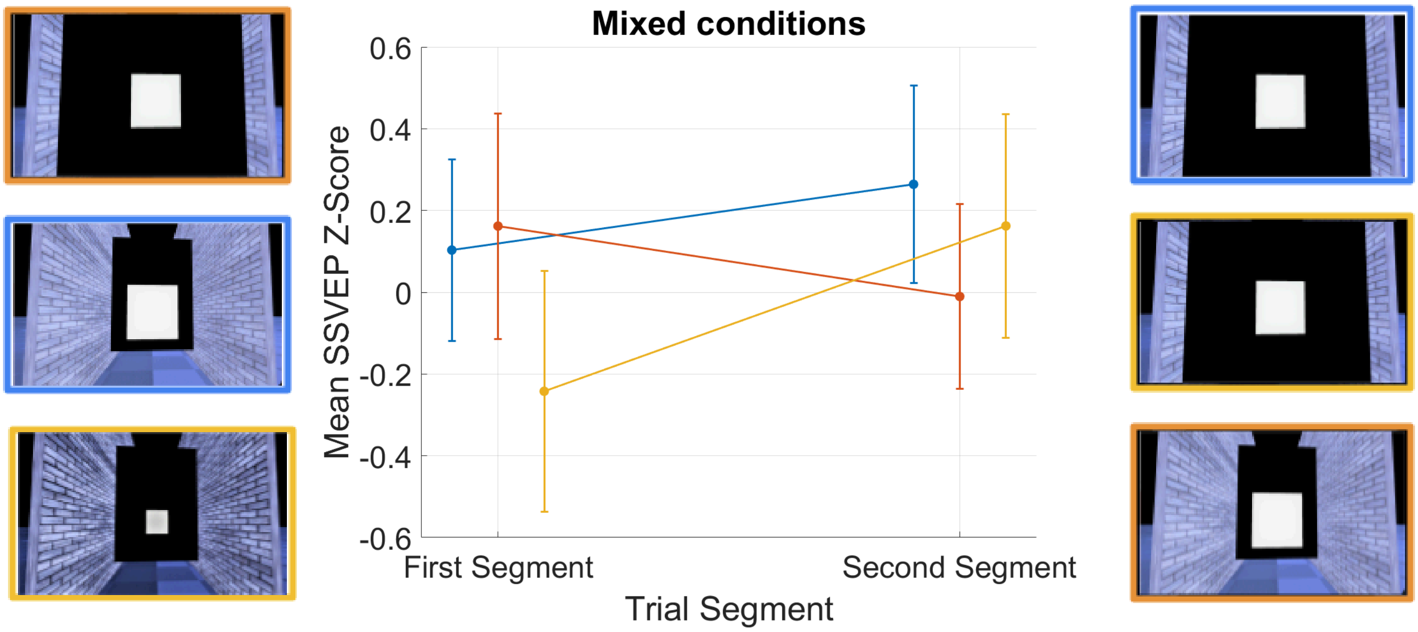


Figure 10: SSVEP strength for stimuli from various conditions and images of the stimuli that elicited these SSVEPS. Error bars represent +/- 2 standard errors.

This is not to say that a neuroscientific explanation ought to be ruled out. Stimuli like those in the retinally constant condition behave in a very counterintuitive and unnatural way. Theories surrounding predictive coding posit that top down signals are generated in line with expectations of the outside world, which are sent to lower level visual processing areas (Rao & Ballard, 1999). Particularly strong expectations of retinal size growth or shrinkage, such as may be associated with stimuli moving straight towards or away from the observer may lead to top down signals modulating lower level visual areas in ways that are congruent with these expectations. It should be noted,however, that SSVEP signals do not seem to vary with perceptual certainty when tested in paradigms that specifically address predictive coding (Gordon, Koening-Robert, Tsuchiya, van Boxtel & Hohwy, 2017). This is not to say that predictive coding related processes play no role in the observed effects of the retinally constant condition, but given the information available no firm conclusions can be drawn.

To confirm the findings of the earlier series of experiments, we also performed a finer grained analysis of the time course of the SSVEP during this experiment. Using a window width of 266ms (4 stimulation frequency cycles) and a step size of 10 data points (or .02 seconds) we obtained the traces shown in figure 11. The facilitation phase where the SSVEP signal increases after stimulus onset again takes a little under 500ms, enforcing the rationale behind excluding this section of the data from analysis. Surprisingly, large peaks after movement onset at 4 seconds and movement offset at 7 seconds remain absent in this data. The most likely reason for this difference is that the stimulation frequency used here (15Hz) was different to that used in the preliminary series of experiments (5Hz). Recent work has revealed the difference in temporal evolution of SSVEPs when different stimulation frequencies are used (Labecki, Nowicka, Wrobel & Suffczynski, 2024). The finding that the transient peaks described earlier seem to be dependent on the stimulation frequency lends credence to the notion that they are attentional in nature, and are closely linked to frequency bands involved with attentional effects in EEG.

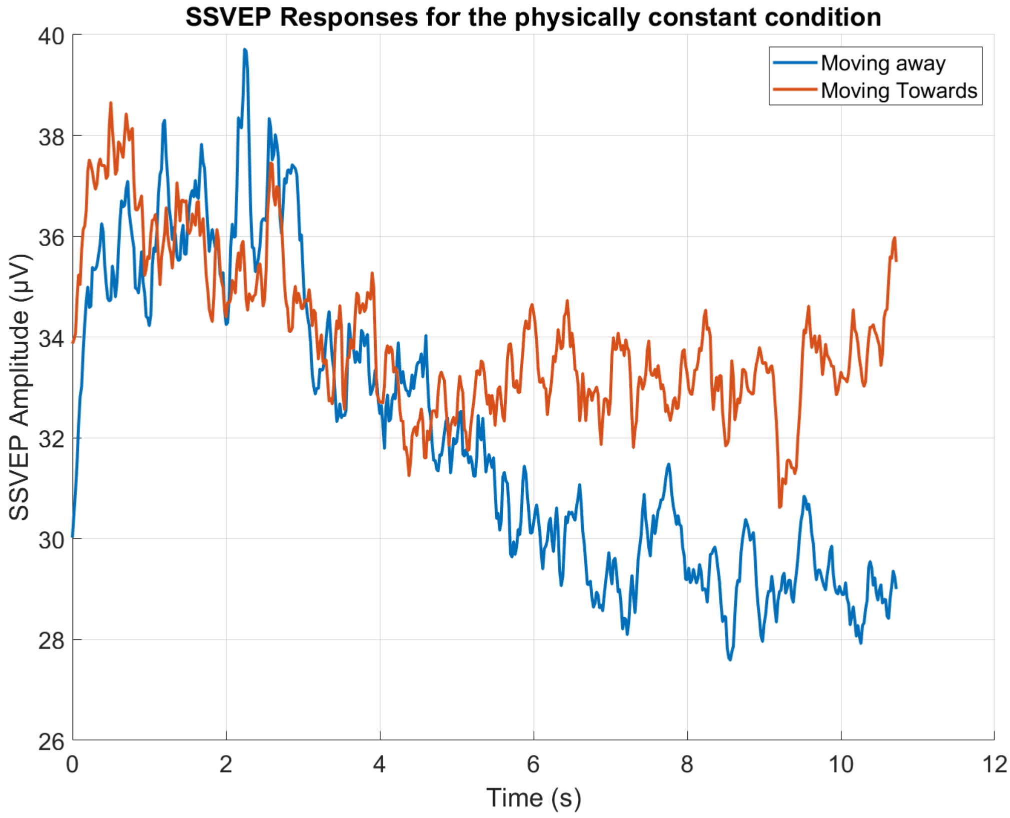


Figure 11: SSVEP strength over time for the physically constant condition. Traces represent all trials split by movement direction, averaged over stimulus size.

Behavioural data were collected after each trial, where subjects attempted to recreate the physical size of the stimulus at an intermediate distance. Error scores, where positive error scores indicate overestimation of the object are depicted in figure 12. In the physically constant condition, error scores were not significantly different from 0, indicating accurate efforts to replicate the size of the stimulus. In the retinally constant condition, larger errors were observed. The size of stimuli moving away, which grew as they moved, was overestimated. Conversely, the size of stimuli moving towards the observer, shrinking as they did so were underestimated in one condition.

Subjects showed high accuracy in reproducing stimulus size in the physically constant condition. Due to the high depth cue salience present in the VR environment and the fact that stimuli did not shrink or grow as they moved, this condition mimicked naturalistic viewing conditions moderately closely. It is therefore not surprising that size perception functioned well, like it usually does in everyday life. Like Kilpatrick and Ittelson (1956) wrote about size-distant invariant perception under normal viewing conditions: "*It is, so to speak, what one would expect the organism to do if there were no reason for doing otherwise.*"

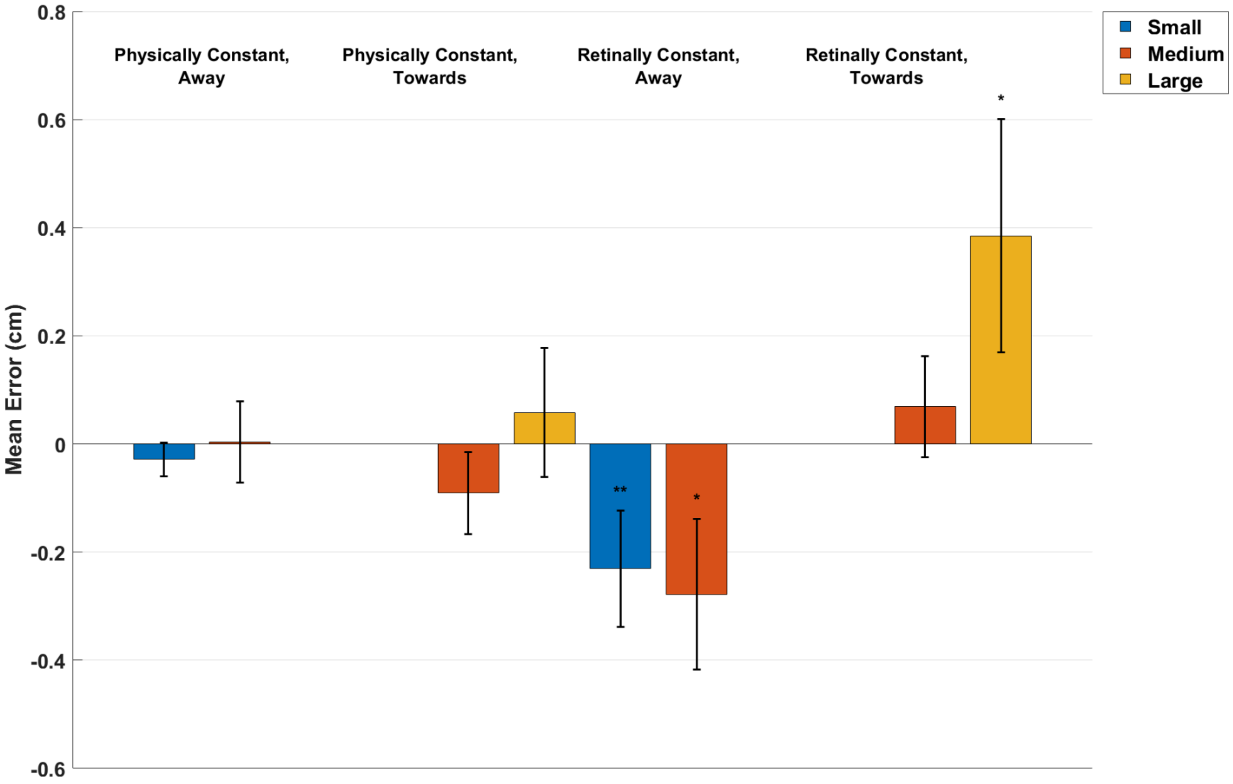


*Figure 12: Error scores in size reproduction task split by stimulus size, movement direction and constancy condition. Positive error scores indicate overestimation. Colour codes for stimulus size. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$*

Still, there is scientific support for the notion that size perception is affected by movement, specifically movement towards the observer, also known as looming motion (Franconeri & Simons, 2003; Rossini, 2014). Brosgole, McNichol, Doyle, & Neylon (1976) found that size constancy deteriorated, and stimuli were perceived as physically larger when they moved towards the observer. Stimuli were also perceived as physically smaller when they moved away. It should be noted that these findings stem from an experiment performed under extremely visually impoverished conditions, where a luminous stimulus is all that is visible. Similarly, Sugarman, McGlinchey & Fortenbaugh (2021) found that stimulus size is overestimated when these are preceded by looming optic flow. This finding is also similar to Hensherson’s (1992) finding, who described how viewing a contracting spiral disk for some time before viewing a stationary stimulus will make that stimulus appear to move in depth towards you.

This movement will then cause the illusory perception that the stimulus is growing. These findings imply that under impoverished or ambiguous viewing conditions, size can be misconstrued due to motion and size constancy may suffer. Our findings show that given sufficient depth information, motion does not disrupt size constancy.

For the retinally constant condition, large errors in size reproduction were observed. Lower precision is to be expected in this condition, because the task is inherently more difficult in this condition. With the physical size changing over the course of the trial, subjects have less time during which they can observe the stimulus at its relevant size, and they are tasked with replicating a dimension that has until recently been in flux. Still, besides more variance, systematic errors can also be observed. Participants tend to misestimate the size of stimuli in the opposite direction of their change, i.e. growing stimuli were reproduced as smaller than they were and shrinking stimuli were reproduced as larger. It appears, therefore, that people bias their responses in the direction of the initial size of the object. This could be due to a strong perceptual prior, where observers tend to assume that a change in retinal size is due to movement rather than due to size, and thus underestimate the degree of change. Alternatively, participants could be replicating the stimulus from memory at the time that it passed the intermediate distance. After all, the stimulus will pass the point 250cm from the observer during the trial, which is where the replication object later appears. In replicating the object to match the impression they had at that time, they will necessarily underestimate the true physical size of retinally constant objects moving away, and overestimate those moving towards them. Either way, literature on size estimation during movement and simultaneous size changing is still limited, however, and more experiments on the combination of movement and size change ought to be done to draw any more definitive conclusions about the mechanisms underlying these systematic errors.

Experiment 2: Presence or absence of stimulus background

Following the previous experiment’s results, many open questions remained. In the retinally constant condition, especially, results defied expectations. A new experiment was decided upon in an attempt to better understand these results.

One aspect of the experimental design that might seem unnatural is the large black background upon which the stimulus is presented. The use of this background was copied from previous research, as it had proven to be part of an effective paradigm there. Nevertheless, this background made for an unnatural element in an experiment that was otherwise making an attempt at greater external validity. The background, making contact with the hallway walls on either side, largely obscures the view beyond the stimulus. This in turn reduces the depth cue salience in the immediate vicinity of the stimulus, which could lead to poorer depth perception and by extension poorer size perception. Conversely, the intersection between the background and the walls on either side of the stimulus could form an extremely informative depth cue, as it marks the exact point along the hallway where the stimulus is located. Further, lacking any background and not having any texture or shading, stimulus distance may become more ambiguous, as it may appear to be floating in mid air at an indeterminable distance. In short, how the presence or absence of this background would affect size perception was an empirical question.

A second and no less important consideration was the role of the background on SSVEP amplitude. Using the SSVEP as a metric for early visual cortex based size perception is only valid insofar as other parameters affecting the SSVEP are held constant. The black background provides high contrast both as a background as well as near the edges of the stimulus. These sharp edges, along with total luminance, are major drivers of SSVEP amplitude (van der Tweel & Spekreijse, 1966). In general, higher contrast leads to higher SSVEP amplitude, until a point of saturation at around 50% modulation of depth (Spekreijse, van der Tweel & Zuidema, 1973; Regan, 1973). It should be noted that modulation of depth can linearly increase SSVEPs given a high enough mean contrast, as demonstrated by Bodis-Wollner, Hendley & Kulikowski (1972). This phenomenon is further complicated by the fact that contrast affects SSVEPs differently based on the stimulation frequency, with low (5Hz) stimulation reaching saturation at higher flicker modulation depths than high frequency flickers (30Hz) (Spekreijse, 1966). Our stimulation frequency being 15Hz, and the 30Hz harmonic being included in the analyses, might be sensitive to saturation. While the depth of modulation is kept at 100% similar to the previous experiments, we decreased the overall luminance to prevent oversaturation effects.

Further, the background might affect SSVEPs by a more indirect influence. Lateral, inhibitory connections in early visual cortex are known to affect neuron population responses to luminance. Lateral effects, presumably functions of these very connections, were observed by Zemon & Ratliff (1984) when measuring SSVEPs. These effects were driven directly by other, adjacent flickering stimuli, however. Polat & Norica (1996) describe long range facilitation and inhibition of SSVEPs, where objects several degrees removed from the inducer affected the resultant SSVEP. More recently, Vanegas, Blangero & Kelly (2015) demonstrated the strong suppression on SSVEPs by the contrast of the surround. The surround in this case meant a computer screen subtending 33.67° by 25.70° visual angle. High contrast in the surrounding visual field, then, can have significant impact on the SSVEP elicited by the central stimulus. As of yet there is no comprehensive analysis describing the exact role of contrast, spatial frequency, luminance and eccentricity as variables affecting SSVEP amplitude. Given the SSVEPs differential responses to the above variables in isolation, a hypothetical analysis like these might well suffer from combinatorial explosion. Still, removing the background in this experiment is likely to also affect SSVEP responses via these surround-suppression type effects.
Overall, we might expect a lower average SSVEP amplitude when removing the background, a cost which may be offset by an increased degree of realism in the virtual environment.

Trials in this experiment could vary along 3 dimensions. Movement direction and retinal vs physical constancy were again included as factors, as well as the presence or absence of a background (figure 13). The experiment contained 80 trials in total, which gave each unique condition a total of 10 trials. All stimuli started out with the dimensions of the ‘medium’ object. Dimensions of this stimulus under the various conditions and locations can be found in tables 2 and 3 of the last chapter. The behavioural task was identical to the previous experiment as well; a size replication task at an intermediate distance. The timing of each trial, too, was identical to the previous experiment, with stimuli staying stationary for 4 seconds, moving over the course of 3 seconds and then remaining stationary for another 4. The reference object which participants had to adjust to replicate the size of the stimulus appeared directly after stimulus offset.

Stimuli were presented in the same virtual hallway condition as used in previous experiments, and the environment was presented using a Tobii Pro VR Integration; an HTC Vive with a Tobii eye tracker integrated into it. This HMD has a refresh rate of 90Hz, a field of view of 110 and a resolution of 2160×1200 (1080×1200 per eye).

8 participants took part in the experiment. EEG data was collected from 13 electrodes (Fz, F3,T7, Pz,P3, P7, O1, Oz, O2, P4, P8, T8, F4) with a sampling rate of 500hz using a Brainvision LiveAmp wireless amplifier (Brain Products GmbH, Germany) and using Brainvision Recorder recording software (Brain Products GmbH, 2024. Electrodes were laid out according to the international 10-20 system. An active reference was used at the FCz location and the ground electrode was placed at AFz. Impedances were kept below 5kΩ.

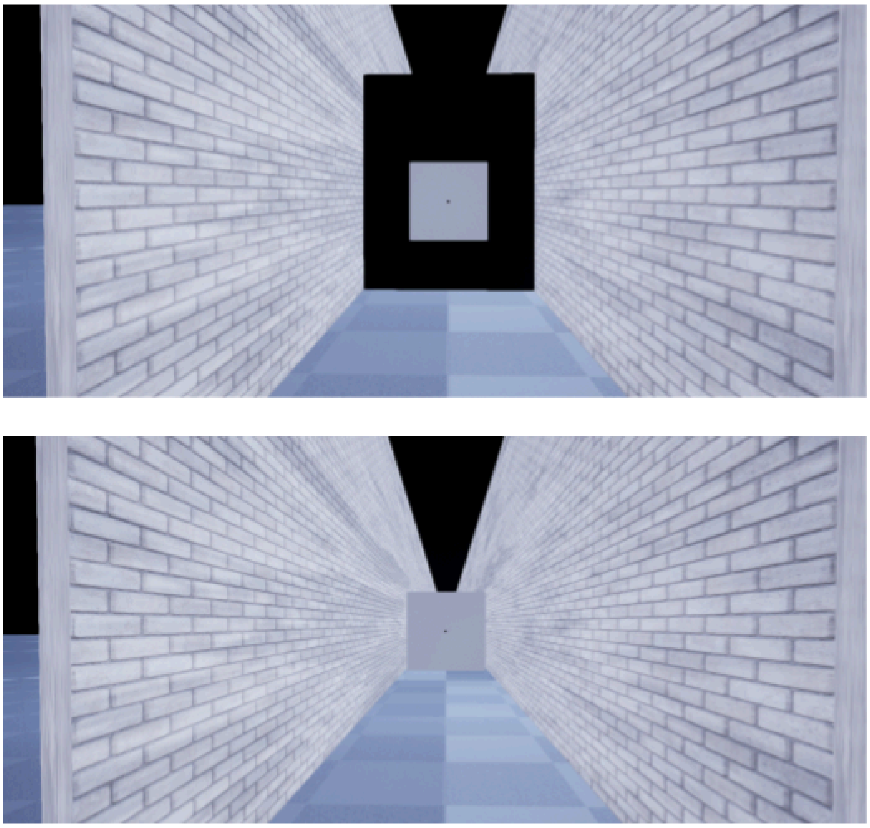


Figure 13: Virtual hallway with stimulus present with (above) and without (below) background

EEG preprocessing was done with a combination of EEGLAB (Delorme & Makeig, 2004) and custom MATLAB scripts. EEG data was filtered with a high pass filter at .5Hz (full width at half maximum of 1Hz) and a low pass filter at 45Hz (FWHM of 11,25Hz). Since using a Laplacian reference is shown to be optimal for analysing SSVEP data (Zhang, Valsecchi, Gegenfurtner & Chen, 2023), the data was re-referenced to a Laplacian reference using the CSD toolbox (Kayser & Tenke, 2006). Data was detrended, and a Tukey window with an alpha of 0.25 was applied to it. Data was then Fourrier transformed. Per trial, frequency spectra from three occipital electrodes; O1, Oz and O2 were averaged into one spectrum. From these averaged spectra, the stimulation frequency and the first harmonic were extracted. These peak values were corrected for the spectral power of their surrounding frequencies, subtracting the average of the nearest eight frequencies from them. SSVEP peak values were then summed to form a single SSVEP value, and standardised to z-scores for each participant.

Figure 14 depicts SSVEPs in the physically constant condition for both movement directions and presence or absence of background. As hypothesized, the SSVEPs are smaller when there is no background present. This trend holds true across movement directions and time points during the trial. More generally, SSVEPs seem to vary with retinal size. Stimuli coming toward the observer, retaining their physical size but increasing their retinal size, induce larger SSVEPs, and vice versa for stimuli moving away from the observer. Similarly to results in the previous experiment, a small effect of movement might be deduced. Stimuli moving to a location elicit a slightly different SSVEP than a stimulus that had started at that same location. E.g. The stimulus having moved towards the observer elicits a larger SSVEP than the stimulus that started close to the observer, despite being identical in size and distance.

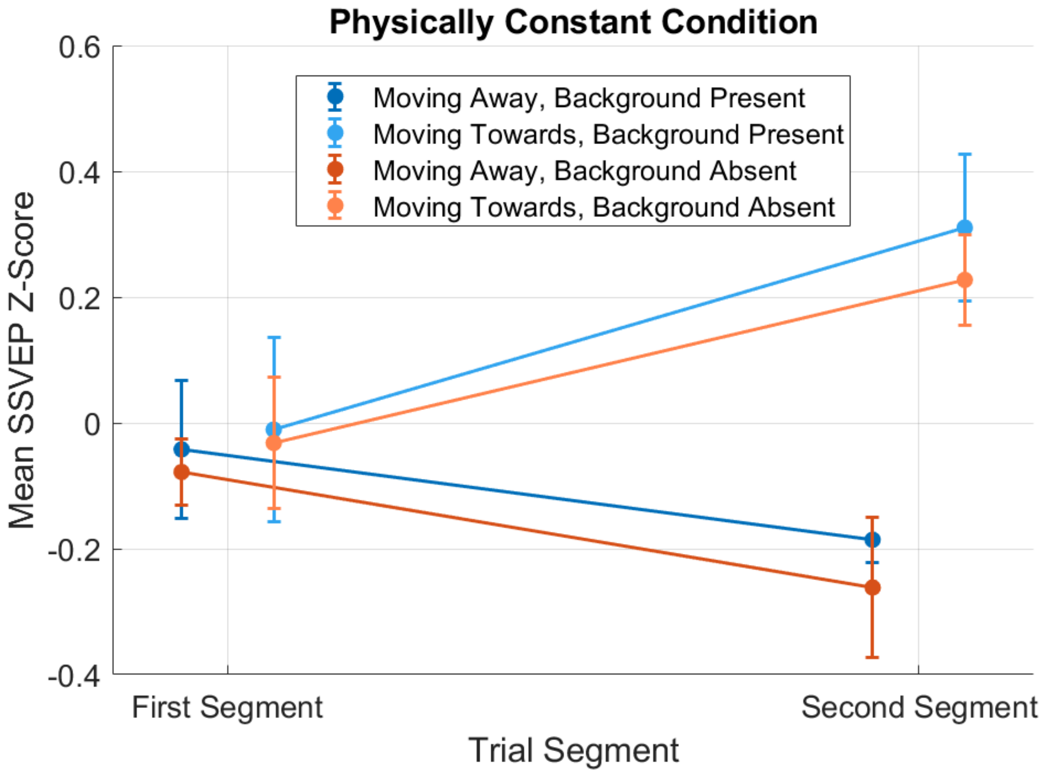


Figure 14: SSVEP strength for stimuli in the physically constant condition. Error bars represent +/- 2 standard errors.

Similarly, the stimulus that moved away from the observer elicits a smaller SSVEP than the stimulus that started at this far distance. The effect sizes are far too small and the data far too noisy for these effects to be significant, but qualitatively this same phenomenon has occurred both in the previous experiment and the current one. Phenomena like these could be indicative of predictive coding; where early visual area activation is influenced by predictable changes in the visual field such as motion (Schellekens, van Wezel, Petridou, Ramsey & Raemaekers, 2016).

In the retinally constant condition, results are again less intuitive and less informative (figure 15). The predicted decrease in SSVEP strength from removing the background observed in the physically constant condition is notably absent here. Variance overall is also greater than in the physically constant condition. Although generally flatter than in the previous experiment, the trends in the data do not indicate perfect retinotopic coding. Nor do they positively correlate with physical size. Notably, stimuli moving away in the retinally constant condition tend to elicit smaller SSVEPs in their second data segment, despite growing physically. This trend of counterintuitive effects of movement and size changes on SSVEP amplitude is in line with the previous experiment.

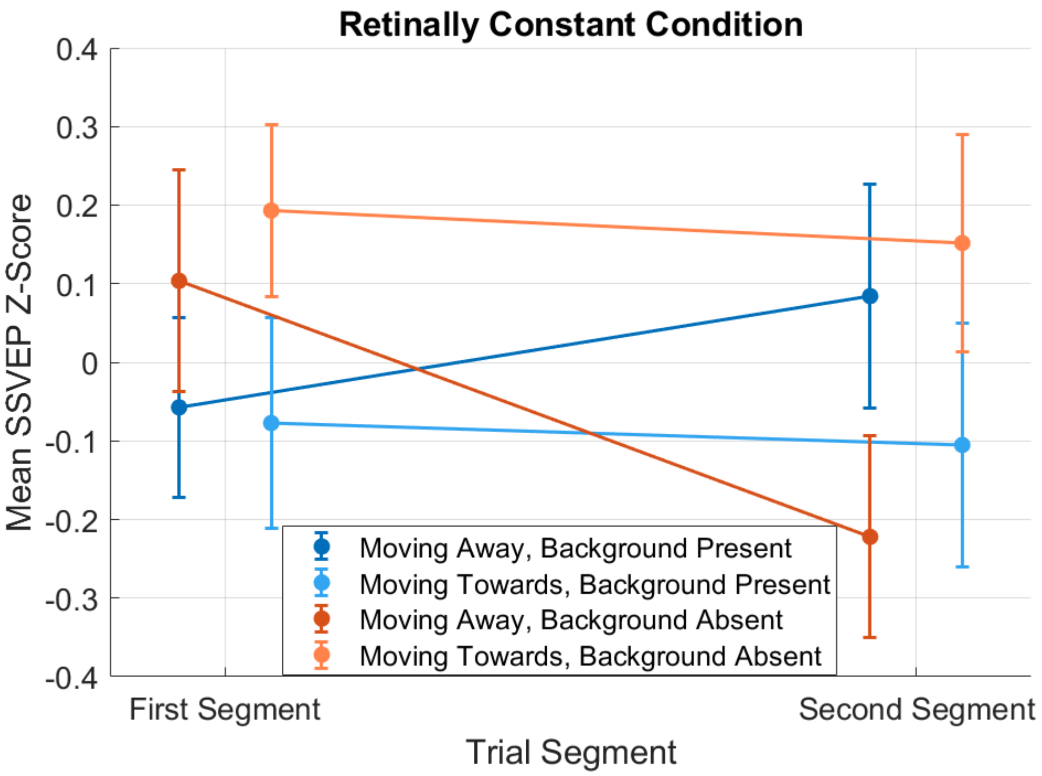


Figure 15: SSVEP strength for stimuli in the retinally constant condition. Error bars represent +/- 2 standard errors.

Despite not being statistically significant in either experiment due to the relatively noisy data, a general trend is starting to emerge that defies expectations. Again, possible explanations could be found when ascribing the effect to predictive coding, combined with a strong prior expectation that variation in retinal size is due to movement as opposed to changes in size, as stipulated by Herschenson's rigidity constraint.

Behaviourally, the presence or absence of a background made little difference to participants' accuracy (figure 16). In replicating the stimulus at an intermediate distance, participants performed accurately in the physically constant condition. Error scores did not differ significantly from 0 in either movement direction, regardless of background. In the retinally constant condition, we replicated the finding of the previous experiment. Objects moving away and growing as they do so are underestimated in size, those moving towards and shrinking are overestimated. Like previously discussed, there are multiple possible reasons for this, and there is little in the literature addressing any of the possibilities empirically.

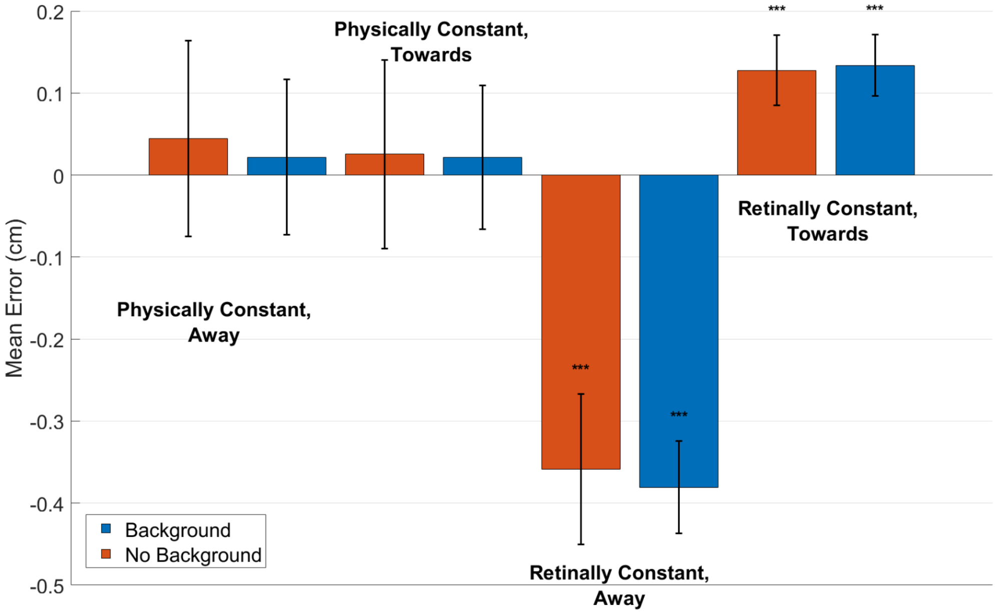


*Figure 16: Size reproduction error scores split by presence of background, movement direction and constancy condition. Positive error scores indicate overestimation. * = p < 0.05, ** = p < 0.01, *** = p < 0.001*

The absence of an effect of a background in participants' behavioural data is reassuring. Although we hypothesized potential changes in accuracy and precision of size perception due to the various ways in which the background can both block depth cues and contribute to them, none of these changes ended up manifesting in the data. With error scores in this and previous experiments not significantly differing from 0, we can conclude we have reached a ceiling effect with regards to performance on this task. This is not a bad thing, since the emphasis of these experiments is to learn more about the neural correlates of size perception, and it is far more valuable to learn about these with regards to normal, functional size perception than to learn about neural correlates of erroneous perception.

With regards to the electrophysiological data, it appears that a background which provides a high degree of contrast does improve SSVEP amplitude, as SSVEPs were notably lower in amplitude without it. It could be argued that this is not inherently problematic as long as data aren't skewed by a floor effect. As long as variations in retinal and physical size effectively drive relative changes in SSVEP amplitude, the SSVEP serves its methodological purpose. Given the increased realism which removing the background provides, and not suffering any methodological cost for this gain, it seems reasonable to leave out backgrounds moving forward.

Experiment 3: Growth and Shrinkage during motion in depth

Having observed hitherto unexplained behavior when people respond to stimuli that are simultaneously moving and changing their size, we designed an experiment to address this phenomenon. Size perception experiments involving moving stimuli are relatively rare, and those combining movement and concurrent changes in stimulus size are extremely rare. Given that size perception relies on the interplay between the sizes of the proximal stimulus, the distal stimulus and the distance to the stimulus, and that determining any of these three factors is subject to noisy sensory processes, one might expect poorer performance when all three of these factors are made variable.

Hershenson (1992) proposed a new version of the size-distance invariance hypothesis which accounts for moving stimuli and makes predictions about how we perceive stimuli that move along the sagittal axis. This kinetic invariance hypothesis (KIH) states that an increase in retinal size of a stimulus will be perceived as an approaching rigid object. Some support for this notion comes from earlier research by Hershenson (1982) himself and Johansson (1975), where this is demonstrated in visually impoverished conditions. A luminous square projected on a screen at a fixed distance in a dark hallway that was made to expand and contract was perceived to be moving towards and away from the observer. Although there is most likely a strong prior in favor of interpreting retinal changes in this manner, it by no means forms a prescriptive rule for the functioning of the visual system. It is no secret that we can very well perceive the growth of objects, and we can even perceive joint changes in size and distance to some degree of accuracy. This can easily be demonstrated by inflating a balloon, letting it go without tying off the end and fixating on the balloon as it flies wildly through the air. You will most likely be able to perceive the fact that the deflating balloon is shrinking despite the fact that its distance to your retina is constantly changing. A more appropriate model for size- and distance perception under normal viewing conditions ought to at least facilitate the possibility of veridically perceiving objects that change their physical size. Such a model could still purely depend on the interrelation between retinal size, physical size and distance, but it would allow for the variability of all three factors.

Gogel (1998) empirically studied size perception of objects which are made to look like they move along the sagittal axis. He notes imperfect size constancy and reports ‘off-size perceptions’, which appear to result from inconsistent combinations of retinal size, physical size and distance. This suggests that observers sometimes fail to appropriately combine size and distance information to form a percept that would be consistent with euclidean geometry (exactly what kind of geometry best describes human optical space remains an open question, see e.g. Luneburg, 1950; Koenderink, van Doorn & Lappin, 2000). It again should be noted that this experiment made use of impoverished visual conditions (monocular viewing, use of a viewing aperture, experiment set in a hallway covered with black cloth) as well as not involving any actual movement, but instead showing expanding or contracting squares at a 2 Hz frame rate on a fixed monitor to mimic movement. Although the questions raised about these intrinsic inconsistencies dealing with size and distance are very interesting, a similar experiment ought to be performed that mimics real life viewing conditions more closely before drawing any conclusions about the abilities of the human visual system. After all, performing experiments under conditions explicitly demonstrated to reduce size constancy is bound to expose errors in size- and distance perception (Holway & Boring, 1941; Lichten & Lurie, 1950).

In this experiment, we aim to address the internal combinations of size and distance information as well as measure the associated neural markers of size perception in early visual cortex. In contrast to our previous experiments, the distance that the stimuli will travel will vary, along with the degree to which they change their size. Stimuli will only move in one direction; away from the observer, and they can either grow, shrink or remain the same size. They will move a variable distance before disappearing. Observers will then be asked to report on how these objects changed their size as well as indicate their final distance. Due to the subtle nature of the change in size (increments between 40% shrinkage and 40% growth in steps of 10%) observers are expected to make errors about the direction of this change. Assuming an internally consistent mental model of size-distance variance, observers are expected to then report the perceived distance which corresponds to their mistaken perception of size change. For instance, if a shrinking stimulus is mistakenly assumed to be growing, the reported distance is expected to be overestimated. After all, a growing object, having a physically larger size, would have to be further removed from the observer to have the same retinal size as a smaller, shrinking object. This type of covariation of size and distance estimates would suggest an internally consistent model of size- and distance perception.

EEG recordings have the potential to be very informative about the role of early visual cortex in size perception. Since previous research has associated size related early visual cortex activity with subjective perception of size (Murray, Boyaci, & Kersten, 2006; Schwarzkopf, Song & Rees, 2011; Schwarzkopf & Rees, 2013), it may be expected that SSVEPs will vary with subjective impressions of size rather than purely retinal size. In other words, a stimulus of 5° visual angle perceived to be a 10cm^2 object at 1.1 meters distance would be expected to elicit a larger SSVEP than that same stimulus if it were perceived as being a 5cm² object at 0.57 meters distance.

The experiment took place in the same virtual hallway as the previous experiments. The experiment again consisted of 80 trials. Trials were randomly distributed across growth and shrinkage rates. Stimuli could either grow or shrink by 10%, 20% or 40%, or not change size at all. The distance traveled by each stimulus was randomized to a value between 150 cm and 350 cm. Stimulus starting position was 150 cm away from the observer. The initial stimulus was identical in all trials, being a grey square subtending 23° visual angle and flickering at 15 Hz with a square wave pattern. The timing of each trial was identical to the previous two experiments, with the stimulus remaining stationary for 4 seconds, moving for 3 seconds and remaining stationary for another 4 seconds before disappearing. After disappearing, participants were presented with a 2-alternative forced choice task where they could indicate whether the stimulus grew or shrank. Following this, they were asked to indicate the location of the stimulus where it had disappeared. This was done by moving a semi-transparent screen to this location using an HTC Vive controller. The screen would appear at the intermediate distance of 250 cm from the observer. There were no time limits on any of the responses.

9 participants took part in this experiment. EEG data collection and preprocessing was performed with the same aperati and using the same methodology as described in the previous experiment. Psychometric functions were created using the Psychinfit 4 toolbox for Matlab (Schutt, Harmeling, Macke & Wichmann, 2015).

Results showed that observers could reliably and accurately tell shrinkage from growth if objects shrank or grew by 40% of their original size (figure 17). This trend appeared to be symmetrical, with 40% growth and 40% shrinkage resulting in equally accurate performance. There does appear to be a slight bias towards reporting objects as shrinking overall, with non-changing objects being reported as shrinking ~70% of the time, and the point of subjective equality being around a 5% growth rate. This might be because of the direction of movement, though since only one direction was included in this experiment, this cannot be definitively concluded.

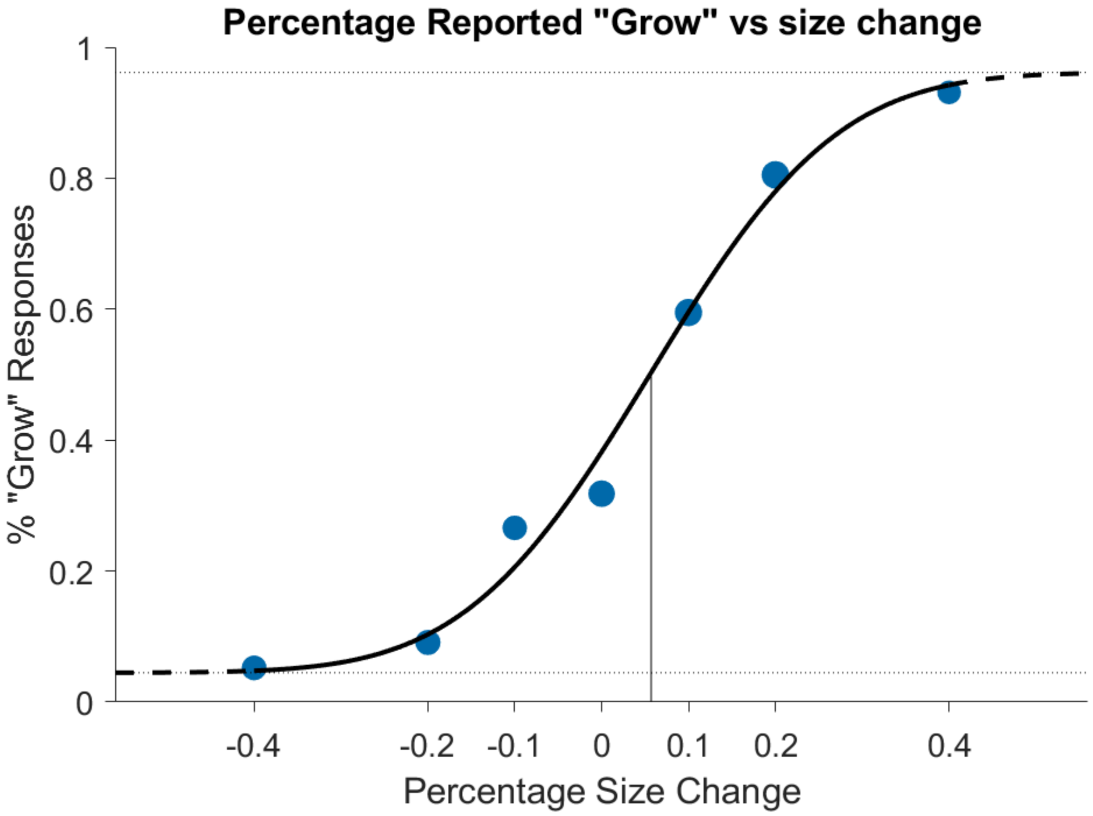


Figure 17: Psychometrically fitted curve fitting percentage ‘growth’ responses to actual size change of stimulus. PSE= 0.0569, 95% CI: [0.0143, 0.0993], slope=0.5550.

Performance on the distance estimation task was incongruently poor compared to the growth/shrinkage judgement task. A general overestimation in estimated distance was observed which seemed to vary with the rate of change in the stimulus’ size (figure 18). A regression analysis was used to calculate the effect of the rate of size change on distance error per subject. Slopes were averaged across participants and compared to 0 using a one sample t-test, which showed a significant deviation from 0 ($t(8) = –3.46$, $p = .009$) with a mean slope of –249.20. These data contain both correct and incorrect judgements of size change by the observers.

The high mean error scores in the distance estimation task are partially driven by a largely right-tailed distribution (skewness coefficient of 1.41, bootstrap 95% CI [1.24-1.58]) (figure 19). Error scores of over 5 meters are not uncommon, which is remarkable since the furthest distance the stimuli could travel from the observer was 5 meters, indicating stimuli were judged to be twice as far away as they actually were. This could be indicative of ambiguity about the distance of objects in the virtual hallway, especially those smaller and further away, a problem thought to have been addressed in the previous experiment. Though errors of over

500 cm are outliers by statistical standards (Q3+1.5IQR), removing them is not appropriate in this situation, since we have no concrete grounds to assume that this was anything other than the observer's best guess.

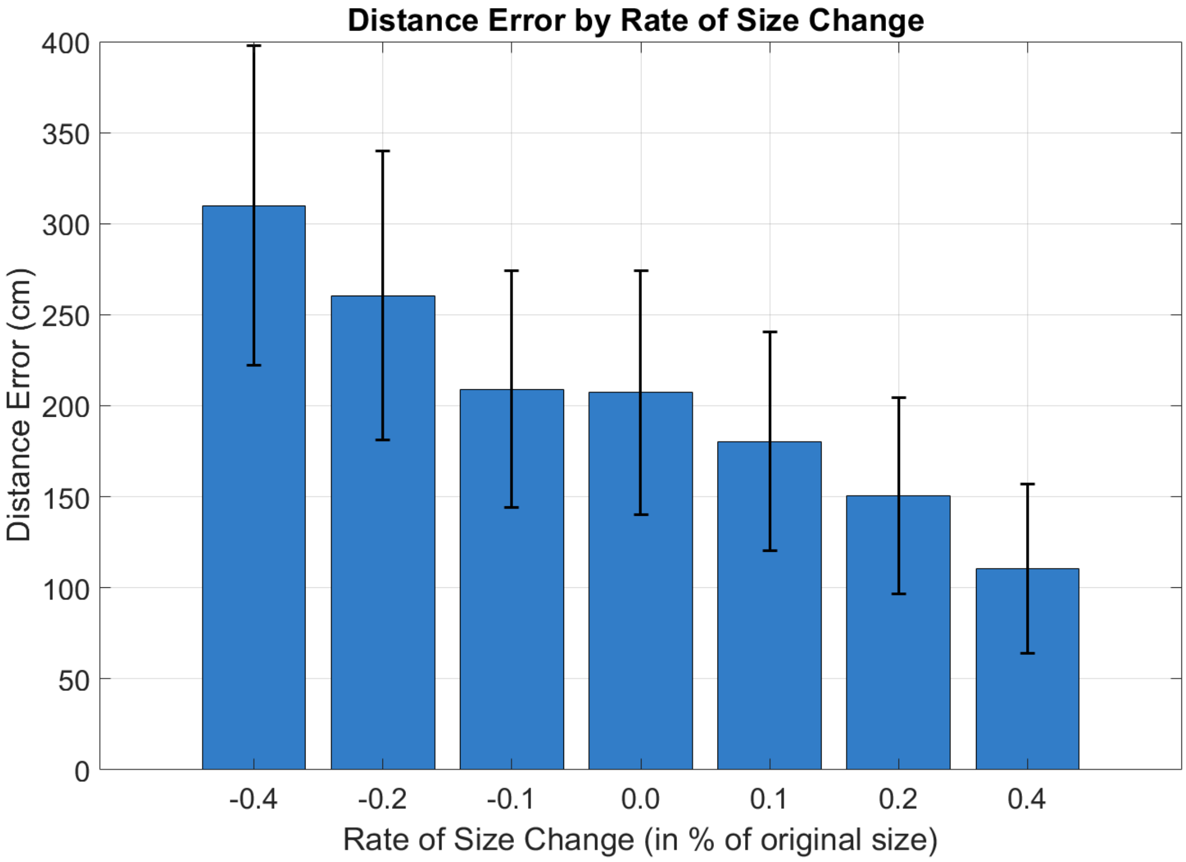


Figure 18: Error in distance estimation task by rate of size change. Positive values indicate overestimation of distance. Error bars represent +/- 1 SE.

We expected accuracy on the growth/shrinkage detection task to influence consequent distance estimation. Indeed, shrinking stimuli that are incorrectly reported to be growing are subjectively located further away than growing stimuli that are incorrectly reported to be shrinking. This indicates an internally consistent model of size and distance, where if the size is incorrectly perceived, the distance is estimated so as to create a situation in which the retinal size makes sense (note that we make no claims about the order in which size- and distance estimations occur). It should be noted that this difference is statistically insignificant, most likely due to the relatively small sizes of these groups. Correctly identified shrinking stimuli are estimated as being even further away, indicating a possible misestimation of the precise degree of shrinkage.

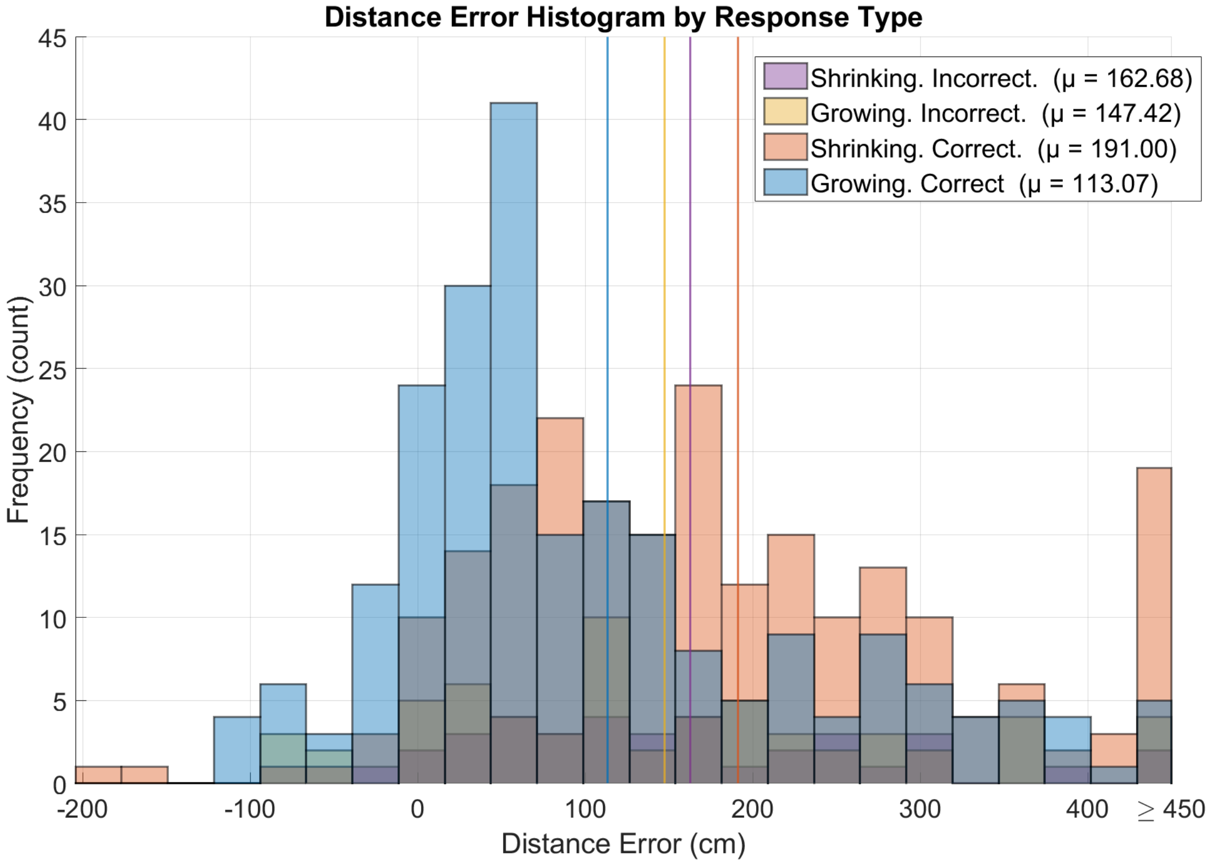


Figure 19: Error in distance estimation split by direction of size change and response accuracy. Positive values indicate overestimation of distance. Vertical lines represent averages.

For the electrophysiological data, a rough correspondence was observed between SSVEP amplitude and the change in stimulus size (figure 20). Generally, stimuli that grew over the course of a trial elicited larger SSVEPs than those that shrank. Data quality unfortunately did not allow a more fine grained analysis of the contributing factors to SSVEP amplitude. Although per-trial information about retinal size and physical size is available, neither of these variables have enough predictive value to be significant in a regression analysis. Comparisons of these models, then, is similarly inconclusive. Although a model using angular size as a predictor yielded a slightly lower Akaike Information Criterion (1810.80) than a model using physical size (1811.39), and a model including both predictors performed slightly worse (1812.72), the differences between these models (ΔAIC < 2) were minimal. This emphasizes the need for purpose-built experiments when it comes to working out the role of retinal size versus physical size in early visual cortex.

Although the experiment provided many interesting results, it has several limitations. Even though it is apparent that observers can reliably detect changes in size once this change is 40% of the original size or more, we do not know what the perceived rate of change is. Just like a 10% decrease in size can evidently be misconstrued for an increase in size, it can equally be misconstrued for a decrease of more than 10%. Consequent estimates of distance of this object are then expected to be erroneous in a manner consistent with this mistake. This type of *within-category error* is likely responsible for a large degree of error in distance estimation observed in this experiment. Another limitation in this experiment is the inability for observers to report objects as neither growing nor shrinking. Like previously discussed, it is likely that there is a strong perceptual prior towards seeing objects as rigid and invariant in their size. Multiple trials in this experiment were likely perceived as not changing in size, though we cannot tell which ones these might have been.

The general overestimation of distance of stimuli in this experiment was surprising. Although it is more pronounced for stimuli that are shrinking, the distances of growing stimuli as well as stimuli that do not change their size are also overestimated. This might be due to the direction of movement, though since this was the only direction of movement in this experiment, we cannot say this for certain based on these data alone. Logically, a general overestimation of distance would imply a corresponding overestimation of size, which would be required to explain the proximal stimulus. Still, the general bias in these results is in the direction of shrinking, with the PSE being at around 5% growth. Further, previous experiments have demonstrated that size perception for stimuli that do not change their size as they move away is quite accurate. If size perception remained accurate in this experiment, the overestimation of distance, at least for these unchanging stimuli, is a definite violation of the size-distance invariance hypothesis. On a large scale, then, these results seem counterintuitive. The missing link in understanding these findings might be in a unique contribution of the direction of movement, or a more refined approach to probing the perceived rate of size change. These are both promising avenues for future research.

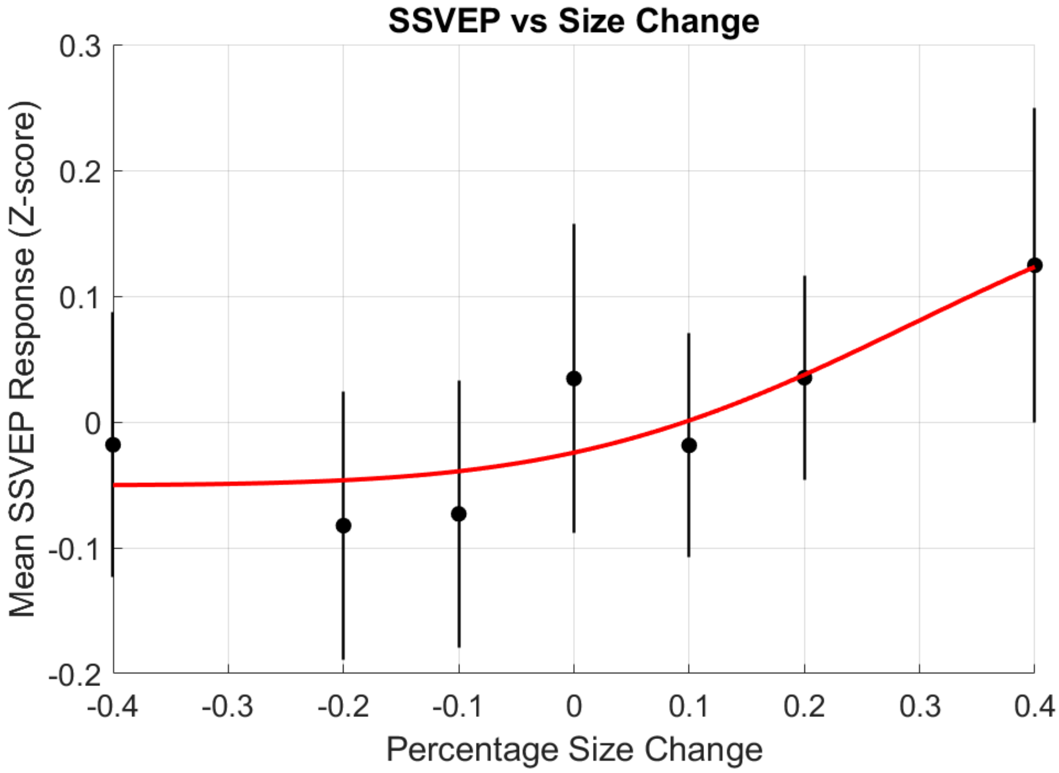


Figure 20: SSVEP amplitude plotted against percentage size change of stimuli.

The SSVEP amplitude proved to be a neuromarker for size, albeit one with a signal-to-noise ratio that leaves something to be desired. Differences in retinal size, physical size and perceived physical size as implied by indicated distance are very subtle in this dataset, and the SSVEP did not have the required specificity to differentiate between them. As the primary goal of this line of experiments is to study the neural substrates of size perception in early visual cortex, and teasing apart the influences of retinal size and physical of stimuli forms is required for this goal, future experiments will again focus on blocked designs where retinal size and physical size are disambiguated in a more categorical way.

The extremely skewed nature of the distance errors does raise some concerns about the virtual environment in which the experiment was conducted. The extremely large misestimations might in part be due to the ambiguity that arises from a floating object in virtual reality that does not appear to be attached to any part of the physical environment. This phenomenon is less pronounced with a larger stimulus that occludes the surrounding surfaces, providing valuable depth cues. It bears reexamining in future experiments.

Experiment 4: Comparing shrinking, static and moving stimuli

An as of yet unexplained phenomenon observed during experiments 1 and 2 is the counterintuitive SSVEPs observed in the retinally constant conditions. When stimuli move away from the observer in these conditions, they physically grow. One would hypothesize that this would lead to the resultant SSVEP either remaining constant (if the SSVEP mainly codes for retinal size) or increasing in amplitude (to the extent that the SSVEP codes for physical size). The observation that the resultant SSVEP amplitude decreases is therefore extremely counterintuitive. The opposite holds true for stimuli that move towards the observer in the retinally constant condition. These findings suggest that there is a different mechanism altogether involved in driving the SSVEP. Transient, attentional effects have already been excluded as described in the section on preliminary experiments, leaving the possibility that more stable representations of objects held in V1 are affected by the movement or change in size of this object.

From a simple stimulus feature driven perspective, the stimuli in the retinally constant condition do two things simultaneously; they move and change their size. A simplified experiment, then, would separate these factors and study them in isolation. Under the hood, however, more complex questions about the neuroscientific nature of this phenomenon persist. A key factor may be that in the retinally constant conditions, stimuli behave in unexpected ways.

The stable representations of objects in V1 which do not merely reflect retinal size, but also physical size and attention are most likely a result of feedback signals from neural regions higher up in the processing order (Murray et al., 2006; Fang et al., 2008). The neural mechanisms of this are not fully clear yet, though recent work in mice has demonstrated the role of cortico-cortical feedback connections from extrastriate cortex on V1, and how these can have either a suppressive or and excitatory effect depending on their location (Fişek et al., 2023). In humans, feedforward and feedback signals to and from V1 are organized along the laminar structure of V1. Feedforward, bottom-up signals stem mostly from intermediate layers (layer 4) (de Hollander, van der Zwaag, Qian, Zhang & Knapen, 2021; Lawrence, Norris & de Lange, 2019), whereas feedback signals to V1 arrive in superficial layers (layers 2 and 3) (Muckli et al., 2015) and deeper layers (layers 5 and 6) (Bergmann, Petro, Abbatecola, Li, Morgan & Muckli, 2024; Aitken et al., 2020).

Recent findings studying these different layers have also found that not all feedback is created equal. There appears to be a difference between internally created feedback and externally created feedback with regards to what cortical layers are more involved. Aitken et al. (2020) found that prior expectations (internal feedback) caused stimulus related activity in the deep layers of V1. Externally created feedback, being induced by the stimulus itself, like perceptual filling in or visual illusions are more associated with activity in the surface layers of V1 (Muckli et al., 2015; Bergmann et al., 2020). It seems possible, then, that whereas the externally driven phenomenon of size constant representations in V1 as measured by SSVEPs reflect one type of feedback (Chen et al., 2019), our findings in the retinally constant conditions reflect another, internally driven type of feedback.

Though viable in theory, is it difficult to address this distinction using the SSVEP methodology. Whereas high resolution (often 7T) fMRI can be used to visualize layer specific neural activity in V1 (Olman et al., 2012), SSVEPs show activity specifically related to the stimulation frequency, and do not inherently reveal anything about their laminar origin. SSVEPs are generally associated with bottom-up processing (Gordon et al, 2019). This is due to them originating in V1, and intracranial recordings in macaques further confirmed that SSVEPs originate in the intermediate (granular) layers of V1 (Salelkar & Ray, 2020). Conversely, feedback signals are generally associated with frequencies in the alpha and beta-band (~7–17 Hz) (Bastos et al., 2015; Michalareas, Vezoli, van Pelt, Schoffelen, Kennedy & Fries, 2016). These frequency bands are indicative of endogenous visual processing mechanisms, however, and as of yet it is not completely clear how these signals would interact with the frequency tagged activity in the granular layers. Learning more about this would require intracranial recordings associated with SSVEP stimulation and its associated feedback mechanisms, and this falls outside of the scope of this thesis.

If SSVEPs from stimuli in the retinally constant conditions are really (partially) changing in the direction of the expected size of the stimulus when it moves, this would imply feedback signals are boosting or inhibiting the bottom up signals to a large extent via the above mentioned mechanisms. These being top down and expectation driven influences, we would also expect that if size changes occurred in less unexpected ways, SSVEPs would reflect the bottom up processes more, and correlate more with their retinal size. In either case, it would be worth measuring SSVEPs in conditions where size change and movement are separated.

Stimuli in this experiment could behave in one of three distinct manners. They could either shrink, move away from the participant or remain stationary (figure 21). The shrinking stimuli would always appear at the far distance (350 cm from the observer), and moving stimuli would always appear at the close distance (150 cm). The far stimulus, on appearance, would subtend the same retinal size as the moving stimulus when it first appeared (23.08°). After shrinking, the stimulus would be the same physical size and be in the same location as the moving stimulus after moving. The stationary stimuli would either appear at the close or the far location, and would be the same size as the moving stimulus. Trials had the same temporal order as in previous experiments, remaining stationary for 4 seconds, changing (moving or shrinking) over the course of 3 seconds and then remaining stationary for another 4 seconds. Stationary stimuli remained still over the entire 11 second presentation time.

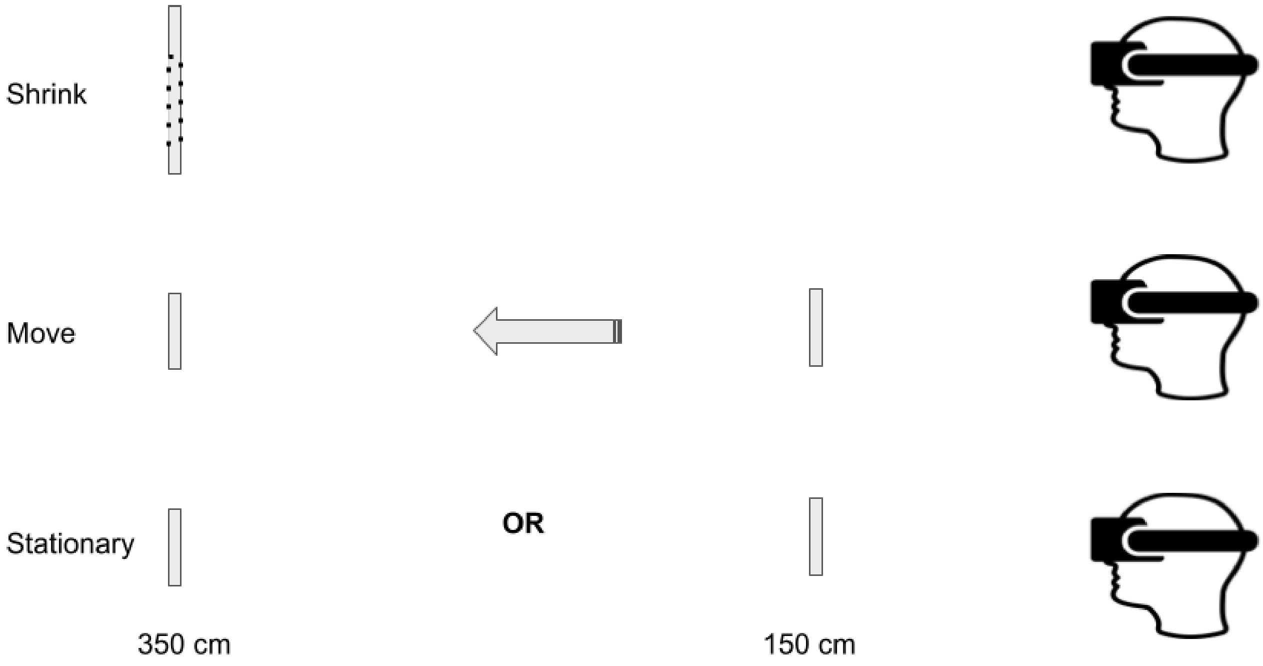


Figure 21: Stimuli conditions. Moving and shrinking stimuli ultimately became stimuli of equal size and distance.

A pilot experiment was conducted consisting of 60 trials; 15 moving, 15 shrinking and 30 stationary equally divided over the two locations. The author took part in this experiment 5 times, and a naive observer took part once, resulting in 6 datasets. This preliminary dataset, though small, indicated a general similarity in the development of the SSVEPs elicited by moving and shrinking stimuli (figure 22). In neither case do the SSVEPs develop in unexpected directions. For the moving condition, SSVEPs in the first segment, when the stimulus is close, are larger than SSVEPs in the second segment, when the stimulus is far away. This was to be expected, as we have previously demonstrated this kind of effect several times already. For the shrinking condition, it appears that in absence of movement towards the observer, the SSVEP decreases with the physical and retinal size of the stimulus. This is also an intuitively logical series of events.

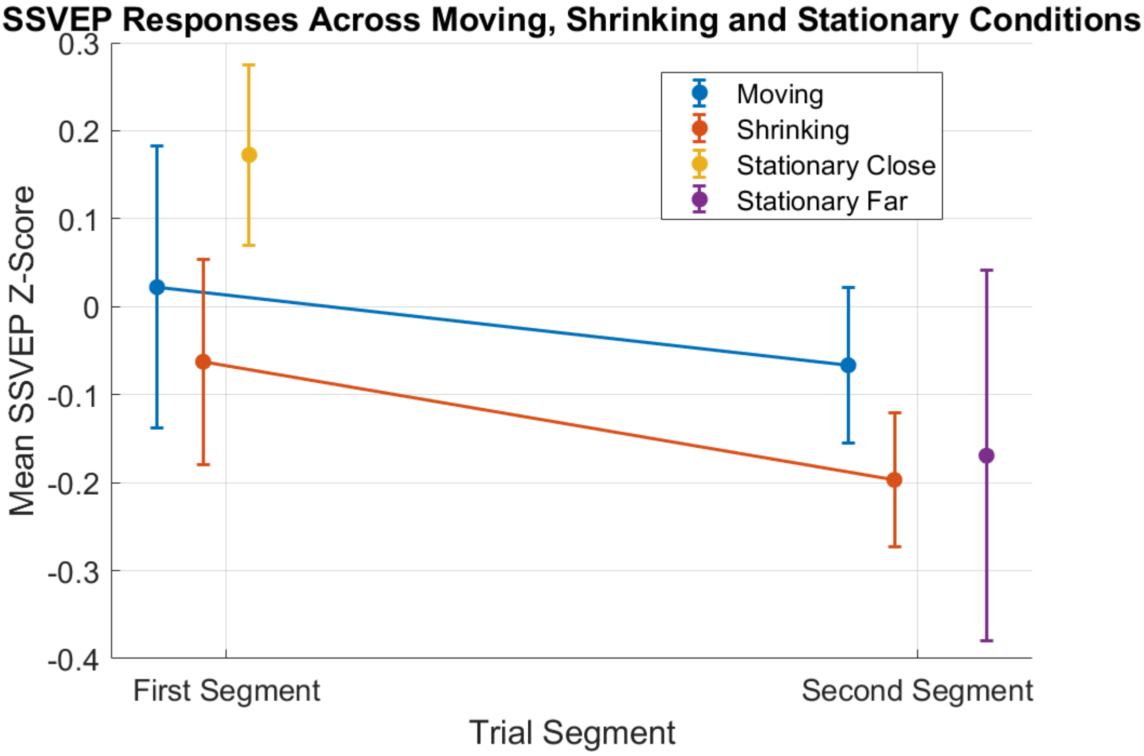


Figure 22: SSVEP amplitudes for moving, shrinking and stationary stimuli. Error bars represent +/- 2 standard errors.

For the full experiment, another variable was introduced. Given the unexpectedly large errors in distance estimation in the previous experiment, we hypothesized that a backgroundless stimulus in virtual reality might come across as being much further away than it really is. In an attempt to remedy this, we introduced a type of frame which would be positioned around the stimulus and which would attach to both walls in the hallway (figure 23). While not blocking the view of the surrounding hallway like the original background used in earlier experiments, this frame would still be a strong depth cue, as it would anchor to the wall laterally to the stimulus. The intention was to provide a stimulus to the observers with a less ambiguous distance.

9 participants took part in this experiment. EEG data collection and preprocessing was performed with the same aperati, software and methodology as described in the previous experiment. Each experiment consisted of 80 trials, distributed over moving, shrinking and stationary stimuli, and split by the presence or absence of a frame around the stimulus. Shrinking stimuli were only presented at the far distance and moving stimuli were only made to move away in order to reduce the dimensionality of the experiment and increase the number of trials per condition, thereby increasing statistical power. The time course of a trial was identical to the pilot experiment. Additionally, participants were asked to indicate the location of where they last saw the stimulus. This was done by moving a semi-translucent screen to this location.

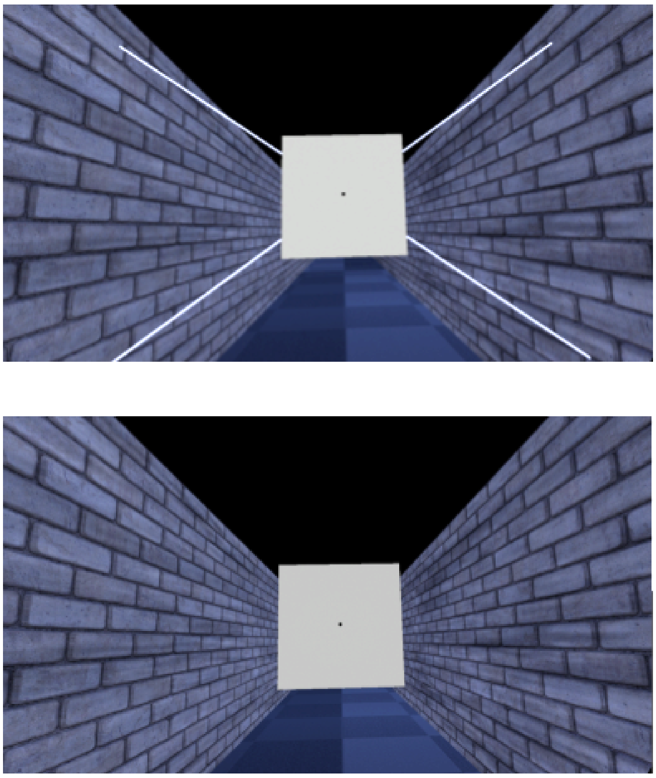
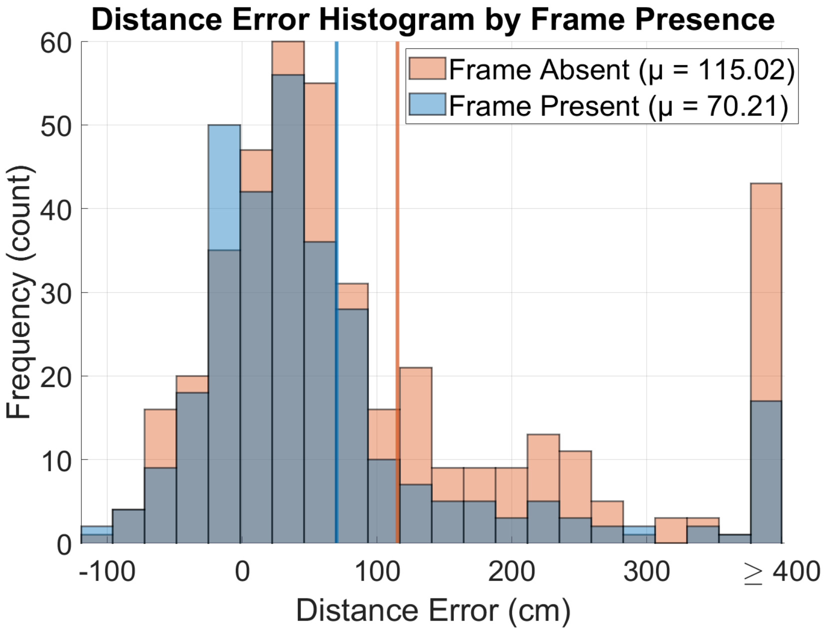
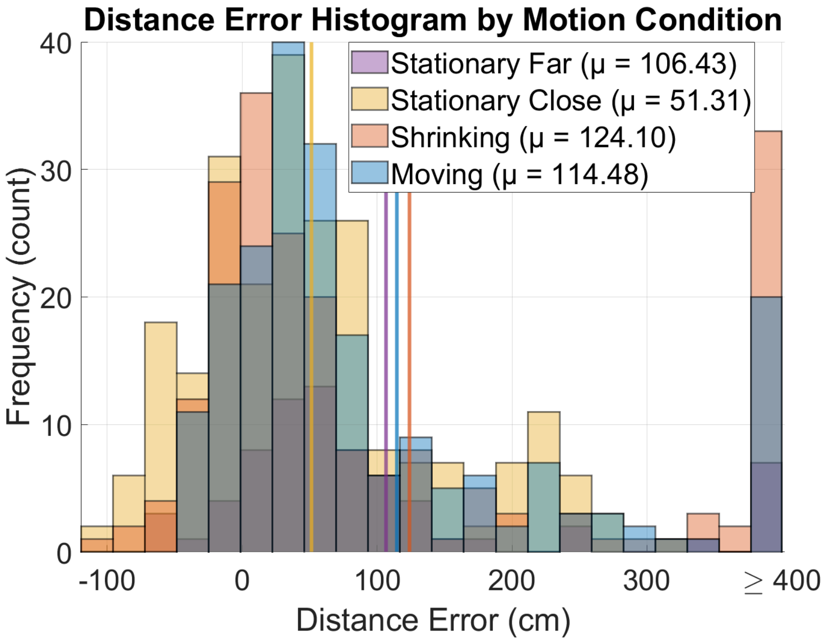


Figure 23: Stimulus with (above) and without (below) the ‘frame’ attachment

Behavioural data shows that overall, distance estimation errors were smaller in cases where the stimulus frame was present. Although not significant when comparing means (*t*(8)=-1.47 *p*=.18 N.S.), the main difference appears to be in the right-tailed distribution (figure 24). Like observed in the data from the previous experiment, errors of 300 to 800 cm occur, seemingly stemming from the illusion that the stimulus is floating very far away down the hallway, possibly near the horizon. The frame seemed to reduce this illusion somewhat, although it does not eliminate it altogether. Also similar to the previous experiment is the finding that stimulus distances are generally overestimated. Interestingly, distance misestimation appears to be a function of distance alone, being unaffected by whether stimuli moved to that location or not.



(A)



(B)

Figure 24: Distance estimation errors split by presence of stimulus frame (A) and by stimulus category (B). Positive values indicate overestimation of distance in centimeters.

Electrophysiological data are generally consistent with those from the pilot experiments (figure 25). Both moving and shrinking stimuli show a decrease in SSVEP amplitude when comparing the first and second segments, indicating that the SSVEP varies with stimulus size. Importantly, neither condition shows counterintuitive results like those in the retinally constant conditions of earlier experiments. This supports the notion that the combination of moving and changing size, especially in an unexpected way, alters the neural computations that give rise to the SSVEP in such a way that it no longer represents size, retinal or physical. What exactly the nature of this shift of the functional significance of the SSVEP is, and how it manifests in the various layers of V1 is still an open question.

A potentially worrisome detail seen in the EEG data is the amplitude of the shrinking stimulus in the first data segment compared to the amplitudes of the moving and stationary stimuli. Recall that the shrinking stimulus, although having an equal retinal size to the moving and stationary (close) stimuli, was physically larger. Previous research suggests that this psychically larger stimulus ought to elicit a larger SSVEP than the physically smaller ones, given equal retinal size (Chen et al., 2019). With EEG data quality being a constant concern when using electrophysiological measures in combination with VR HMDs, this may be a result of noise, although it bears further investigation.

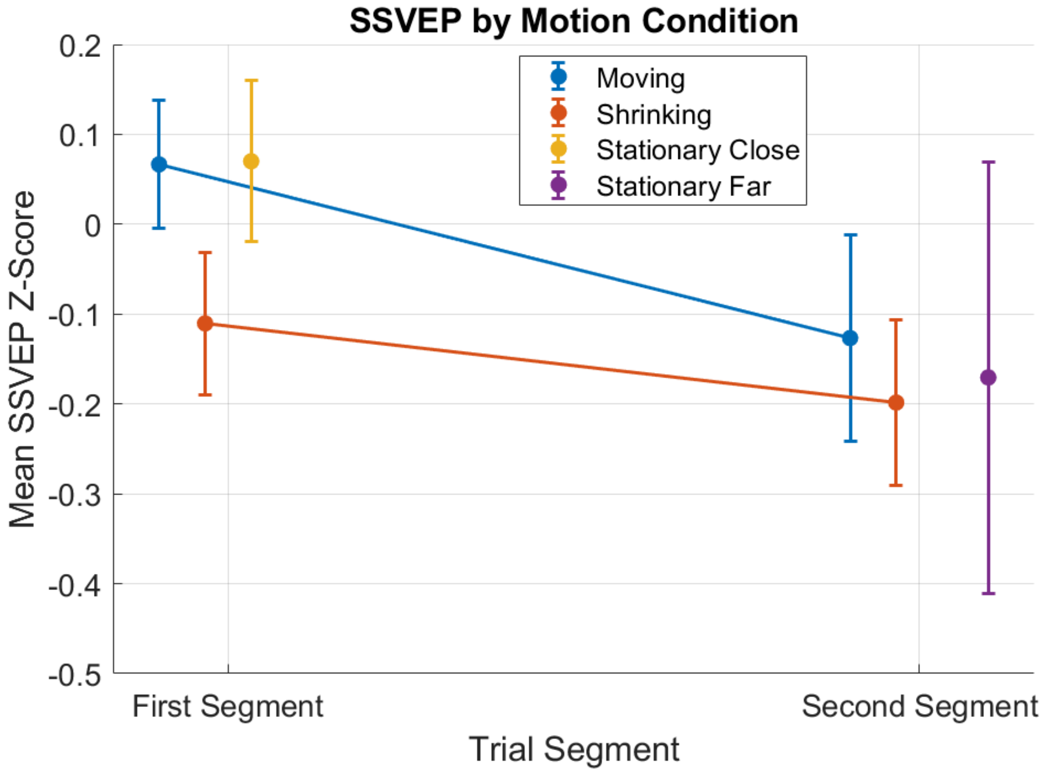


Figure 25: SSVEP amplitudes for moving, shrinking and stationary stimuli. Error bars represent +/- 2 standard errors.

In conclusion, separating movement with changes in size leads to SSVEP amplitudes that again coincide with stimulus size. This indicates that, if the previously observed trends in the retinally constant conditions hold, unexpected combinations of movement and size change elicit activity in early visual cortex that causes the SSVEP to no longer code for either retinal or physical size. This phenomenon is highly likely a result of feedback processing. How exactly this process would work, whether it be due to higher up visual areas sending inhibitory or excitatory signals to retinotopic areas of V1 expected to be activated based on predictions extracted from stimulus movement, or some other mechanism, is as of yet unknown.

A major issue preventing the drawing any definitive conclusions at this stage is the poor data quality. Wearing a VR HMD over top of the EEG cap introduces a high degree of noise to the signal, mostly due to the physical interference of this HMD with the electrodes. Although sample sizes used in previous experiments might be sufficient in traditional SSVEP experiments, when using VR these must evidently be increased to offset this lower data quality.

Experiment 5: The effect of motion alone.

Due to the lackluster data quality and the resultant lack of statistically significant differences between conditions in EEG data, we decided to again reduce the dimensionality of our experimental setup and increase the sample size. In an attempt to definitively establish whether or not stimulus movement alters the SSVEP evoked by a given stimulus, we designed an experiment that contrasts stationary stimuli with moving stimuli. Additionally, given the minor differences in methodology from Chen et al. (2019) such as the presence of a fixation dot on the stimulus, the frame surrounding the stimulus and the texture of the virtual walls and floor, we included stationary stimuli of such retinal and physical sizes as to allow us to replicate the findings of partial size constancy found by this earlier paper.

Based on data from the first two experiments which involved physically constant stimuli moving either towards or away from the observer, it appeared that movement alone might have contributed to the final SSVEP amplitude. Stimuli moving away from the observer elicited smaller SSVEPs than those of equal size and distance that had not moved (yet). The opposite was true for stimuli moving towards the observer. Three conditions split over the first two experiments contained conditions that allowed for the calculation of an effect size of this phenomenon. By comparing the difference between SSVEPs elicited by two stimuli of equal physical size at two distances before and after movement (the stimuli having switched places by moving), we can deduce a possible unique contribution of movement. Subtracting the difference in SSVEPs from stimuli pre movement from those obtained post movement resulted in positive values for all three conditions, meaning that the difference in SSVEPs was made larger due to the movement. Table 4 contains these calculations.

Dataset	Close stimulus (Before)	Far stimulus (Before)	ΔBefore movement	Close stimulus (After)	Far stimulus (After)	ΔAfter movement	ΔAfter – ΔBefore
Experiment 1	0.03762	–0.08849	0.12611	0.3574	–0.2793	0.6367	**0.5106**
Experiment 2 Condition 1	–0.042	–0.011	0.031	0.310	–.185	0.4950	**0.4840**
Experiment 2 Condition 2	-0.078	-0.032	0.0460	0.227	-0.262	0.4890	**0.4430**
Mean	—	—	—	—	—	—	**0.4792**
SD	—	—	—	—	—	—	**0.0341**

Table 4. Values and differences between SSVEP Z-scores before and after moving. All stimuli were of equal physical size, and did not change their physical size over the course of moving.

Taken on average the effect size (calculated by dividing the mean difference by the standard deviation) appears extremely large (Cohen’s D= 14.05). It should be noted that this calculation does not factor in the noisy nature of the data in the experiments; the individual comparisons themselves might not be significant. It is therefore likely that this effect size is an overestimation. Since this experiment will only contain one movement direction in which we will compare moving stimuli versus stationary ones, and the above difference scores are based on the comparison between two stimuli that both move, we must adjust the expected effect size.

Assuming the added effect of movement on resultant SSVEPs is symmetrical for both directions of movement, the expected effect size when comparing a stimulus after moving away with an identical stimulus that had not moved would be the above calculated effect size divided by two (7.03). A power analysis using G*Power 3.1 (Faul, Erdfelder, Buchner & Lang, 2009) was unable to find an adequate sample size prediction, since its calculations do not allow for effect sizes greater than 1. Using a hypothetical effect size of .99, however, it predicted that a sample size of 3 would be required to achieve 99% power (1- chance of false negative). A priori, then, our chances of finding this effect looked almost certain.

The goal of this experiment was twofold: to find out whether stimulus movement alone affected neural processing of stimulus in early visual cortex in any given way, and to confirm that the currently used methodology would even be capable of correctly capturing this neural processing by replicating a previously published finding.

Stimuli in this experiment were divided across movement (stationary vs moving) and size (small, medium and large). The experiment consisted of 85 trials in total, with 45 trials containing moving stimuli (15 of each size, all moving away from the observer) and 40 trials containing stationary stimuli (10 small and close, 10 large and far, 10 medium and close, 10 medium and far). The behavioural task was a distance estimation task, identical to that of the previous 2 experiments.

The stimuli were flickering squares with a 10 Hz frequency, oscillating in a square wave pattern with a duty cycle of 0.5. A fixation dot was presented in the middle of each stimulus, and a *frame* consisting of diagonal lines connecting the stimulus to the adjacent walls in the virtual hallway was included.

22 naive participants took part in the experiment. EEG data collection and preprocessing was performed with the same aperati, software and methodology as described in the previous experiment. This experiment did contain eye tracking for the first time, and trials containing blinks were excluded from analysis.

Figure 26 displays the SSVEPs for the moving, medium sized stimulus in blue, with static stimuli corresponding to this stimuli's size and distance when near and far in orange and yellow respectively. Although the difference in SSVEP amplitude between the static stimuli might appear larger than the difference between the first and second segment of the moving stimulus, the important comparison is the non-significant difference between the moving and stationary stimuli. The first segment of the moving stimuli data does not differ significantly from the stationary medium stimulus at a close distance (*t*(21)=1.179, *p*=.25 N.S.). This is to be expected, since at this point these stimuli are identical in every respect. More importantly, there is no significant difference between the second segment of the moving stimuli data and the stationary stimulus presented at a far distance (*t*(21)=.205, *p*=.84 N.S.). Although identical in size and distance, any effect of movement on the elicited SSVEP would have been expressed here. It therefore seems that, despite the near certainty of finding this effect as predicted by the power analysis, movement alone does not alter the representation of objects in early visual cortex.

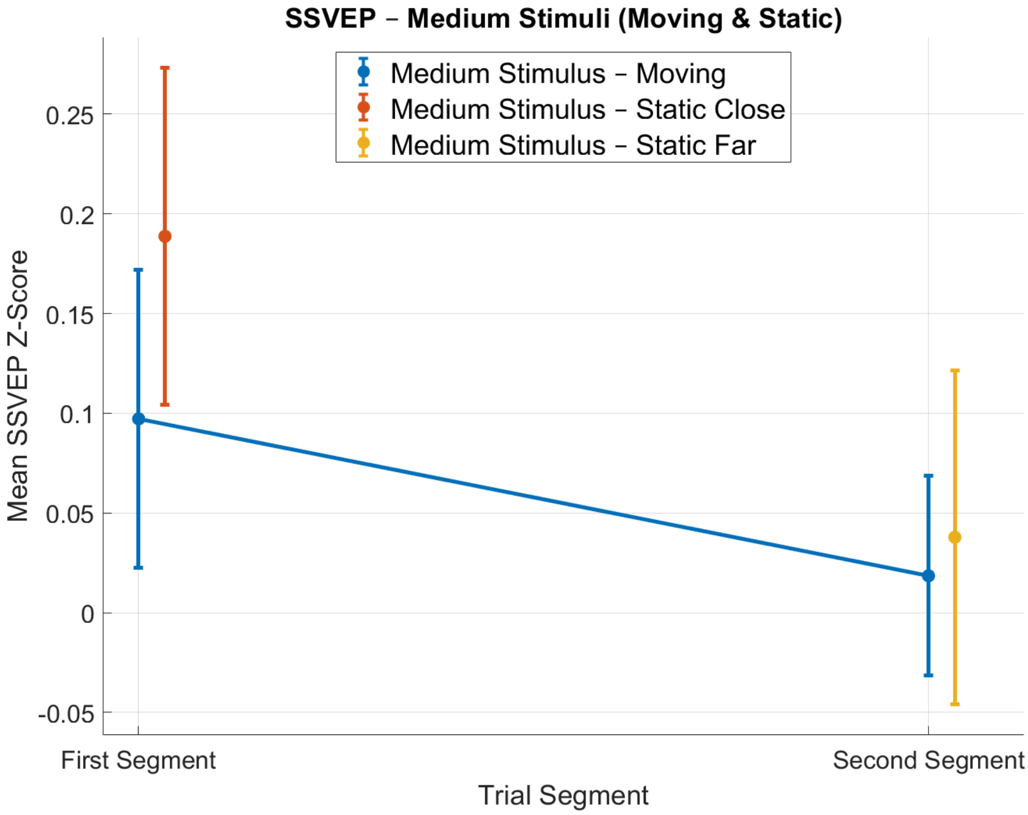


Figure 26: SSVEP amplitudes for a moving, medium sized stimulus, and corresponding static stimuli of equal size and distance. Error bars represent +/- 2 standard errors.

Figure 27 shows the SSVEPs elicited by the stationary stimuli. These stimuli can be paired in two distinct ways. Two stimuli had an identical physical size; the medium stimuli which could appear at either near or far distance. Two sets of stimuli had an identical retinal size; the medium stimulus presented far away and the small stimulus presented close subtended 10 degrees visual angle, whereas the medium stimulus presented close and the large stimulus presented far away subtended 23 degrees visual angle. Comparing these retinal and physical sizes, it is possible to calculate whether the SSVEPs are more determined by retinal size or physical size. Although using the same formula to calculate this size constancy index as Chen et al. (2019) results in a modest degree of size constancy at 18.3%, this formula does not include the SSVEP of the large stimulus presented far away:

$$SSVEP\ size\ constancy\ index\ = \frac{S(medium\ close) - S(small\ close)}{S(medium\ far) - S(small\ close)} \times 100$$

We can devise a similar formula instead using the physically large stimulus and leaving out the small stimulus.

$$SSVEP\ size\ constancy\ index\ = \frac{S(medium\ far)-S(large\ far)}{S(medium\ close)-S(large\ far)}\times 100$$

This formula similarly compares the contribution of physical size in the numerator with the contribution of retinal size in the denominator. Using this formula we obtain a negative size constancy index of -19.4%.

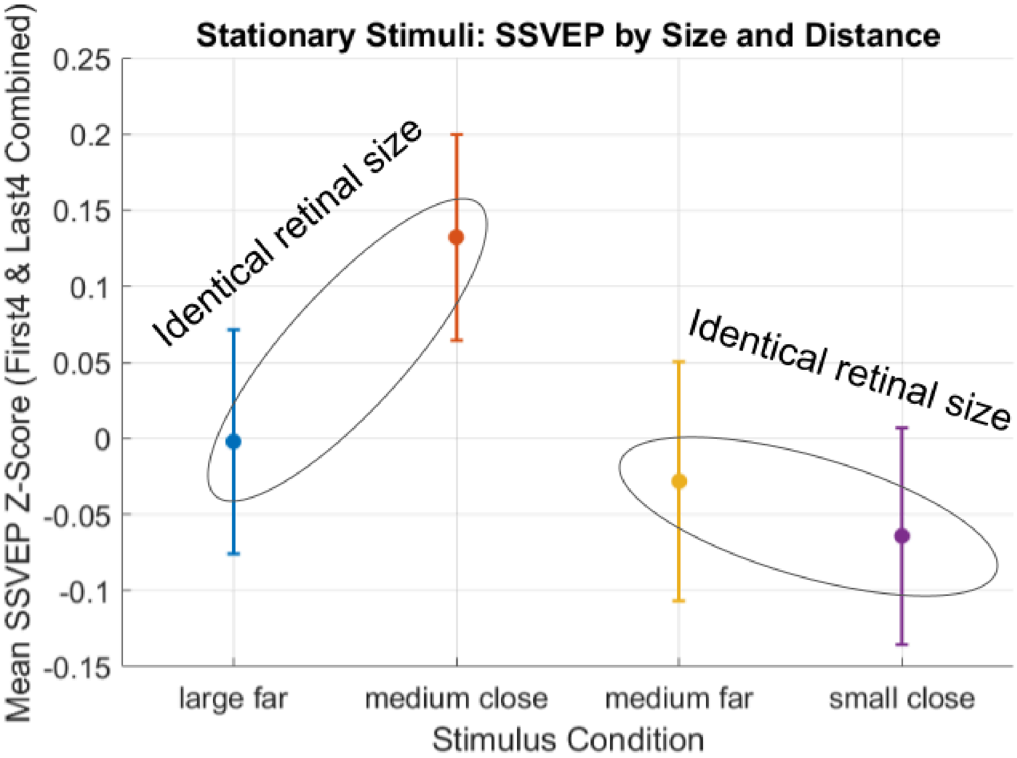


Figure 27: SSVEP amplitudes for all stationary stimuli of various sizes and distances. Error bars represent +/- 2 standard errors.

On the whole, then, these findings fail to replicate the partial size constancy found by Chen et al. (2019). This failure to replicate thereby also casts doubt on the other findings of this experiment that fail to find an effect of motion overall.

The behavioural data is generally consistent with the findings of previous experiments, in that they reveal slight overestimation of distance to the stimulus (figure 28). These overestimation errors again appear to have a long right tailed distribution, indicating that the previously described illusion of stimuli sometimes appearing as being very far away, possibly appearing to be at optical infinity, still occurs sporadically. This illusion seems to be related to the size and distance of stimuli, with 42% of error scores of more than 500 cm being in response to small stimuli, and 31% occurring for medium stimuli presented far away. The distance to large stimuli is never overestimated by more than 500 cm.

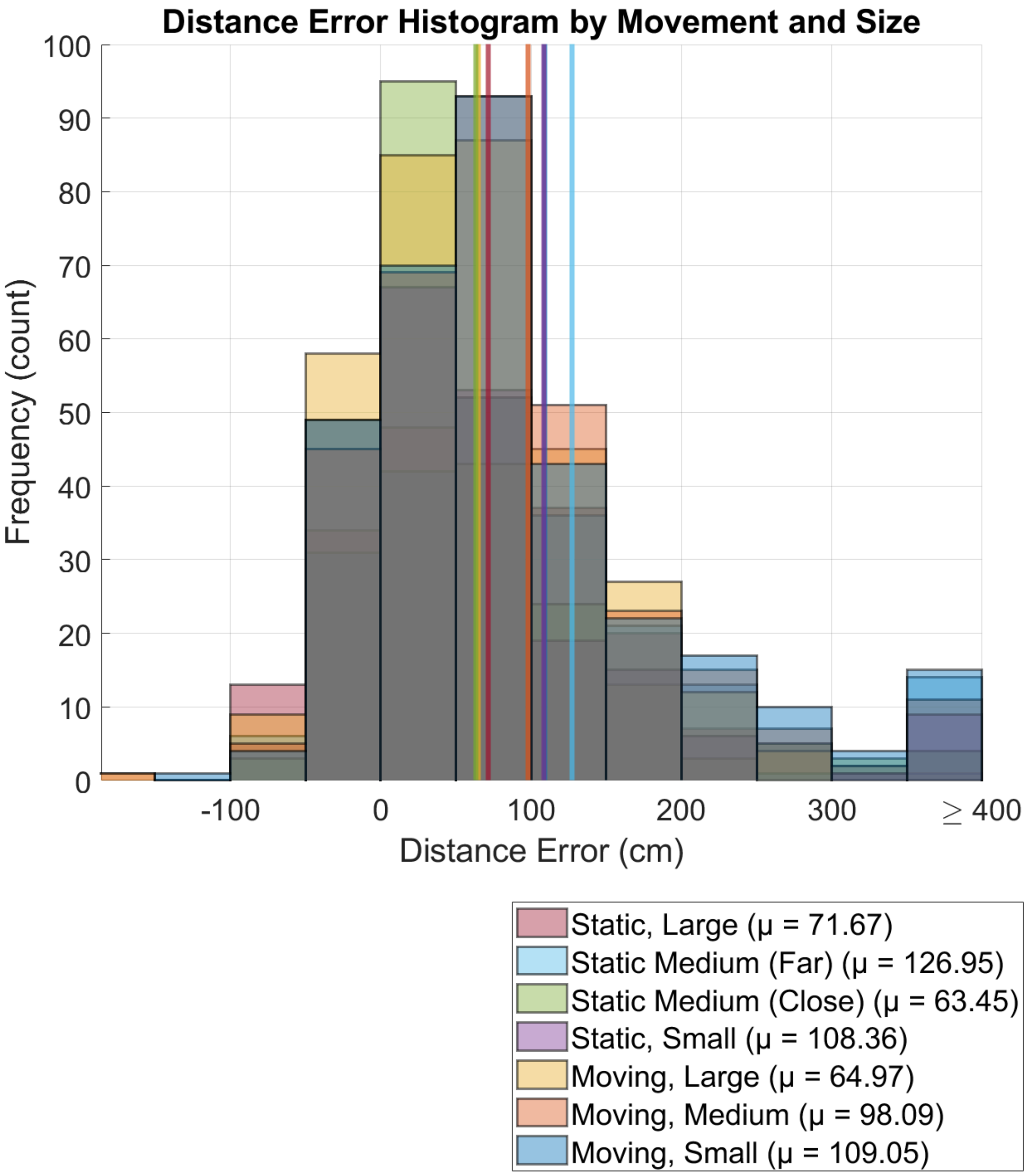


Figure 28: Distance estimation errors split by movement condition and stimulus size.

In conclusion, although we cannot find support for a unique contribution of stimulus movement to SSVEP strength, we also fail to find support for size constant representations in early visual cortex, a phenomenon which has been established using both EEG and fMRI (Chen et al, 2019; Murray et al. 2006). There are several possible contributing factors to this failure to replicate.

First, poor EEG data quality has persistently plagued our findings over the many experiments described here. This is, to some degree, inevitable when combining virtual reality with EEG techniques. Although there is electromagnetic interference caused by VR head mounted displays (HMDs), this is generally contained to 50 Hz line noise and the refresh rate of the screens (Weber et al., 2021). This noise, in the case of data presented here, is filtered out and is of little concern. The more troublesome source of noise is that of mechanical interference with the electrodes. With the types of virtual reality headsets available, there will be a constant interference between the head straps holding the headset on and the EEG electrodes in the cap underneath.

Second, there are more general properties of vision in VR that are known to affect size and distance perception. The vergence accommodation conflict (VAR) refers to the conflicting depth cues generated when wearing a VR HMD. Although observers will generally display vergence behaviour consistent with the virtual distances of objects, the lens will remain accommodated to the distance of the screen displaying this information (Lin & Canny, 2022). This cue conflict has been demonstrated to lead to discomfort (Hoffman, Girshick, Akeley & Banks, 2008) and as well as poorer depth perception, both in and outside of VR (Wang et al., 2024; Durgin, Proffitt, Olson & Reinke, 1995). Distance perception might therefore be distorted in a way that cannot be readily measured by behavioural tasks also taking place in VR, like in our experiments.

To remedy both these issues, it would be necessary to fabricate an experimental setup that allows for independent manipulation of stimulus size and distance outside of VR.

The moving monitor setup

Given the failure to replicate findings from other SSVEP experiments, and given the overall variability in findings resulting from our EEG experiments taking place in virtual reality, we decided to repeat several of our previous experiments outside of VR. This would mean constructing a device that both allowed for stimulus presentation at various distances, and stimulus presentation during movement.

In the size- and distance perception literature, it is rare to find experiments that make use of physically moving stimuli. Even experiments that ostensibly mean to study sagittal motion specifically often make use of illusory motion, Virtual 3D scenes or random-dot stereograms (e.g. Noguchi & Taya, 1981; Gogel & Eby, 1997; Nefs, O'Hare & Harris, 2010; Pai, Rolin, Fooken & Spering, 2017). Notable exceptions do exist, such as Gogel & Tietz (1992), where experimenters manually moved a cart carrying the stimulus along a track towards or away from the observer. A major reason for this dearth of actual, physical motion in perception research is one of practicability. Constructing a mechanized stimulus presentation device is costly, and generally assumed to be unnecessary as other mechanisms of inducing perceived motion are deemed to be good enough. In our case, replacements for real movement have proven to be unsatisfactory, and a device was constructed, impracticable or not.

The eventual moving monitor apparatus consisted of an aluminum frame carrying a 22 inch Samsung Syncmaster bx 2240 monitor with a refresh rate of 60Hz (figure 29). This screen could be moved back and forth and had a range of motion of 60cm. This was made possible by a Festo ELGC-BS-KF-60-200-12P spindle axis, which could move the monitor from one end of the range of motion to the other in 3 seconds. The spindle axis was remotely controlled via MATLAB's Instrument Control Toolbox (The MathWorks, Inc., 2022). Commands were transferred from the control computer to the spindle axis via a series of submodules (a router, a Festo CPX-AP-I4IOL-M12 and a Festo CPX-AP-I-EP-M12). Because the stimulus presentation was programmed in Psychtoolbox-3 for MATLAB (Kleiner, Brainard, Pelli, Ingling, Murray & Broussard, 2007), this connection between MATLAB and the spindle axis allowed for the initiation of monitor movement at specific and predetermined times during stimulus presentation.

Figure 29: The moving monitor. A 22 inch monitor attached to a spindle axis, allowing the screen to be moved back and forth over a range of 60cm.

As a preliminary test of this new setup several datasets were collected with the author acting as the participant. The stimuli were modeled after the stationary stimuli of the previous experiment. That is to say: 4 stimuli overall, with one pair of stimuli being identical in physical size, and two pairs being equal in retinal size. Medium sized stimuli were flickering squares of 11.9cm per side with a square of .25cm per side acting as a fixation dot in its center. These stimuli subtended 4,87° visual angle at the far distance of 140 cm, and 8.81° visual angle at the close distance of 80 cm. The large stimulus was 20.8 cm and subtended the same visual angle when far away as the medium stimulus when it was close. The small stimulus was 6.8 cm and subtended the same visual angle when close as the medium stimulus when it was far away. The stimulation frequency was 8.57 Hz with a duty cycle of .57. Stimulus presentation time was identical to previous experiments at 11 seconds. Six sessions of this experiment were performed, each of which consisted of 80 trials.

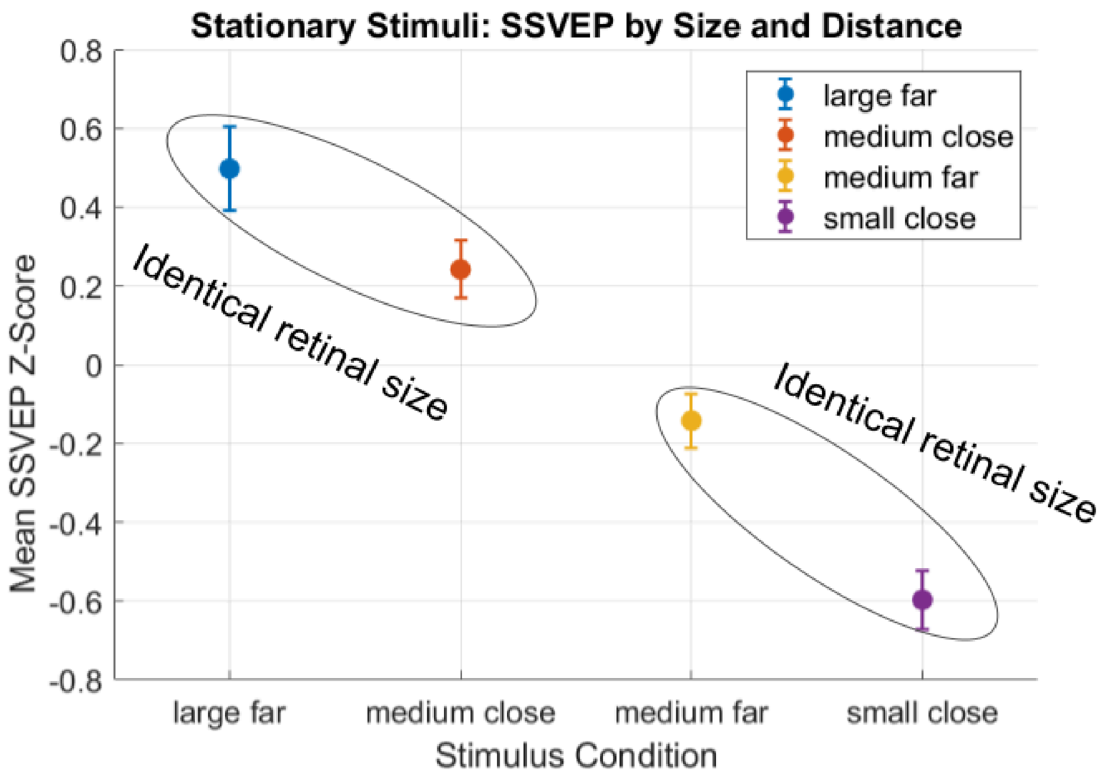


Figure 30: SSVEP amplitudes for all stationary stimuli of various sizes and distances. Error bars represent +/- 2 standard errors.

Results were remarkably different from those in the previous experiment. Where the previous experiment in VR failed to find any evidence of size constancy in the SSVEP data, these pilot data showed significant evidence of size constant representation of objects in early visual cortex (figure 30). Applying the previously used formula: $SSVEP\ size\ constancy\ index = \frac{S(medium\ close) - S(small\ close)}{S(medium\ far) - S(small\ close)} \times 100$, these data show a size constancy index of 178.3%. This result, though apparently indicating a degree of size *overconstancy* (Teghtsoonian, 1974), is more indicative of a shortcoming of the above formula. Notably, the two medium sized stimuli do not elicit an identical SSVEP, which would be indicative of perfect size constant representations in the EEG signal. Instead, the retinally smaller medium stimulus elicits an SSVEP with lower amplitude, indicating a contributing role of retinal size, and a size constancy index of less than 100.

An improved metric for size constancy can be calculated with the following formula: $SSVEP\ size\ constancy\ index = (1 - \frac{S(medium\ close) - S(medium\ far)}{Range}) \times 100$, where *Range* is the entire range between smallest SSVEP and largest. Here, the size constancy index = 100 when the physically identical stimuli elicit identical SSVEPs, and <100 when the retinally smaller stimulus elicits a lesser SSVEP. Division by the entire range of SSVEP values places the difference observed between physically identical stimuli in context, and makes the resulting index more robust against outliers, since it uses all available data points. Using this new formula, SSVEP data from the pilot experiments reveal a size constancy index of 66.21%.

It should be noted that because all six datasets were collected from a single participant, variance in the data is lower than may be expected from an equally sized dataset collected from six different participants. Nevertheless, the data quality as expressed by variance in SSVEPs is noticeably lower than in experiments using VR. This is also evident in the signal to noise ratio (SNR, defined as the stimulation frequency power divided by the averaged power of the 6 nearest adjacent frequencies) being 2,92. For comparison, the SNR in the previous experiment comparing moving to stationary stimuli in VR was roughly half this at 1.63.

In conclusion, the moving monitor setup appears to solve several major problems encountered during previous testing in VR. For one, the data quality is significantly improved, the unobstructed electrodes of the EEG cap result in a much higher SNR than previously encountered. Second, we were able to replicate earlier findings demonstrating partial size constancy in early visual areas, which implies that the earlier failure to replicate was somehow caused by factors specific to the methodology used in VR. Still, with these data only being preliminary and originating from a single observer, the next step would be to perform a conclusive experiment employing a much larger sample. This experiment will not only contrast physical, real world movement with its virtual counterpart, but it will also revisit retinally constant stimuli, which resulted in such counterintuitive results in the initial experiments described here.

Experiment 6: Neural correlates of size perception in dynamic conditions in virtual reality and real life

Using the physical apparatus described in the previous section in conjunction with a virtual reality version thereof, this experiment set out to conclusively answer the foremost question of this thesis, namely whether movement affects object size encoding in early visual cortex. Secondarily, it aims to address methodological concerns about the validity of virtual reality as a research tool, particularly its combination with electrophysiological recordings.

To reiterate, partial size constancy has been demonstrated in early visual cortex. An ever-growing body of evidence suggests that V1 representations encode perceived size rather than only retinal size (Fang, Boyaci, Kersten, & Murray, 2008; Sperandio, Chouinard, & Goodale, 2012; Pooresmaeili, Arrighi, Biagi, & Morrone, 2013). Presumably, the observed size constancy in V1 emerges due to feedback from higher order visual areas (Clavagnier, Falchier & Kennedy, 2004;He, Mo, Wang, & Fang, 2015). Support for this also comes from studies about the time course of size constancy. Chen, Sperandio, Henry & Goodale (2019) demonstrated that size constancy takes at least 150ms to manifest, indicating processing beyond V1 (Lamme & Roelfsema, 2000). More recently Yeh, Gayet, Kaiser, & Peelen (2024) showed that size representation can be decoded from EEG data as early as 90ms after stimulus presentation, but scene modulated size representation occurs around 200ms after stimulus presentation, suggesting that size representations in V1 are adjusted after depth information is retrieved from the surrounding visual environment. Zeng, Fink & Weidner (2020) linked this time course more specifically to neuroanatomical regions. Using TMS, they showed that while early visual areas are critical for size perception in general, TMS disruption of early visual areas disrupted size perception more strongly after higher order regions had been involved. In short, size constant representations are known to exist in early visual areas, and the presence of these size constant representations is linked to functional size perception.

Our own experiments (experiments 1-5) found novel and counterintuitive results when using moving stimuli, though these same experiments failed to replicate established findings from the literature. The present experiment directly contrasts moving stimuli with static stimuli, as well as virtual reality with physical setups, using a large sample size.

Methods

Main experiment

Participants

24 participants (16 female, 8 male, aged 19-54, average age 30) participated in the experiment. All participants had normal or corrected-to-normal vision. All participants gave their informed consent prior to starting the experiment. The experiment was approved by the local ethics committee (LEK 2020-0038) and was performed in accordance with the declaration of Helsinki.

The experiment consisted of two conditions; one physical, presented on a moving monitor, and one virtual, presented in VR. In the moving monitor condition, stimuli were presented on a Samsung Syncmaster bx 2240 monitor with a refresh rate of 60Hz. This monitor was attached to a spindle axis allowing it to move back and forth between and/or during trials (figure 31A). The axis had a length of 60 cm, and it took 3 seconds for the monitor to travel this distance. In the VR condition, a digital, 3D reproduction of this setup was used, and it was presented using a VR head mounted display (HMD) (figure 31B). The VR HMD used was a Tobii Pro VR Integration; an HTC Vive with a Tobii eye tracker integrated into it. This HMD has a refresh rate of 90Hz, a field of view of 110 and a resolution of 2160×1200 (1080×1200 per eye).

EEG data was collected from 13 electrodes (Fz, F3,T7, Pz,P3, P7, O1, Oz, O2, P4, P8, T8, F4) with a sampling rate of 500hz using a Brainvision LiveAmp wireless amplifier (Brain Products GmbH, Germany) and using Brainvision Recorder recording software (Brain Products GmbH, 2024. Electrodes were laid out according to the international 10-20 system. An active reference was used at the FCz location and the ground electrode was placed at AFz. Impedances were kept below 5kΩ.

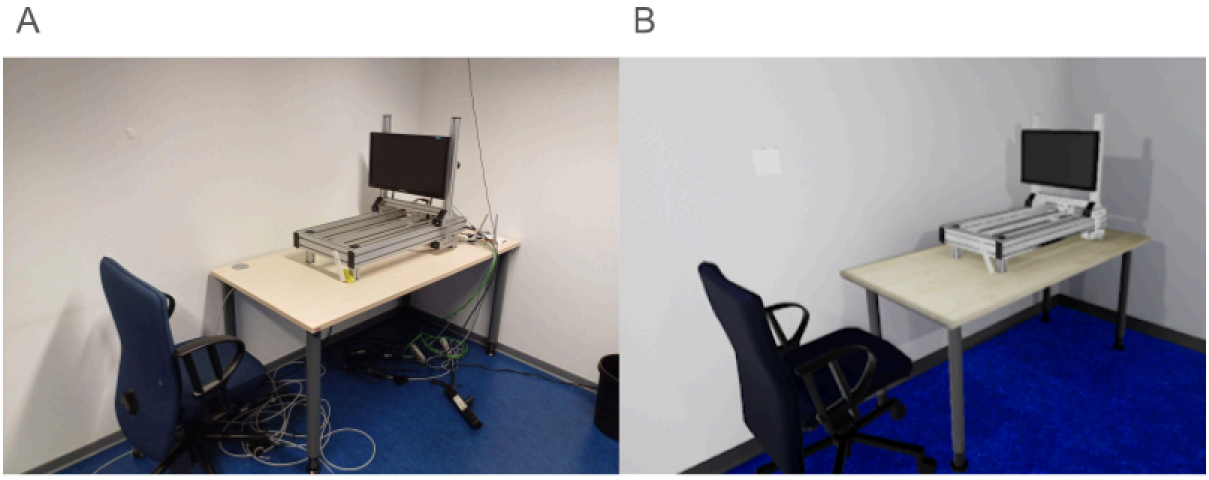


Figure 31: Comparison of the physical experimental setup (A) and virtual experimental setup (B)

The VR and moving monitor conditions were equal with regards to the amount of trials, blocks as well as the size, distance and presentation duration of the stimuli. Each experiment contained 80 trials split into a two-by-two design with as factors: starting position (close or far away) and type of constancy (physical constancy or retinal constancy). Trial presentation order was fully randomised. Trials were divided into 8 blocks of 10 trials each, and blocks were separated by break screens. Inter-block breaks could last as long as the participant wished, but on average only lasted 3 seconds.

Stimuli

The stimuli were squares flickering at 7.5Hz with a duty cycle of 50% and a fixation dot in the middle. Viewing distances were 80cm when the screen was in its near location, and 140cm when the screen was far. A standard stimulus had 11.9cm sides, subtended 4.87 degrees visual angle when far, and 8.51 degrees when near. A retinally constant stimulus that had moved towards the observer would subtend 4.87 degrees all throughout the trial, shrinking to 6.8cm as it moved towards the participant. A retinally constant stimulus that had moved away from the observer would subtend 8.51 degrees all throughout the trial, growing to 20.8cm as it moved away. Stimuli were presented for a total of 11 seconds. They remained stationary for 4 seconds, then moved either away or towards the observer over the course of 3 seconds, and then remained stationary for another 4 seconds before they disappeared (figure 32).

After the stimulus disappeared, a grey square would appear on screen, and participants were tasked with adjusting the size of this new square to that of the stimulus as it most recently appeared. Participants did this by means of a keyboard in the physical condition and an HTC Vive controller in the virtual condition. After adjusting this grey square, participants could prompt the next trial to start by a keypress.

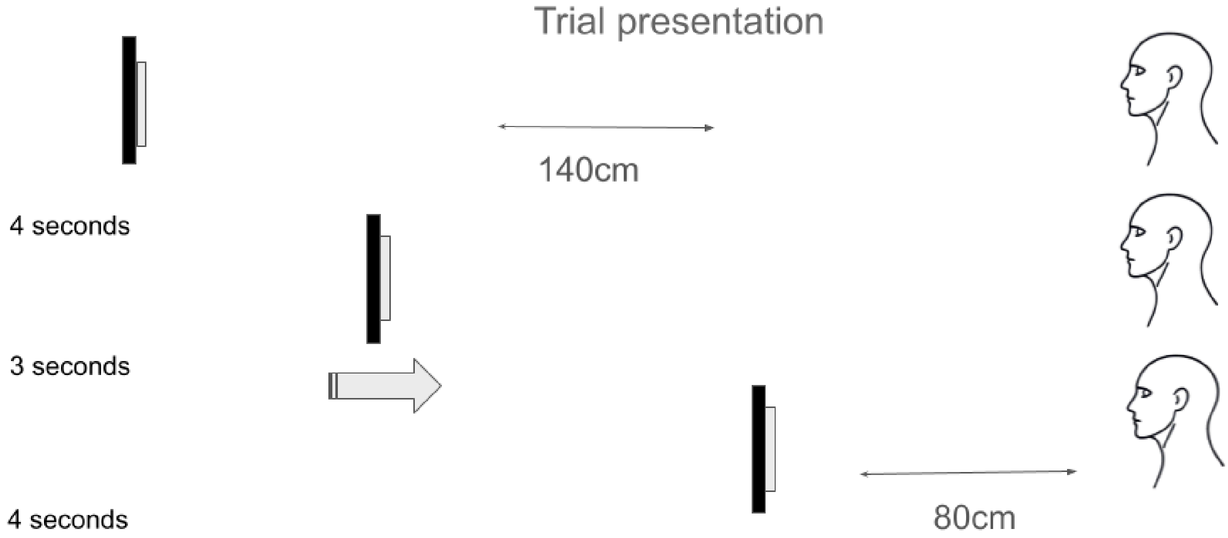


Figure 32: progression of a single trial. Stimuli would appear, remain stationary for 4 seconds, move over the course of 3 seconds and remain stationary for another 4 seconds before disappearing. Although this figure depicts the screen moving towards the observer, the screen could both move towards or away from observers.

Control experiment

Participants

11 participants (7 female, 4 male, aged 20-54, average age 31.5), who had previously participated in the main experiment, took part in this experiment. This experiment was again split into a physical and a virtual part, and took place in the same laboratory space and under the same conditions as the main experiment. All apparati used, both for presentation of the experiment and EEG data collection were identical to the previous experiment.

Stimuli

Stimuli in this experiment, again, were squares flickering at 7.5Hz. Contrary to the main experiment however, stimuli in this experiment did not move. They also remained on screen for 11 seconds before disappearing. There were 40 trials split across 3 stimulus sizes and 2 stimulus distances. At the near distance, either a small (6.8cm) or a medium sized (11.9cm) stimulus could appear. At the far distance, either a medium sized (11.9cm) or a large (20.8cm) stimulus could appear. These sizes were such that at the two distances, 80cm, and 140cm from the observer, the small stimulus would subtend the same angular size as the medium stimulus at the far distance (4.87°), and the large stimulus would subtend the same angular size as the medium one at the close distance (8.51°). This provides one pair of stimuli that are equal in physical size, but not retinal size, and two pairs of stimuli that are equal in retinal size, but not physical size. After stimulus presentation, the same size-matching task was presented as in the previous experiment.

EEG preprocessing

EEG preprocessing was done with a combination of EEGLAB (Delorme & Makeig, 2004) and custom MATLAB scripts. EEG data was filtered with a high pass filter at .5Hz (full width at half maximum of 1Hz) and a low pass filter at 45Hz (FWHM of 11,25Hz). Since using a Laplacian reference is shown to be optimal for analysing SSVEP data (Zhang, Valsecchi, Gegenfurtner & Chen, 2023), the data was re-referenced to a Laplacian reference using the CSD toolbox (Kayser & Tenke, 2006). Data was segmented in accordance with the experiment’s trials. For the main experiment, two segments were created for each trial. The first lasting from 500ms to 4000ms, and the second lasting from 7500ms to 11000ms. This was done to avoid the most egregious of the attentional effects associated with stimulus presentation and movement. Data was detrended, and a Tukey window with an alpha of 0.25 was applied to it. Data was then Fourrier transformed. Per trial, frequency spectra from three occipital electrodes; O1, Oz and O2 were averaged into one spectrum. From these averaged spectra, the stimulation frequency and the first two harmonics were extracted. These peak values were corrected for the spectral power of their surrounding frequencies, subtracting the average of the nearest eight frequencies from them. SSVEP peak values were then summed to form a single SSVEP value, and standardised to z-scores for each participant.

Results

Main experiment

EEG

Paired t-tests were performed to compare SSVEP strength in various conditions, and a correction was applied based on the false discovery rate (FDR) using the Benjamini-Hochberg procedure (Benjamini, & Hochberg, 1995). For the virtual reality condition, results showed no difference in SSVEP strength between near and far presentation distance for stimuli that had constant retinal size, for stimuli either moving towards ($t(21) = 0.0536$, $p = 0.9578$) or away from the observer ($t(21) = 1.7016$, $p = 0.1036$) (figure 33). Conversely, SSVEP strength was different for stimuli that retained their physical size, both when moving towards ($t(21) = -5.2470$, $p = 0.0000$) and moving away ($t(21) = 4.1631$, $p = 0.0004$) from the observer. Further, a cross condition comparison of stimuli of equal physical size and distance showed they did not differ from each other significantly for the far object($t(21) = -0.4113$, $p = 0.6850$), but did for the close object ($t(21) = 3.1941$, $p = 0.0044$).

For the moving monitor condition, the same tests were performed (figure 34).There were no significant differences in SSVEP strength in the retinally constant condition, neither for the moving away condition,($t(23) = 1.11$, $p = 0.28$), nor for the moving towards condition ($t(23) = -0.58$, $p = 0.57$). For the physically constant condition, however, the SSVEP strength differed significantly both in the moving away condition ($t(23) = 5.3838$, $p = 0.0000$) and the moving towards condition ($t(23) = -5.7821$, $p = 0.0000$). A cross condition comparison of stimuli of equal physical size and distance showed no significant difference in SSVEP strength, neither for the far ($t(23) = 0.1311$, $p = 0.8969$) nor for the close object ($t(23) = 0.4797$, $p = 0.6360$)

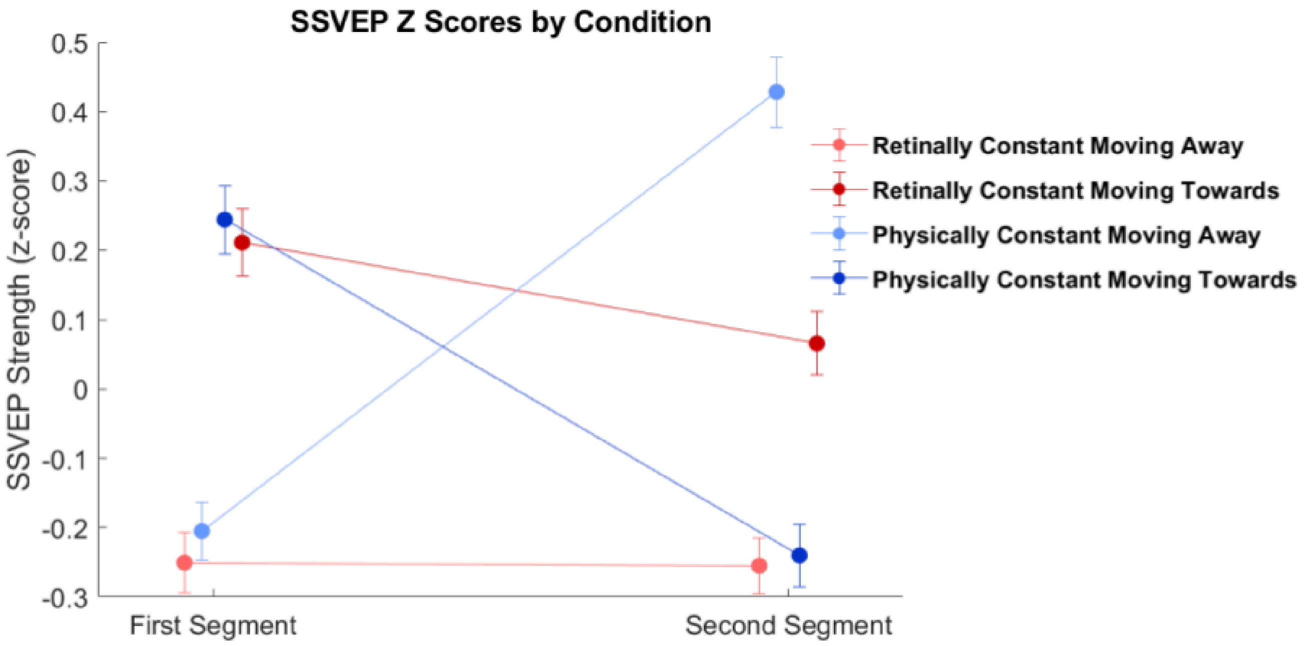


Figure 33: SSVEP strength for retinally and physically constant stimuli before (left) and after (right) moving in the virtual reality condition.

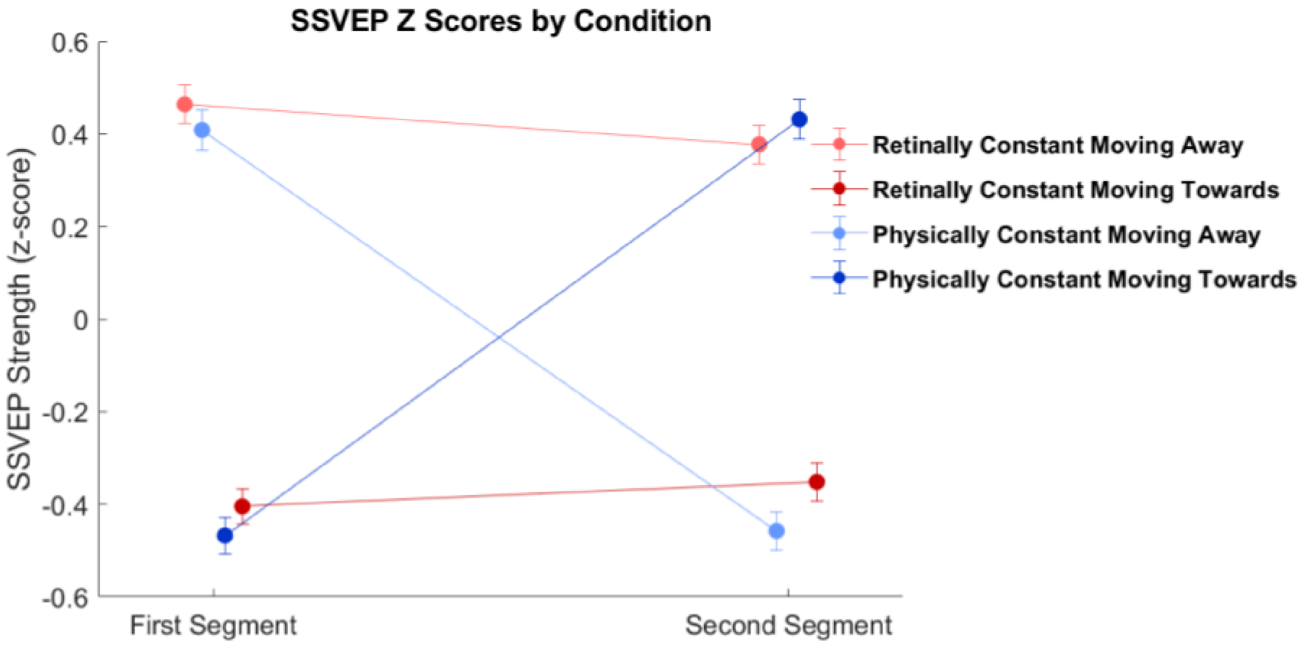


Figure 34: SSVEP strength for retinally and physically constant stimuli before (left) and after (right) moving in the moving monitor condition.

Size constancy indices were calculated, which express the degree to which the SSVEPs are driven by the retinal size or the physical size of a stimulus (Chen, McManus, Valsecchi, Harris & Gegenfurtner, 2019). Indices are derived via the following formula.

$$SSVEP\ size\ constancy\ index\ = \frac{S(medium\ close) - S(small\ close)}{S(medium\ far) - S(small\ close)} \times 100$$

A size constancy index of 100 implies that the SSVEP is driven entirely by physical size, indicating perfect size constancy in the neural signal. A size constancy index of 0, on the other hand, would indicate that the SSVEP is driven entirely by retinal size, with no role of physical size.

In the virtual reality condition the SSVEP size constancy index was 0.99 indicating a lack of size constancy in the neural signal (figure 35).

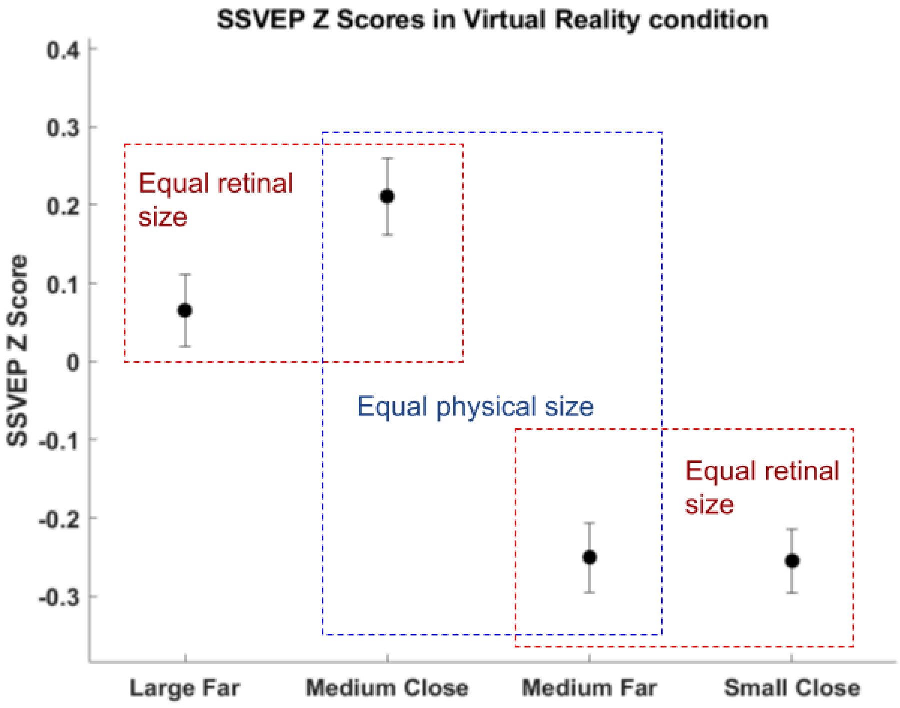


Figure 35: SSVEP strength for stimuli in the retinally constant condition in virtual reality.

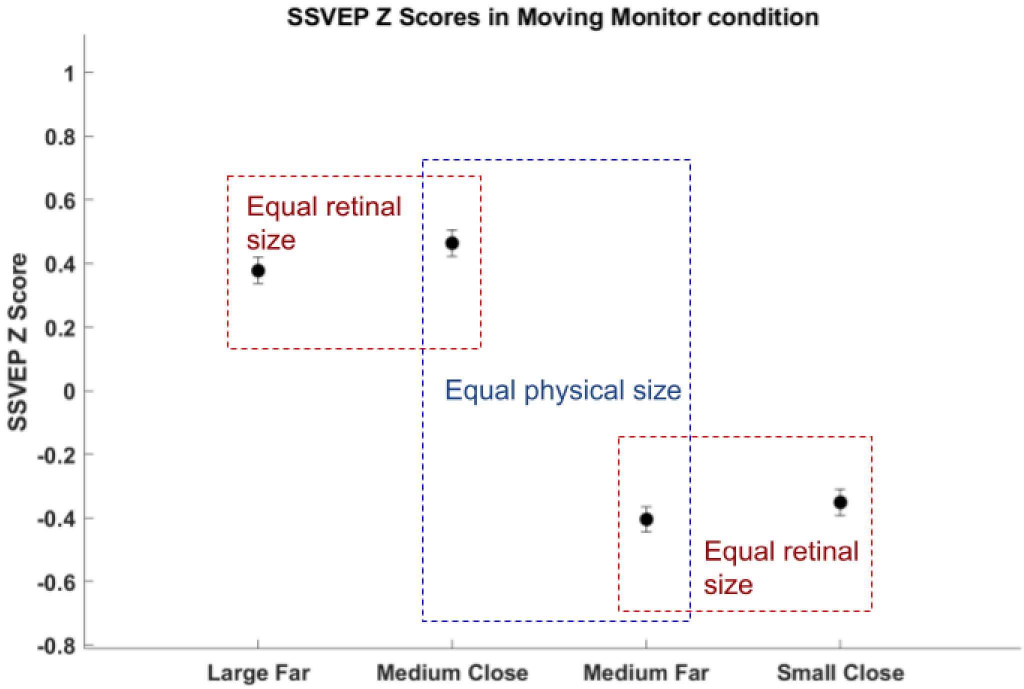


Figure 36: SSVEP strength for stimuli in the retinally constant condition in the moving monitor condition.

In the moving monitor condition the SSVEP size constancy index was -6.49, indicating a lack of size constancy in the neural representation of the stimuli (figure 36).

Behavioural data

To study performance on the size-matching task, error scores were calculated by subtracting the size of the stimulus from the size of the recreated square. These error scores indicate potential over- or underestimations of the stimulus, with positive values indicating stimuli being recreated as larger, and negative scores indicating stimuli being recreated as smaller than the original stimuli. For the moving monitor condition, performing t-tests and correcting for multiple comparisons by applying the FDR, only recreations of retinally constant stimuli moving away differ (marginally) significantly in size from the original stimuli ($t(23)=-3.06$, $p=0.0056$). Further, a repeated-measures ANOVA was conducted, showing a main effect of motion direction, $F(1,23)=4.91$, $p=0.037$), but no main effect of type of constancy (retinal vs physical) ($F(1,23)=0.83$, $p=0.37$) and no interaction effect ($F(1,23)=1.77$, $p=0.20$) (figure 37).

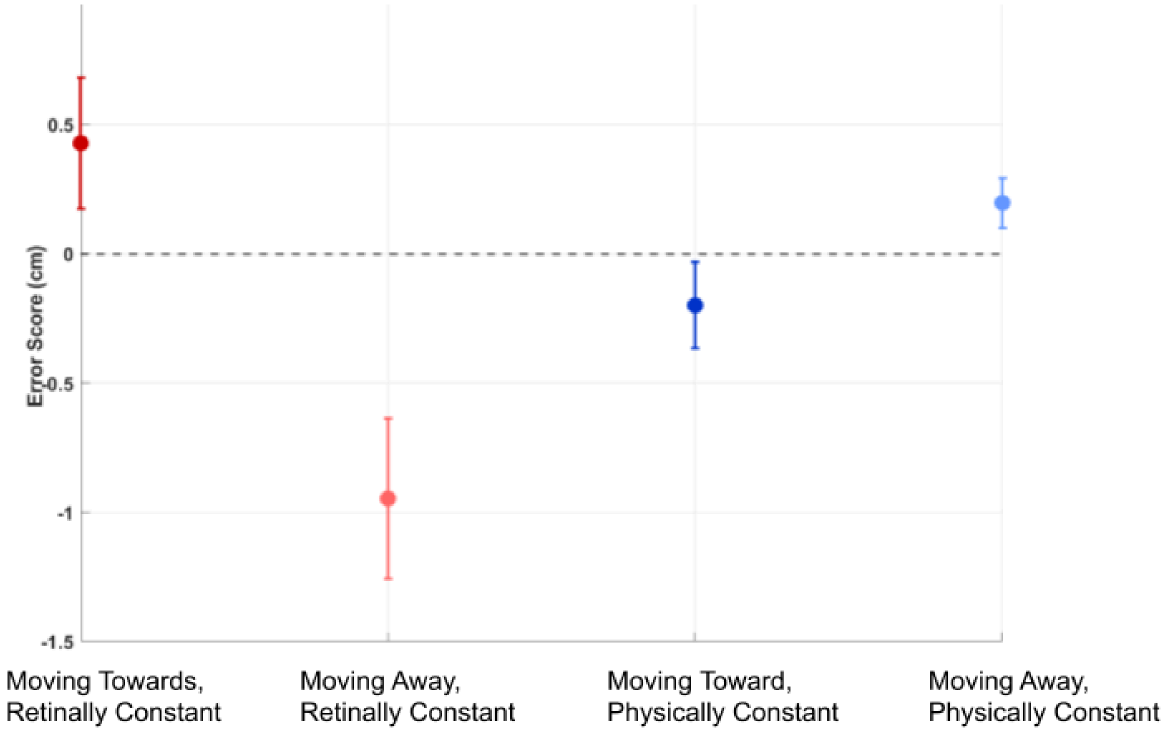


Figure 37: Behavioural error scores in the moving monitor condition. Positive values indicate that the response was larger than the original stimulus.

Similarly to the moving monitor condition error scores, when tested by t-tests and corrected by means of FDR, only differ significantly from 0 for retinally constant stimuli moving away from the observer ($t(21) = -2.78$, $p = 0.0111$). A repeated measures ANOVA revealed no significant effects (figure 38). Error scores are similar between the moving monitor and virtual reality conditions, following the same ordinal pattern and having similar magnitudes.

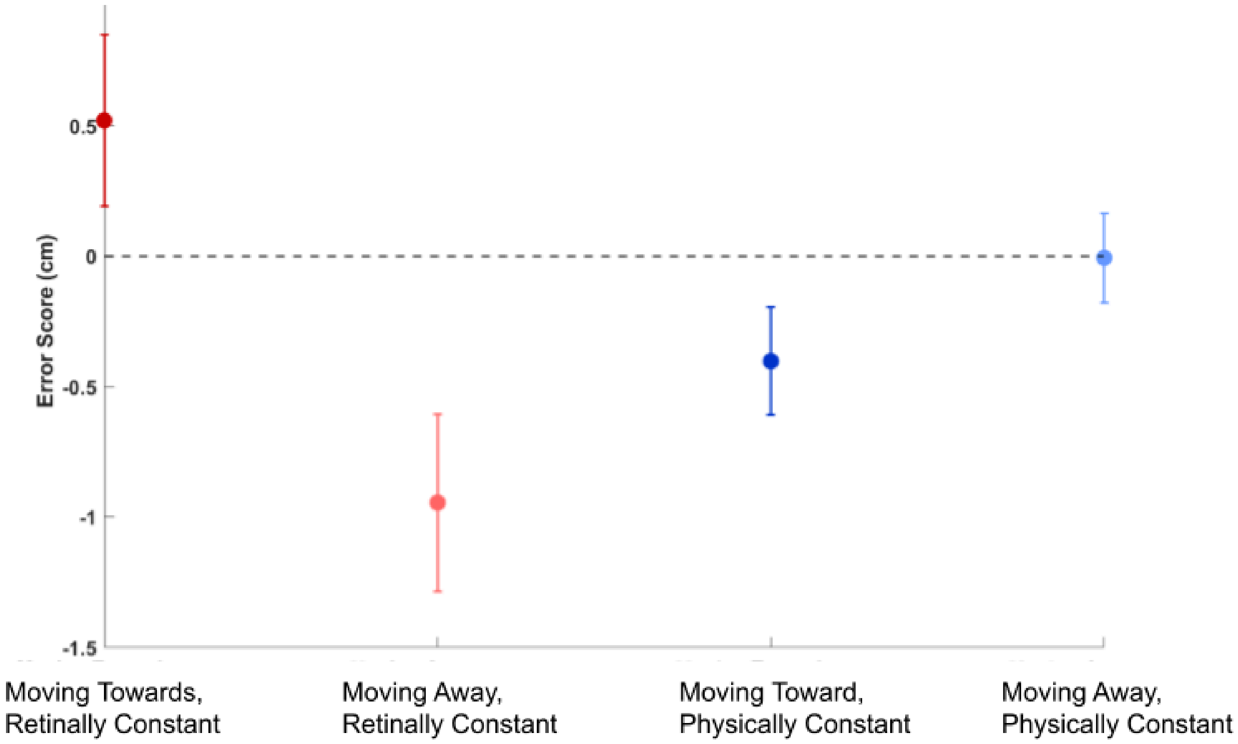


Figure 38: Behavioural error scores in the virtual reality condition. Positive values indicate that the response was larger than the original stimulus.

Control experiment

In the control experiment, stimuli presented in the moving monitor condition elicited SSVEPs with a size constancy index of 36.94, while in virtual reality, the size constancy index was 9.44 (figure 39). Both conditions showed partial size constancy, consistent with previous findings.

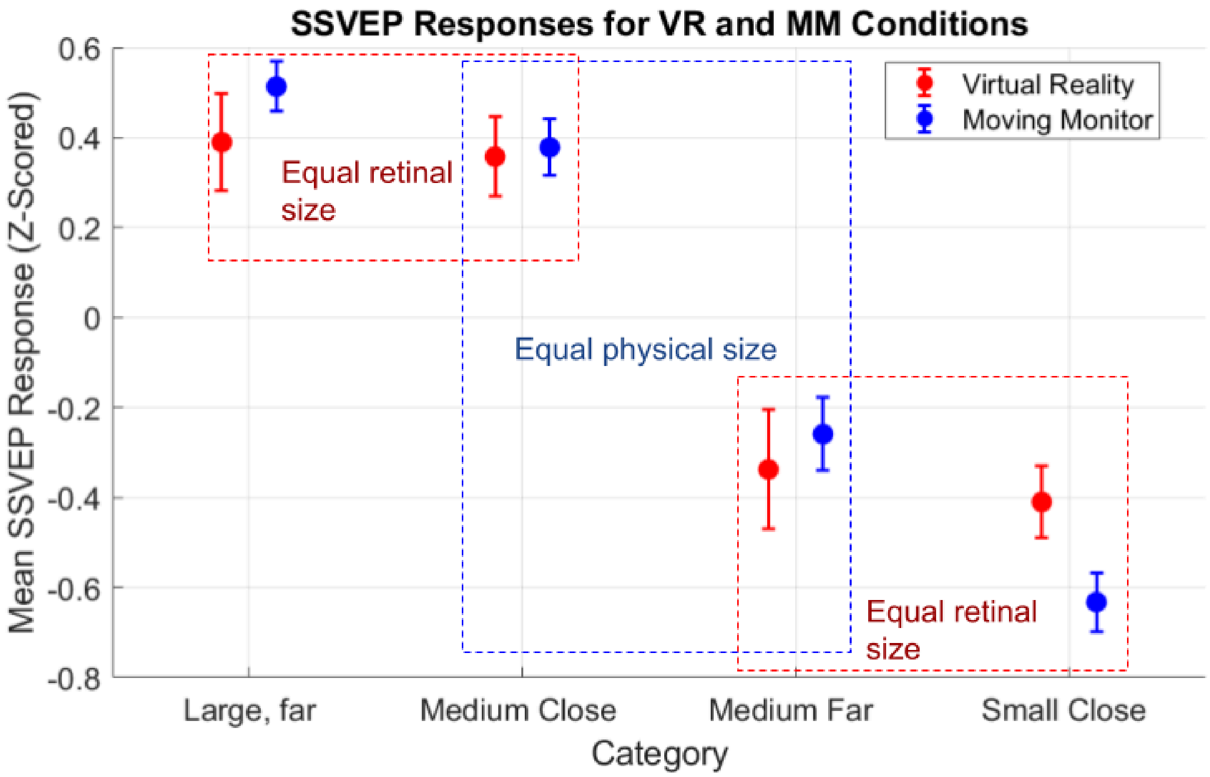


Figure 39: SSVEP strength for stationary stimuli in the virtual reality and moving monitor conditions.

The goal of this experiment was to learn more about the neural representation of size constancy in early visual cortex under dynamic circumstances. Using moving stimuli in an environment rich with depth cues, we attempted to study size constancy in more ecologically valid conditions.

Surprisingly, we found no encoding of physical object size in early visual areas of objects that had just moved. The SSVEP signal seemed to have encoded retinal size only, implying that V1 takes on a more retinotopic organisation when representing items in a visual environment where objects are moving and changing their size. It is important to contrast this with the findings of our control experiment, which does show partially size-constant representations of objects when viewed in static conditions, in accordance with previously published findings (Murray et al., 2006; Chen et al., 2019).

Our findings complicate current theories about the neural mechanisms of size perception. One theory concerning the adaption of neuron population receptive fields in relation to perceived size, which has been demonstrated in both non-human primates and humans (Ni, Murray, & Horwitz, 2014; he et al., 2015), does not rule out partially size constant representations when viewing objects in motion. Still, motion related size illusions have been shown to result in counter-intuitive activity in V1 (Whitney et al.,2003), and while illusory size changes are reflected in V1 activity, illusory motion is not (Ho & Schwarzkopf, 2022).

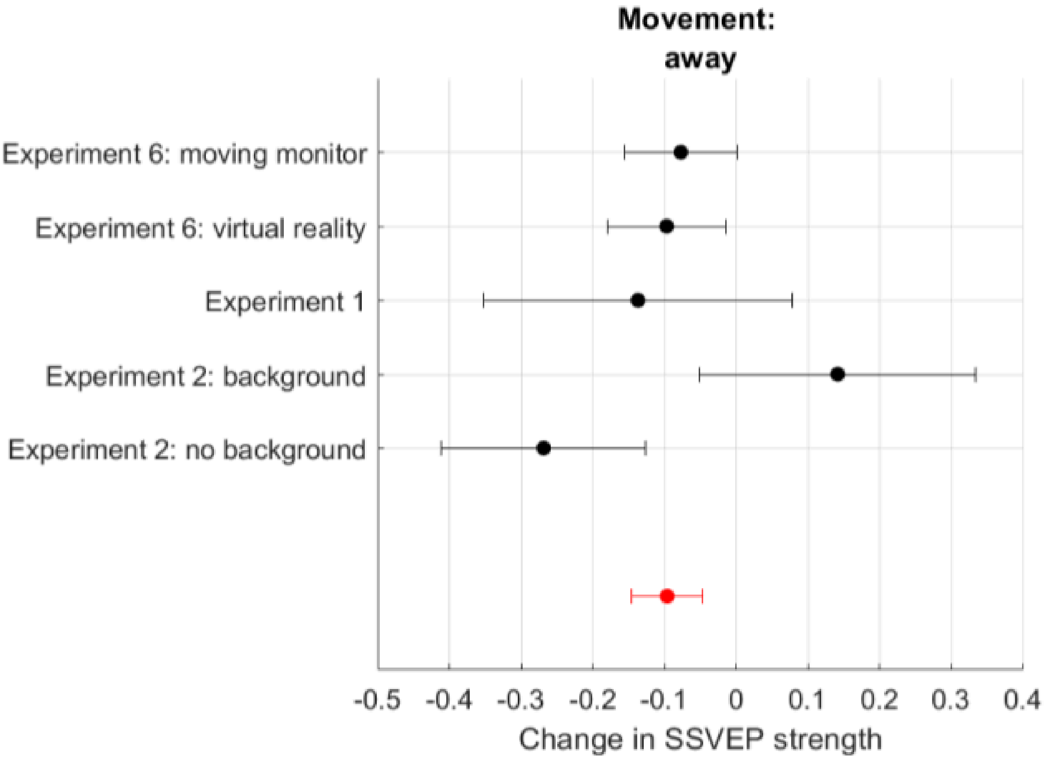
Although we cannot rule out the existence of partial size constancy in early visual areas in dynamic conditions based on these data, our findings do indicate a fundamental difference in how size is coded in early visual cortex when the object has just moved. Our findings essentially indicate a gap in our understanding of how this size coding works. Converging evidence from other neuroimaging techniques will be required to make any more conclusive statements about how objects are represented in V1.

Meta analysis

Over the many experiments that make up this thesis, retinally constant stimuli have consistently caused counterintuitive results. It was these counterintuitive findings that played a large role in the development of the hypothesis that stimulus movement and size change have unique contributions to their representations in early visual cortex. Despite these trends never reaching effect sizes that were statistically significant, the same trends kept appearing across experiments using different stimulus sizes and distances, different stimulation frequencies, visual environments and of course different participants. Each time, the trend in the data was in the opposite direction from the physical size change, with growing stimuli that moved away leading to slightly weaker SSVEPs, and shrinking stimuli moving towards the observer eliciting slightly stronger SSVEPs.

There are 3 different experiments that involved the use of retinally constant, moving stimuli. The first experiment, the second experiment (split over two conditions depending on the presence of a background behind the stimulus) and the sixth experiment comparing moving monitor and VR conditions (which consisted of two experimental sessions). This totals five data sets from three unique groups of participants, which can be used to perform a miniature meta analysis of this phenomenon. These five datasets contain a combined total of 2360 trials involving retinally constant, moving stimuli during which EEG data was collected.

Figure 40 displays two Forrest plots showing the change in SSVEP for retinally constant stimuli, split by direction of movement. Although stimuli that move towards the observer, overall, tend to elicit a slightly larger SSVEP after moving and shrinking, this effect is not significant (meta-effect = 0.0045, 95% *CI* [–0.089, 0.098], *SE* = 0.048, *Z* = 0.095, *p* = 0.924). Stimuli moving away from the observer, on the other hand, show a meta effect that is negative and on the cusp of significance (meta-effect = –0.096, 95% *CI* [–0.193, 0.0005], *SE* = 0.049, *Z* = –1.95, *p* = 0.051).



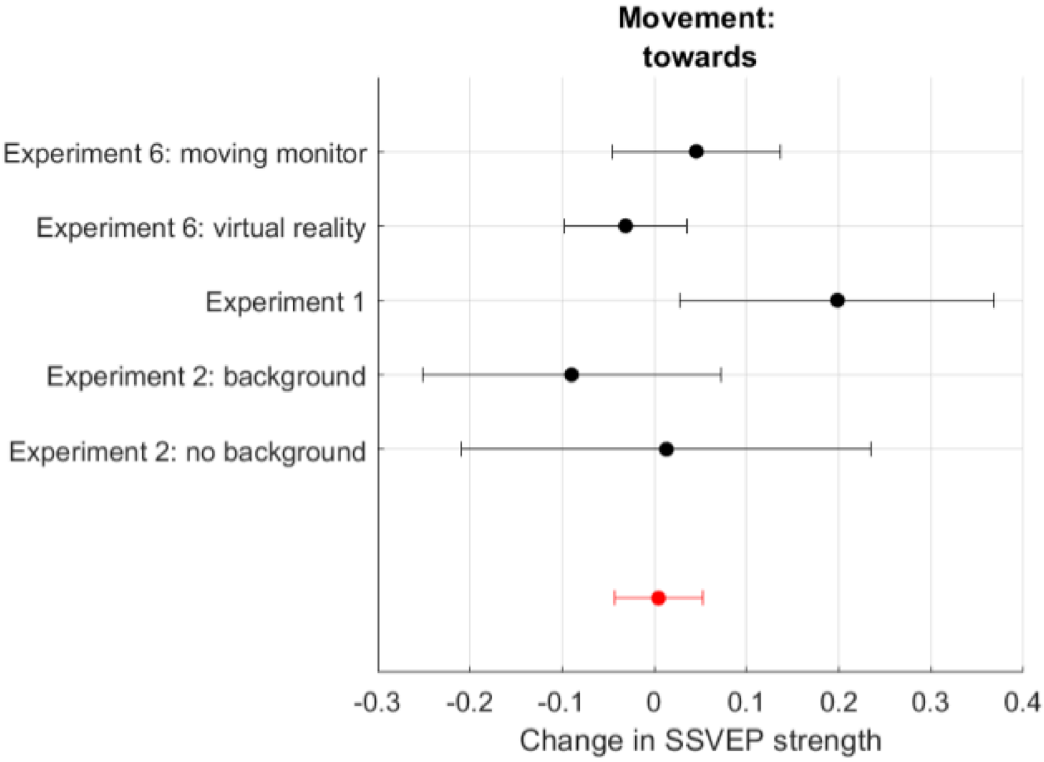


Figure 40: Forrest plots for the change in SSVEP strength elicited by retinally constant stimuli, split by direction of movement. Error bars are +/- 1 SE. The red data point signifies the summary effect estimate.

Much like the individual experiments that first yielded these data, the meta-analysis' results are ambiguous, and are not able to definitively demonstrate either the presence or absence of an effect. If we allow some statistical leeway, and we take the meta effect with a p value of .051 to be significant, this would mean there is an asymmetry in which visual information from sagittally moving stimuli is processed. This is not without precedent, as stimuli moving towards the observer have been demonstrated to elicit differential neural responses to those moving away (Hervais-Adelman, Legrand, Zhan, Tamietto, de Gelder & Pegna, 2015; Thieu, Ayzenberg, Lourenco & Kragel, 2024). Our findings add another layer of complexity, however, with concurrent changes of size.

In conclusion, there appears to be weak support for the notion that retinally constant stimuli moving away elicit weaker SSVEPs, despite physically growing. There is no support for a similar effect for stimuli moving towards the observer, which concurrently shrink as they move. This effect, if reflective of differential neural processing, is very small in its magnitude but could still be indicative of unique top-down predictive processes that emerge from expectations concerning the retinal sizes of moving objects.

Conclusions: on virtual reality in size and distance estimation tasks and neurophysiological experiments.

The vast majority of work described in this thesis has taken place in virtual reality. Although in the literature several important differences are described between real life size/distance perception and that in VR, we found VR to be an adequate, valuable research tool. Especially given our goal to investigate dynamic size constancy, the freedom afforded by virtual reality with regards to the creation of stimuli is unmatched by any physical device. Despite our eventual retreat to more traditional, physical devices, we consider virtual reality to be a promising technology that has much potential in advancing our knowledge about human size and distance perception.

Behavioural data

When it comes to evaluating VR as a research tool for investigating visual perception, it is Important to distinguish between two kinds of discrepancies typical to VR. One kind relates to differences between real life and VR, and is based on a direct comparison of real world based measurements and VR based measurements. An example of this is compression; where distances in VR are perceived as smaller than in real life. The other kind of discrepancy relates to how perceptual phenomena might differ from their real world counterparts, and rely on measurements taken exclusively within VR. An example of this would be the measuring of size constancy using exclusively virtual stimuli. This distinction is not always clearly delineated in the literature, and may lead to confusion when one experiment reports underestimation of sizes in VR (Hornsey, Hibbard & Scarfe, 2020) and another overestimation (Kim, 2018). These findings are not necessarily contradictory, because Hornsey et al. (2020) compare virtual stimuli to real world stimuli, and Kim (2018) compares virtual stimuli to other virtual stimuli. We will briefly go over the peculiarities of size and distance perception in VR as they have been described in the literature, and discuss our findings in this context.

Differences between virtual reality and real world perception

First, it bears investigating how VR differs from the real world. Perhaps the most prominent and well replicated finding involving VR and size/distance perception is the consistent underestimation of distances seen in VR as compared with the real world. Early work by Witmer and Sadowski (1998) first demonstrated that observers tend to underestimate distances in VR. When blindfolded, observers would only walk 85% of the distance to a previously shown goal in VR, compared to 92% of the distance in real life. Consequent research confirmed this principle, with distance judgements in VR, as measured by this walking method, being around 74% of the actual distance to target (Renner, Velichkovsky & Helmert, 2013). This phenomenon is not only apparent when observers walk to reach a target, with verbal reports of distance also undershooting the mark (Kelly, Doty, Ambourn & Cherep, 2022).

This underestimation of distance is tightly linked to size estimations. Since distances are underestimated, it is expected that object size is underestimated too (Loomis & Knapp, 2003). After all, the size-distance invariance hypothesis posits that an object of a given retinal size at a closer distance must be physically smaller. Virtual objects, when presented on a screen are indeed perceived as smaller than their real life counterparts (Stefanucci, Creem-Regehr, Thompson, Lessard & Geuss, 2015). Similarly, objects presented in VR tend to be perceived as slightly smaller too (Siegel, Kelly & Cherep, 2017; Hornsey, Hibbard & Scarfe, 2020; Itaguchi, 2021).

Our findings show no significant difference in size estimation between VR and real life viewing conditions. On the contrary, errors in the stimulus reproduction task of the virtual reality / moving monitor experiment are exceedingly similar between real life and virtual reality conditions (figure 41). The average difference between VR and real life error scores is less than 1mm (0.79mm). This demonstrates that when the virtual environment used contains adequate depth cues and is nearly identical to its real world counterpart, size estimations are also nearly identical.

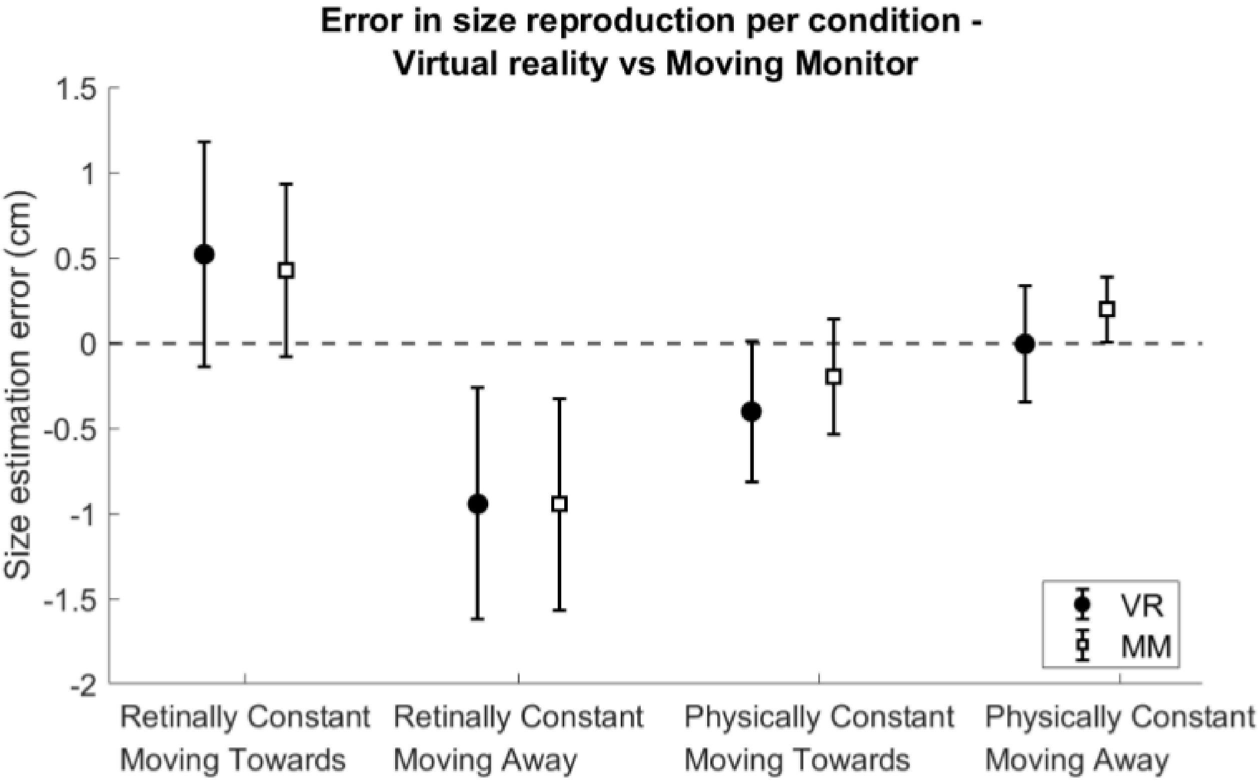


Figure 41: Error in size reproduction per stimulus condition in VR and real life viewing conditions. Error bars represent +/- 2 standard errors.

While we cannot make any direct claims about whether the distances and sizes in VR were perceived as smaller than those in the real world, since we have no conditions where the stimulus is perceived in VR, and reproduced real life or vice versa, the fact that the internal consistencies in VR also manifest in real life is excellent news for those who wish to use VR as a research tool in size and distance perception research. Even if there is an absolute difference in size and distance estimations between VR and real life, as long as observers show identical effects of identical magnitudes this would be irrelevant for our purposes of studying how human visual perception works.

Given the data as depicted in figure 41, we can conclude that retinally constant stimuli that move away from the observer, physically growing as they move, are slightly underestimated in size. The magnitude of this effect, in both VR and real life, is about 1 cm, which constitutes an underestimation of the final object size of ~5%. This effect is presumably a function of the combination of object movement and simultaneous object growth. It is entirely possible that the object in virtual reality, if it were to be reproduced in real life, would be reproduced at a smaller size, perhaps by a factor of 25% as indicated by Kelly (2023). However, if object movement and growth are the factors of interest, like they were in our case, this absolute difference is irrelevant. More important is the fact that we observe that the direction and magnitude of the effect is identical in VR and real life, allowing future experiments that only make use of VR to be generalizable to real life viewing conditions, given that the effects are expressed in relative units of size and distance, and not absolute ones.

Whereas some authors explicitly warn against transferring findings from VR to real life conditions (Rzepka, Hussey, Matlz, Babin, Wilcox & Culham, 2022), our findings seem to suggest this kind of transferral could be valid. What could be the reason for this lack of difference between VR and real life? For one, the degree of distance underestimation in VR has diminished with the use of more advanced, modern HMDs (Kelly, 2023). Nevertheless, the specific HMD used in our experiments (HTC Vive) has been shown to still elicit an underestimation of distance of ~18% (Kelly, Cherep & Siegel, 2017). This 18% should also be seen in the context of a size underestimation of 5% which occurs in real life viewing conditions, further narrowing the difference between VR and real life.

A second consideration is the quality of the virtual environment. It seems intuitive that sizes and distances would be easier to judge in more true to life representations of an environment. Our experiment comparing real world viewing conditions and VR made use of a photorealistic virtual copy of the physical experimental setup. However, research into the effects of photorealistic environments as compared with more simplified graphics suggests that improved visuals do not necessarily improve size and distance perception (Thompson, Willemsen, Gooch, Creem-Regehr, Loomis & Beall, 2004; Kunz, Wouters, Smith, Thompson & Creem-Regehr, 2009). Vaziri, Liu, Aseeri & Interrante (2017) report differences in distance perception when comparing natural viewing with viewing the same scene through a live-video feed, but no difference between this live feed and a simplified video feed with much reduced graphic fidelity. This implies that the graphical quality of the virtual environment has little, if any, effect on perceived size and distance.

Nevertheless, the photorealistic nature of our virtual environment might still have contributed to the lack of difference between our VR and real life conditions. Interrante, Ries & Anderson (2006) found that differences in distance estimation between VR and real life disappeared when participants were first made familiar with a certain room before being made to enter a photorealistic copy of this room in VR. The authors ascribe this accurate perception to a cognitive effect of familiarity with the space being occupied. Although we made no explicit efforts to familiarize our participants with the room or the apparatus, our participants were familiarized with the environment by simply being present in it during the pre experimental briefing and EEG preparation. It is therefore plausible that, similarly to participants in Interrante et al.’s (2006) experiment, our participants observed the virtual environment as being functionally identical to its real world counterpart, which led to size judgments that were identical to those observed in the real world. If this explains our findings, this might take away from the usefulness of VR as a research tool. After all, if a real world counterpart to any virtual environment is required for accurate size and distance perception, there is little added benefit to performing any experiments in VR.

In conclusion, our findings that size judgements are identical in VR and real world viewing conditions are likely due to a combination of modern hardware and a high fidelity virtual environment that mimicked the physical experimental setup. Although promising at face value, more research is required to confirm which of these factors contributes most to the observed results, and to what extent these findings will transfer to novel virtual environments that lack a real world counterpart. More importantly, the relative effects of movement and size change manifested themselves in the same way in VR and real life. This implies that future experiments investigating these phenomena making use of solely VR may indeed have their findings generalized to normal everyday vision. Whether this also applies to other visual phenomena remains to be tested, but our findings should inspire optimism.

Size constancy in VR

Aside from the above described differences in judgements of absolute distance and absolute size between VR and real life viewing conditions, there appear to be discrepancies in size- and distance scaling that are unique to VR. Hornsey et al. (2020) find that size constancy in VR is imperfect; when replicating the size of an object at various distances, observers create larger objects at further distances. This increase in size of the replicated item implies underconstancy; and in the context of this experiment where participants were physically holding the item they were replicating, this is ascribed to an underestimation of distance to the stimulus. Kim and Yoo (2024) performed a size constancy experiment where both the target and the replicated stimulus were presented in VR. Contrary to Hornsey et al. (2020), they found that the replicated stimuli were smaller than the target stimulus. Given that observers accurately reported the distances to each stimulus, this implies an overconstancy in VR that cannot be explained by errors in distance perception.

Contrary to both these experiments, our data shows that size constancy can be properly maintained in VR. Behavioural results in VR from experiments 1, 2 and 6 show accurate replication of the physical size of moving stimuli, demonstrating properly functioning size constancy. Note that even though we did not calculate size constancy indices which are calculated as *recreated item size/stimulus object size*, if the replication size is statistically indistinguishable from the stimulus size, this implies size constant perception. This functional size constancy may be partially explained, again, by the nature of the virtual environments used in our experiments. All virtual environments we used were designed to have a high degree of depth cue salience. Previous research using a Cave automatic virtual environment (CAVE) found that virtual environments with rich depth cues facilitate size constant perception more than environments with poor depth cues (Luo, Kenyon, Kamper, Sandin & DeFanti, 2009; Murgia & Sharkey, 2009). Luo et al. (2009) also report near perfect size constancy in environments with rich depth cues.

One factor that may help explain the difference between our findings and those of Kim and Yoo (2024) is the distance at which stimuli were presented. Kim and Joo (2024) report a trend towards the breakdown of size constancy that starts when stimuli are presented at 6 meters distance from the observer, with significant non-constancy being reported from distances of 8 meters and above. In our experiments, as well as other in the literature that report functional size constancy (e.g. Luo et al, 2009) stimuli are presented at closer distances, with our maximum presentation distance being 3,5 meters from the observer. It is possible that even in depth cue rich virtual environments, size constancy starts to falter at greater distances. However, more research would be required in order to map out this phenomenon more clearly.

In conclusion, size constancy in VR, at least within the parameters of our experiments, appears to be perfectly functional. Despite differences between real and virtual world judgements reported by Hornsey et al. (2020), when observations and responses are both made within VR, observers seem to perceive things in this virtual world in a size constant manner. This is reassuring, because as long as the logic and mechanisms of size and distance perception as they occur in VR mimic those as they occur in real life, findings made within VR will be able to be generalized to normal, everyday perception.

Distance overestimations in VR

One recurring phenomenon we encountered during our VR experiments was the tendency for observers to overestimate the distance to the stimulus. In experiments 3, 4 and 5, participants were tasked with indicating the location at which the stimulus was last seen. Generally, participants indicated a distance beyond the actual location at which the stimulus was presented, and these overestimations were occasionally very large. This phenomenon is not indicative of a difference in distance estimation between VR and real life, as both stimulus presentation and distance estimation were performed in the same virtual space. Rather, it seems to be specific to the type of virtual environment used in these experiments. Overestimations of particularly large magnitudes, such as twice the actual distance or more, would occasionally occur. These seem to indicate a particular kind of illusion where the stimulus appears much further away. As seen in experiment 5, this illusion is particularly prone to occur for small stimuli that are presented far away, and as evidenced from experiment 4, it is more likely to occur when the stimulus is not anchored to its immediate surroundings. The effect of stimulus size is most likely due to the fact that smaller stimuli do not occlude as much of the hallway, providing less salient depth information. Similarly, when there is no physical attachment between the stimulus and the surrounding, it is sometimes hard to judge where along the hallway the stimulus is, giving rise to a perception of ambiguous distance.

The lack of texture and shading seen on the stimulus further contributes to the ambiguity of its distance. Texture, shading and especially their combination are useful cues to size, distance and shape perception (Blake, Bülthoff & Sheinberg, 1993; Kemp, Cesanek & Domini, 2023). Given that our main goal was to extract SSVEPs from our observers, stimuli could not have either a texture or shading applied to them, at least not in conditions where they moved. Due to the sensitivity of SSVEP amplitude to luminance and local contrast, shading and texture at various distances would have caused tremendous confounds, and have rendered our electrophysiological results moot. Some degree of depth cue ambiguity must therefore be taken for granted. In sum, the large overestimations of distance observed in experiments 3, 4 and 5 were likely due to specific stimulus properties, and this overestimation of distance can partially be remedied by anchoring the stimulus to the immediate surrounding.

Virtual Reality and Electroencephalography

The combination of virtual reality and electroencephalography (VR-EEG) has seen a rise in popularity with the prevalence of commercially available VR HMDs. Current VR-EEG applications are wide ranging, with clinical applications, brain-computer interfaces (BCIs) and, like described here, fundamental research. There is some methodological overlap between the experiments of this thesis and BCI research. BCIs often make use of SSVEPs, and recent work has aimed to design specific SSVEP stimuli for use in VR, mapping out what parameters for SSVEPs work best in 3D environments (Zhu, Yang, Ding, Wang, Gong & Fu, 2023; Liu et al., 2024). Although our own work has still made use of 2D stimuli, Zhu et al.'s research offers promising future avenues using 3D stimuli as SSVEP inducers in virtual environments.

With regards to size and distance perception, some unique aspects about neurophysiology in VR as compared with 2D depth perception have been identified. Tang et al. (2022) find that depth cues presented in immersive VR elicit different neural responses to those presented in 2D, with differences present in ERP amplitude (specifically P2) and low frequency (delta and theta band) activity. It should be noted that these effects are likely not unique to VR, but rather are related to how stereoscopic information is processed as opposed to 2D depth cues (Cepeda-Freyre, Garcia-Aguilar, Eguibar & Cortes, 2020; Guo et al., 2022). In any case, depth perception related EEG activity in VR being more similar to real world conditions than previously used 2D stimuli is a promising finding, and supports the notion that VR has the potential to combine the internal validity of lab-based experiments with the external validity of real world viewing conditions.

Issues with data quality

The foremost issue we encountered using VR-EEG in the experiments described in this thesis was related to data quality. Noisy SSVEP averages plagued our findings all throughout this project and caused us to steadily reduce experiment complexity in order to have more trials per unique condition in each experiment. To illustrate, experiment 5 was intended to be conceptually simple, with the main comparison of interest being between static stimuli and moving stimuli. Trials in this experiment were split by movement vs no movement, and within the movement condition, split across three sizes. This resulted in 15 trials per participant in the condition where a medium sized stimulus moved away from the observer. 22 participants were able to provide usable datasets which were used for consequent analysis. As far as

EEG experiments go, this is a respectable sample size. Indeed, the physically constant stimulus from this condition evoked a larger SSVEP when close to the observer, and a smaller SSVEP when far away, in line with expectations and in line with later findings in experiment 6. However, this difference was not significant. Based on the data we have, we can make tentative predictions of how many participants this experiment would have needed to find a significant effect. Assuming that the means found in these data are representative of population means, and using the standard deviation of these subject-level differences as an estimate of population variability, we can run a Monte Carlo simulation to estimate statistical power for any given sample size. Based on these calculations, we would require 166 participants for the difference in SSVEPs to be obtainable with 80% power. This is simply not a feasible number, and raises questions about the viability of trying to answer our research questions with these methods. Of course, the above calculation is a rough estimate based on the available data, but it does paint a picture of the variability present in our data, and the magnitude of the sample size needed to find significant results. Data quality is therefore a major limiting factor in VR-EEG research.

Solutions for data quality

There are several ways of improving data quality, which involve methods in data collection and data processing. Concerning data processing, clever methodologies may be used to extract more meaning from inherently noisy data. A recent, and very helpful improvement came from Zhang, Valsecchi, Gegenfurtner & Chen (2023), who demonstrated that using a Laplacian reference is able to increase SNRs when measuring SSVEPs as compared to other reference methods. All data presented in this thesis have, after the above paper was published, been re-referenced to use this Laplacian reference. Another recent example is rhythmic entrainment source separation (Cohen & Gulbinaite, 2017), which makes use of filters in order to more clearly delineate the stimulation frequency from surrounding frequencies. Such advances in signal processing techniques have the potential to increase data quality, and thereby increase the value of VR-EEG as a research method. As more advanced data processing techniques are developed, more nuanced research questions may be addressed using VR-EEG, leading to increased potential of this technique.

On the data collection side of the equation, advances are being made as well. One might assume that having an electrical device like a VR HMD mere millimetres away from EEG electrodes might negatively affect EEG data quality. Fortunately, noise introduced by VR HMDs is generally restricted to specific line noise at the frequency of the alternating current

(50Hz in our case, since all data was collected in Germany) and the frequency of the refresh rate (90Hz in our case, since we used the HTC Vive). Studies have found that, generally speaking, electrical interference from VR HMDs does not significantly affect EEG data quality (Weber et al., 2021; Cavallo, Brubaker, Bruckner & Castro, 2023). The more pressing issue is the physical interference between the HMD and the EEG electrodes. Many commercially available HMDs make use of headstraps or headbands, which hold the display up to the eyes of the observer. These bands usually wrap around the top and sides of the head, and rest on the electrodes of the EEG cap. Given that EEG electrodes are exceedingly sensitive to mechanical perturbation, this is a major source of interference.

To prevent this, several solutions have been proposed. Integration of electrodes into the HMD is one of these, with complexity of this integration ranging from simply sticking electrodes to the inside of an HTC Vive face mask (Zhang et al., 2021) to specifically designed dry sponge electrodes integrated into frontal, central and occipital aspects of a VR HMD headband (Li et al., 2023). There are also purpose built, commercially available solutions that replace the standard HMD headband entirely with an electrode-containing headband and which appear to record EEG data of acceptable quality (Lee, Kim & Ahn, 2024). The advantage of these systems is that it reduces the amount of mechanical forces enacted upon the electrodes. The shearing forces between the top of the electrodes worn in a standard EEG cap and the overlaying headstraps is eliminated in these cases. Of course this does not eliminate the movement related noise, but it likely reduces it significantly.

An alternative, and less resource intensive solution to the problem of mechanical interference in VR-EEG is the use of specialized headstraps. Tauscher, Schottky, Grogorick, Bittner, Mustafa & Magnor (2019) modified the headstrap of an HTC Vive, cutting away material to make room for underlying electrodes of a standard EEG electrode cap. This simple but effective solution is a worthwhile alternative for those not willing to invest in fully VR integrated EEG solutions.

Looking ahead, the incorporation of (self-initiated) movement in experiments into size- and distance perception seems to be a promising avenue. Size constancy has been demonstrated to benefit from self initiated movement as compared with object movement (Combe & Wexler, 2010), and self initiated movement has been demonstrated to elicit unique activation in V1 in mice, even in absence of any visual input (Keller, Bonhoeffer & Hübener, 2012; Saleem, Ayaz, Jeffery, Harris & Carandini, 2013). Incorporating self initiated movement in a VR based experiment on the neural correlates of size constancy could therefore reveal very interesting results. Of course, allowing for free movement of the observer would increase noise, and data might suffer from the same low SNRs reported in our experiments. However, given the above mentioned remedies for this kind of noise and potential future solutions for enhanced data analysis techniques, an experiment like this might be in the realm of possibility (Giangrande, Botter, Piitulaenen & Cerone, 2024).

Overall, there are reasons to be optimistic about the use of VR-EEG techniques. Although we experienced significant issues with data quality, and future avenues of research that would incorporate more naturalistic behaviors from observers are likely to introduce even more noise, much work is being done to remedy these issues. With advancements in reduction of noise at the recording level, data analysis techniques that enhance SNR and techniques to isolate motion specific artefacts, the usefulness of VR-EEG for studying neural activity in naturalistic conditions is ever increasing.

Conclusions: on the neuroscience of size perception

The main, overarching goal of the experiments was to learn more about the neuroscience of size perception, with findings about behavioural responses and implications about the use of virtual reality being of secondary importance. Indeed, we can glean several things about size perception in early visual cortex from our findings, although these should all be interpreted within the context of the methodology used. The SSVEP is a complicated phenomenon, and even though we use its magnitude as a more or less direct metric for object size in the experiments described here, many other factors affect it. We will start with a recapitulation of the various determinants of SSVEP amplitude, paying particular attention to the ones most relevant to our work, and expounding on how our experiments incorporated them.

The many determinants of SSVEP amplitude

Contrast is one of the most important determinants of SSVEP amplitude. Higher contrast generally results in higher SSVEP amplitude, with the SSVEP amplitude being proportional to the logarithm of the contrast (Campbell & Kulikowski, 1972). There are limits to this effect however, with saturation occurring at higher contrasts where the SSVEP no longer significantly increases (de Lange, 1958; Spekreijse, 1966). This saturation effect is further dependent on the temporal and spatial frequencies used, with saturation being reached sooner for higher temporal and spatial frequencies (Zemon & Ratliff, 1984). Saturation is often reported to occur at contrasts of anywhere between 10% and 50%, depending on spatial and temporal frequencies. Whereas our initial, preliminary experiments made use of high contrast stimuli with contrasts of 100%, we soon reduced this to contrasts of ~10% to avoid saturation effects and to increase observer comfort and reduce fatigue (Chang, Baek, Lee & Park, 2014; Ladouce, Darmet, Torre Tresols, Velut, Ferraro & Dehais, 2022). With this lower degree of contrast, stimuli in our main experiments were unlikely to reach a saturation point. This allowed us to interpret any lack of increase in SSVEP amplitude as being reflective of size perception in early visual areas, as opposed to an effect of saturation. This supports its use as a metric for the neural correlates of size perception.

Overall luminance is another, related, determinant of SSVEP strength. A lower background luminance has been shown to increase SSVEP amplitude (Spekreijse, 1966; Zhang & Chen, 2022). Stimuli using the same contrast are therefore capable of eliciting a stronger SSVEP, given lower background luminance. Zhang & Chen’s (2022) paper does not fully disambiguate between luminance and contrast, however, since the contrast against the higher luminance background was necessarily lower than that against the low luminance background. Recently, Hedjar, Martinovic, Andersen & Shapiro (2025) attempted to differentially manipulate contrast and luminance, finding that SSVEPs generally correspond more to contrast, although at higher frequencies (>5Hz) and higher contrasts (>15%), this effect becomes less clear and SSVEPs neither cleanly responding to contrast or luminance modulation. Although preliminary experiments as well as experiment 1 of this thesis made use of a large black background behind the flickering stimulus, reducing overall luminance, this was removed in experiment 2. Like one could predict from the literature, this did decrease overall SSVEP amplitude. However, the SSVEP was still strong enough to be distinguishable, and more importantly, it did not seem to reduce the magnitude of the differences between large and small stimuli. We therefore deemed the removal of the background to be a reasonable choice, enhancing the perceived naturalness of the virtual environment while retaining the SSVEP’s usefulness as a metric for the neural correlates of size perception.

The removal of this black background did emphasize the potential effect of another SSVEP amplitude determinant, namely surround suppression. Surround suppression is a neurophysiological phenomenon where neural responses to a stimulus are suppressed by the presence of surrounding stimuli, which are not in the original neuron’s receptive field (Allman, Miezin & McGuinness, 1985). High contrast surrounds have the potential to reduce SSVEP amplitude (Vanegas, Blangero & Kelly, 2015). Suppression effects are generally stronger when the surrounding visual field matches the stimulus’ spatial orientation, and is stronger in the periphery than in foveal vision (Baker, Vilidiaite & Wade, 2021). The removal of the stimulus background increased surround contrast, and the eccentricity of this high-contrast surround would vary with the size and distance of the stimulus. It is currently unclear what the exact relation is between degree of contrast, its eccentricity and the degree of suppression on SSVEPs induced by a centrally attended stimulus, a comprehensive study systematically varying these conditions has not yet been performed. It is therefore difficult to predict what, if any, effect the surrounding visual environment had on the SSVEPs reported in our experiments. Again, we must refer back to the findings of experiment 2, which showed the continued variability of SSVEP magnitude with stimulus size, and rely on this as support that despite possible effects of surround suppression, the SSVEP retained its purpose as a metric for the neural correlates of size perception.

The stimulation frequency itself is a crucial factor in SSVEP strength. Early work on SSVEPs focussed on 10 Hz as a stimulation frequency, inspired by the fact that endogenous activity at this frequency is high in resting subjects. This is what we now know as alpha activity, and was first described as the Berger rhythm (Berger, 1929; Adrian & Matthews, 1934). The current consensus on SSVEP frequencies is that stimulation frequencies in the alpha band (8-12 Hz) do produce the largest SSVEPs (Regan, 1966; Norcia, Appelbaum, Ales, Cottereau & Rossion, 2015). Although this is not an absolute rule, and what frequency evokes the strongest SSVEP is affected by stimulus properties and recording location (Srinivasan, Bibi & Nunez, 2006). Overall, SSVEP amplitude tends to decrease drastically at stimulation frequencies above 20 Hz, and are relatively low at lower frequencies of < 5 Hz (Spekreijse, 1966; Herrmann, 2001). Duart, Quiles, Suay, Chio, Garcia & Morant (2020) specifically compared SSVEP responses at low (5 Hz), middle (12 Hz) and high (30 Hz) frequencies, finding the best SNR to occur at 12 Hz. Our experiments made use of stimulation frequencies between 5 and 15 Hz based on the above findings. Of course, this frequency band corresponds broadly to the alpha frequency band (8-12 Hz), which is tied closely to attentional processes, and can vary based on observer attentiveness and task demands (Klimesch, 2012). To prevent any variation in alpha band activity from being confused with size-related SSVEP effects, we subtracted the average of the 8 nearest frequency bins from the stimulation frequency during data processing. This removes any variation in alpha activity from the SSVEP signal.

Fatigue and cortical / retinal adaptation are further factors influencing SSVEP amplitude. Over the course of SSVEP experiments subjects may get fatigued, decreasing the SSVEP amplitude and SNR (Cao, Wan, Wong, da Cruz & Hu, 2014). Certain stimulation frequencies and contrasts cause more fatigue than others, though there is often a tradeoff between stimuli that offer better SNRs and those that are more comfortable to observe (Chang et al., 2014; Tian et al., 2024). To avoid a systematic confound of fatigue or accumulation of general discomfort, in experiment 6 the order of presentation of the experiments (the VR version and the moving monitor version) was randomized. In our previous experiments either trials or blocks were randomized.

Aside from fatigue effects which tend to affect SSVEP amplitude over a longer time course, recent work has shown more short term fluctuations in SSVEP strength. Labecki, Nowicka, Wrober & Suffczynski (2024) demonstrated that SSVEP amplitude varies significantly over the course of 60 seconds, with amplitude variance depending on stimulation frequency. Certain, higher stimulation frequencies, like 15 and 20 Hz display a sharp rise in power which then exponentially declines over time, whereas a stimulation frequency of 5 Hz results in a steady increase in power. It therefore seems like the SSVEP is not as steady as the name may imply, with the differential effects of stimulation frequency indicating the separate contributions of the parvo- and magnocellular pathways to the SSVEP. Theirs is not the only paper that investigated SSVEP time courses, as in recent years several experiments examined just this phenomenon, concurrent with our own preliminary experiments on the time course of SSVEPs (Prado-Gutierrez et al., 2019; Labecki, Nowicka & Suffczynski, 2019; Zhang, Valsecchi, Gegenfurtner & Chen, 2023). Zhang et al. (2023) reports the longest time course yet, tracing SSVEP strength over the course of 150 seconds. They find that the SSVEP, after an initial peak and exponential decrease, like reported by other authors, does stabilize after ~45 seconds, and remains steady from there on.

In our own experiments, we also observe similar fluctuations of SSVEP amplitude over time. Figure 42 displays the data from the control experiment of experiment 6. The stimuli in this experiment were stationary, and were presented for 11 seconds at a time. Any fluctuation in SSVEP amplitude is therefore not due to the stimulus moving or changing size. Similarly to the findings described above, we observe an increase in SSVEP amplitude, followed by a decrease. The closest simile to published work is to that of Zhang et al. (2023), who used the same stimulation frequency at the data in figure 42; 7.5 Hz (though Zhang et al. (2019) use chromatic stimuli compared with our achromatic ones).

This does not mean, however, that we need to correct our previous data where we compare data segments from the first four seconds before movement, and the last 4 seconds after movement for this kind of trend. Note that the data as presented in figure 42 were created using a moving window analysis (for more on this, see appendix A “Time-Frequency Optimization for rapidly changing SSVEPs”). The previously cited papers with the exception of Labecki et al. (2019) who used a band pass filter to isolate the stimulation frequency, also use this kind of analysis. Given the short windows from which the constituent frequencies are extracted, it is not possible to correct these data for the amplitude of endogenous EEG activity at the stimulation frequency, and the resultant data contain the sum of the contribution of the stimulus and whatever endogenous brain activity occurs at this same frequency. Despite not using a moving window analysis, the methods of Labecki et al. (2019) also cause the resultant data to include the contribution of the flickering stimulus as well as endogenous brain activity at this frequency. Given that we know that EEG frequency band power changes with sustained attention (especially alpha and theta, in which SSVEP stimulation frequencies regularly fall), these changes might go some way to explaining the shifts in SSVEP amplitude (Ray & Cole, 1985; Benwell et al., 2019). Indeed, if we take the same data as presented in figure 42, and instead plot the average SSVEP amplitude from the first and last 3.5 seconds of each trial, after subtracting the power of the nearest 8 frequency bins (+/- 1.1 Hz) to the stimulation frequency, no difference is observed over time (figure 43). Comparisons between SSVEP amplitude during the first and last 3.5 seconds are not significant for all categories. If there is a difference in absolute SSVEP amplitude during the segments that we use for data analysis, it is not statistically significant, and relative differences between the differently sized stimuli are retained across the span of each trial.

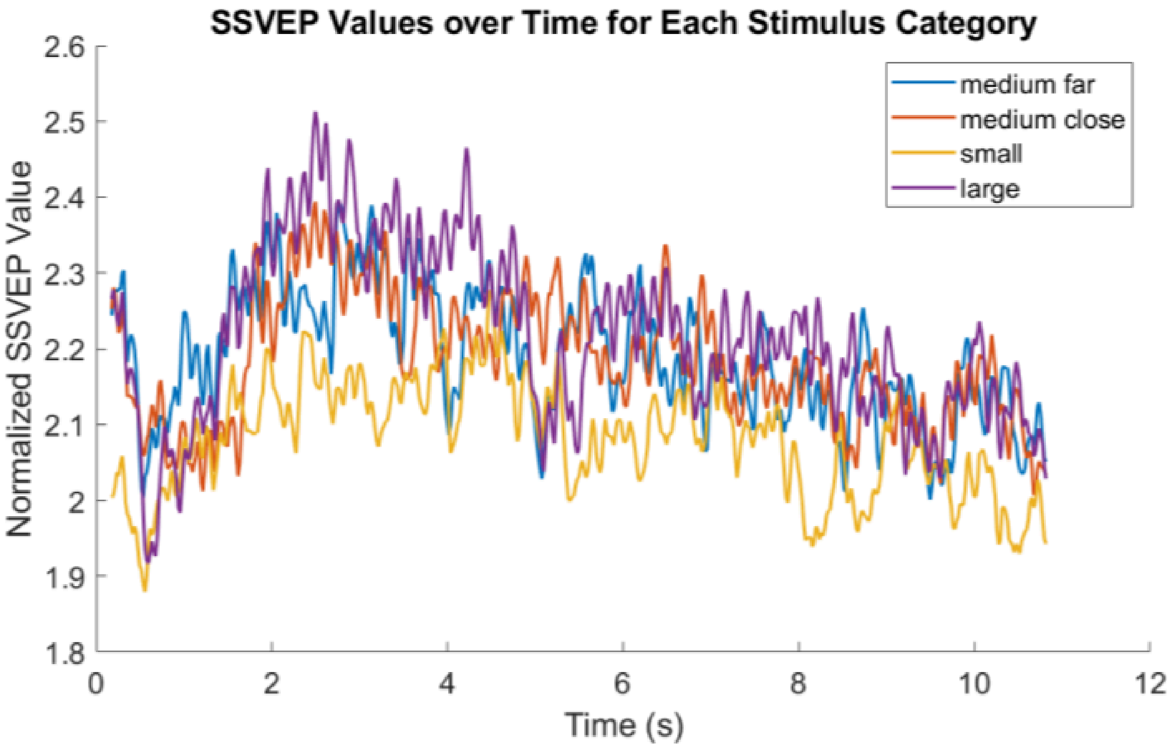


Figure 42: Normalized power of stimulation frequency and first harmonic over time, using bins widths of 350ms and step sizes of 20 ms.

Even though there are systematic short term fluctuations in SSVEP amplitude that might interfere with its use as a marker for size perception, data processing techniques are capable of removing these fluctuations. By removing the amplitude of the nearest surrounding frequencies, the endogenous EEG activity in the frequency band nearest the stimulation frequency is removed from the signal, leaving a cleaner indication of the contribution of the stimulus to the measured brain activity.

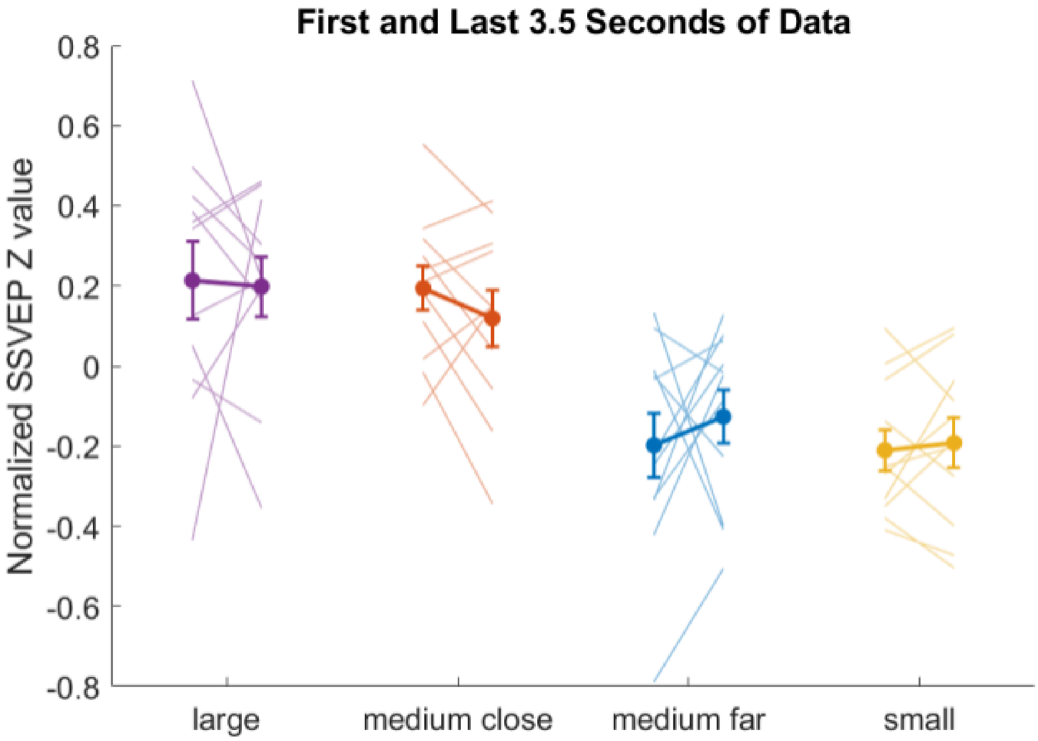


Figure 43: Comparison of SSVEP amplitude during the first (left) 3.5 seconds of data and last (right) 3.5 seconds of data, split per category. Thick lines indicate group averages, thin lines subject averages.

The properties of the waveform of the stimulation is another relevant variable. There are many types of waveforms, though sinusoidal and square waves are most commonly used in SSVEP applications. Though both square waves and sinusoid waves have been used since the earliest SSVEP experiments, and both types are still used today, recent work suggests that square waves are capable of generating higher SNRs (Teng et al., 2011; Chen, Wang, Zhang, Xu & Gao, 2019; Panitz, Gundlach, Boylan, Keil & Muller, 2023). Another related property is the duty cycle, which concerns the proportion of each on-and-off cycle that the stimulus spends in its ‘on’ state. Recent work on this found that a 50% duty cycle, in which the stimulus spends equally long periods of time in its bright and dark states, to be optimal for SSVEP generation (Chen et al, 2019; Shishavan et al., 2024). In all our experiments, we have made use of square wave stimuli with a 50% duty cycle, or as close to 50% as possible given the framerate, which appears to be the optimal waveform for our purposes.

A final determinant of SSVEP amplitude of note is attention. Attention has been shown to increase SSVEP amplitude (Morgan, Hansen & Hilyard, 1996). SSVEP enhancement has been shown in relation to feature based attention, spatial attention and even object based attention (Andersen, Fuchs & Muller, 2011; Adamian, Anderson & Hilyard, 2020). This reliable variance of SSVEP amplitude with attention has made the SSVEP a valuable tool for studying attention, though it means a potential source of noise for our experiments. It should also be noted that there is an intrinsic connection between size perception and attention, with exogenous and endogenous attention affecting size perception in different ways (Han, Tan, Zhuang & Qian, 2022). SSVEPs might be affected by attention via some unique mechanism, or attentional effects might be linked to the size perception of that object in our experiments. Regardless, aside from asking our participants to attend to the stimuli and inferring from their behavioural data that they did, generally, pay attention to the stimuli when it was presented, we have little means of establishing that their attention was maintained over the course of the experiment. We can only hope that variations in levels of attention fluctuated randomly across trials and categories, so that there are at least no systematic attentional biases in our data. Similarly to fatigue effects, our random presentation order of stimuli presumably helped to mitigate any systematic effects of attention in our data, and we may assume that the recorded SSVEPs reflect size perception in early visual cortex first and foremost.

Because of the large number of relevant variables that might affect the SSVEP amplitude, it is important to carefully consider the stimulus parameters before drawing any conclusions about underlying effects related to the subject being studied. It is not unrealistic to think that experiments might fail to replicate older findings due to a mismatch in one of the above listed factors. Indeed, a pre-registered study by Adam, Chang, Rangan & Serences (2021) failed to find evidence of feature-based attention SSVEP enhancement, although this is a well established phenomenon. The authors conclude that, among other things, the specific stimulation frequencies used might have contributed to the failure to replicate. Adam et al. (2021), by virtue of being a pre-registered study, allows us to learn about the potential finickiness of the SSVEP. There are likely many more examples of failures to replicate SSVEP findings which, due to factors like publication bias, have never made it into the body of published literature.

The above summation of the many variables that affect SSVEP amplitude, and the corresponding explanation of how they were controlled for in our experiments hopefully lends a good deal of credence to our assumption that the SSVEP, as we recorded it, forms a reliable indicator of size perception in early visual cortex. We cannot exclude the possibility that certain variables like attention and fatigue affected our data to some degree, nor can we exclude the possibility of some as of yet unknown interaction effect between the determinants described above affecting our results. Nevertheless, having controlled these variables to the best of our knowledge, which in turn is based on the most recent findings from the published research on this topic, we shall now discuss the neurophysiological findings and their implications about the neuroscience of size perception.

The neural mechanisms of size perception in dynamic conditions

Overall, we have demonstrated the use of SSVEPs as a general biomarker for size perception when viewing dynamic stimuli. Experiment 3 showed general agreeance between behavioural psychophysical measures of size perception and corresponding SSVEP strengths. Psychometric curves fitted to these data did indicate that electrophysiological measures were a lot noisier than behavioural measures, showing lower sensitivity. Still, it demonstrated the added value of SSVEPs as measures of size perception, which may be recorded in conjunction with the use of VR. Future experiments further exploring size and distance perception may therefore rely on SSVEPs as an implicit metric of size perception without explicitly querying observers.

The main goal of the experiments described in this thesis, however, was to more specifically delineate between the roles of retinal and physical size in object representations in early visual cortex. That the early visual cortex (partially) codes objects in a size constant manner has been long established. This phenomenon has been demonstrated using multiple methods, including fMRI (Murray et al., 2006; Fang et al. 2008), event related potentials in EEG (Chen et al., 2019; Yeh et al., 2024), SSVEPs (Chen et al., 2019) and TMS (Zeng et al., 2020). We therefore started with the assumption that for moving stimuli, some degree of size constancy was a given. Partial size constancy for moving stimuli would imply that stimuli that maintained their physical size as they moved would result in SSVEPs that were similarly more stable. For stimuli that retained their retinal size as they moved, this would mean some variance of the SSVEP with their physical size. Results from our experiments soon challenged these assumptions.

The most unexpected trend in the data, which was observed to various degrees in experiments 1, 2 and 6, was the increase of SSVEP strength for shrinking stimuli, and the decrease of SSVEP strength for growing stimuli. This trend (it is important to not classify it as an effect, as analyses on individual experiments never reached statistical significance) was repeatedly observed across multiple experiments. Aggregating all available data and performing a meta analysis, results revealed that this trend was not significant by the smallest of margins (p=.051) in one direction (movement away from observer), and not significant the other (movement towards observer). Still, the suspicion that this effect might have been real was a large factor in the design of experiments 4 and 5. The phenomenon could have been indicative of a particularly strong predictive effect of higher level areas on V1 object representation, which overpowers the role of retinal size representation. It is probably wise not to speculate too much on statistically insignificant phenomena, though, and we must conclude that we lack sufficient evidence to make any definitive statements about the presence or absence of this phenomenon.

The most substantial conclusions may be drawn based on the data of experiment 6, consisting of its 4 experiments both in VR and real life, and making use of dynamic and static stimuli. The results indicate that there is no size constancy in early visual areas when viewing objects that have just moved, while replicating the finding that there is partial size constancy in early visual areas when watching static stimuli. Does this mean that object representations in V1 cease to be size constant in more dynamic situations, or are our findings indicative of the limitations of the SSVEP technique? This finding deserves a more nuanced and fine grained analysis.

Recent findings have highlighted the importance of different layers of V1 in feedforward and feedback processing. Whereas layer 4 is mainly associated with feedforward processing, passing visual information along from the LNG to higher order visual areas, deep and shallow layers of V1 are more associated with feedback processing, receiving predictions or prediction error information. A hindrance to the interpretation of our data comes from the lack of our understanding of the specific source of the SSVEP with relations to the layers of V1. Although the SSVEP is generally assumed to reflect more bottom-up, feedforward processing on a global scale (Gordon et al., 2019), it remains unknown how this relates to the feedforward/feedback processing locally within V1. If, for instance, SSVEPs are mainly generated and contained within layer 4 of V1, and feedback processes affecting the size constant representation of a given object act in and via the deep and shallow layers of V1, our findings simply illustrate a small part of the equation. It is entirely possible that object representations in V1 are partially size constant even in dynamic situations, but the SSVEP method simply fails to capture this. This notion can be related to empirical findings somewhat, insofar as the degree of size constancy found in neurophysiological data from fMRI is generally larger than that found in SSVEP results. This would correspond to fMRI data reflecting activity across all layers of V1, including all feedback phenomena, and SSVEP data capturing the more feedforward processes contained in the granular layer of V1. At the moment this is mere speculation, and high resolution fMRI experiments would be required to more clearly delineate the role of various V1 layers to size constant perception.

Alternatively, assuming that the SSVEP is an appropriate measure for measuring size representations in V1, our findings indicate a qualitative difference between perception of static and dynamic stimuli. Indeed, if during dynamic conditions V1 as a whole codes for retinal size instead of physical size, this indicates a shift in bottom-up vs top-down emphasis. This might be explained by the increased need for processing of bottom-up information when the bottom-up information stream is conveying novel information, such as that about moving objects in the visual field. Some similar concepts have been reported in the literature, including the seminal paper by Rao and Ballard (1999) on predictive coding, in which the authors discuss the relative influence of top down and bottom up signals based on the noisiness of sensory input. Additionally, gamma band activity, thought to carry feedforward information from V1 to higher level visual areas dominate when novel visual information is presented, and the more feedback related frequency bands of alpha and beta will become more prevalent under sustained, unchanging visual input (Semoda et al., 2022). A shift in V1 activity between bottom-up and top-down dominated activity is therefore to be expected based on whether there is novel visual information.

More specific to motion, recent work by Ho & Schwarzkopf (2022) has shown that for static stimuli, V1 codes the illusory position of static stimuli, but shows veridical, retinotopic representations of moving stimuli. The authors showed a dot variant of the Müller-Lyer illusion, and resultant data showed that V1 representations of the central dots of outward pointing fins were further apart than those of inward pointing fins, despite physical distance being identical. This kind of representation of perceived location coincides with size-constancy studies using fMRI, showing perceived size to influence V1 representations. However, when the authors showed stimuli evoking a curveball illusion, where a straight moving stimulus consisting of an orthogonally moving sinusoidal grating is perceived as moving off in one direction, V1 representations accurately showed a straight movement path. Similarly to our findings, moving stimuli induce more retinotopic representations in early visual cortex than static ones. It therefore seems that motion is a critical factor in how visual information is represented in the early visual cortex.

The role of motion, or dynamic and naturalistic visual stimuli in general, therefore seem like a promising topic for future research. Given that size constancy, and perceptual constancy more generally are phenomena by which we perceive the world as constant despite changing sensory input, it stands to reason that neuroscientific inquiries into this phenomenon also make use of stimuli that elicit changing sensory inputs. Since motion has been shown to challenge previous findings on the role of early visual cortex, more use of dynamic stimuli could teach us more about how our visual system works under everyday conditions, and might challenge previously held conceptions which stem from research using strictly controlled, static stimuli.

Another important conclusion concerns the link between V1 activity and perceived size. There is much evidence to support a bidirectional, direct relation between size representations in early visual areas and our conscious size perception. Illusory effects of size are reflected in V1, and perturbation of V1 affects size perception. We demonstrate that this link is not without exceptions, since behavioural measures of size perception indicated accurate judgements of object size, in the absence of size-constant representations of these objects. More research is required to elucidate the exact connection between early visual representations of objects and our conscious percepts.

On a more pessimistic note, our results also reveal weaknesses of the methods used here. Experiments 4 and 5, although specifically designed to test the effects of motion, growth and shrinkage failed to find any such effects. Moreover, experiment 5 failed to replicate a well established finding, casting doubt on the data more generally. This leaves the role of object growth / shrinkage on resultant V1 representations of these objects in limbo, and perhaps future experiments might more conclusively investigate such an effect.

Conclusions: on size perception

Overall, behavioural measures of size perception revealed mostly accurate, size constant perception in our experiments with physically constant, moving stimuli. This was, mostly, to be expected. Despite some previous experiments indicating that motion under certain conditions might harm size perception (Brosgole et al., 1976; Suchow & Alvarez, 2011; Mruczek et al., 2022), gradual motion is generally thought to improve size perception accuracy, or at least not impede it (Combe & Wexler, 2010). Results from experiments 1, 2 and 6 showed little to no deviation from veridical perception of size for physically constant, moving stimuli, indicating size constant perception. Although this may not sound like the most riveting conclusion, accurate size perception is beneficial to our conclusions insofar that it means that our neurophysiological findings correspond with normal, functional visual processes.

Deviations from veridical size perception were repeatedly found when objects changed size as they moved. This was again found in experiments 1, 2 and 6. The direction of this effect was consistently in the direction of the original size of the object. That’s to say: the size of shrinking objects was overestimated, and the size of growing objects was underestimated. We cannot attribute this phenomenon to size change alone, since this was always coupled with movement (experiment 4 did contain shrinking stimuli that did not move, though the behavioural task here was not one of size replication). Nevertheless, this finding corresponds to the notion that there is a strong perceptual prior towards perceiving objects as rigid, referred to as the invariance principle by Herschenson (1992). Although participants were able to detect that the object changed size, the magnitude of this change was underestimated.

A more nuanced experiment of the ability to detect changes in size with concurrent movement was performed in experiment 3. From this experiment, we can conclude that even subtle changes in object size can be reliably detected while the object is moving. Further, even when judgements of the direction of size change are incorrect, observers reported a perceived distance to the stimulus which was more congruent with their perception. It therefore seems like some degree of size-distance invariance remains at play even when stimuli are moving and changing size at the same time. A more direct assessment of the invariance principle might be attempted in future experiments, by allowing participants to report a lack of size change. This would allow us to probe the strength of this perceptual bias in a more fine grained way.

Overestimation of distance to the stimulus was another recurring phenomenon in experiments 3, 4 and 5. We believe this to be specific to the virtual environments used in these experiments, and is discussed in further detail in the section “Distance overestimations in VR”.

Overall, our behavioural findings indicated a perceptual prior towards seeing objects as rigid, and maintain their physical size, leading to misestimations of size when objects physically changed their size. In the absence of physical size changes, size constancy was maintained when viewing moving stimuli in both virtual and real life viewing conditions.

A note on textbooks

Although most references made to scientific work of the past in this thesis refer to specific research articles in peer reviewed journals, I would be remiss not to mention the textbooks that have been a great help in providing a more broad overview of certain topics, as well as helpful guides in finding primary sources. Although not directly referenced or quoted in this thesis, the following textbooks have served as sources in the writing of it.

Arnold, N. (with Saulles, T. D.). (1999). *Horrible Science: Bulging Brains*. Scholastic Hippo.

Gable, P., Miller, M., & Bernat, E. (Eds.). (2022). *The Oxford Handbook of EEG Frequency*. Oxford University Press.

Howard, I. P. (2012a). *Perceiving in Depth, Volume 1: Basic Mechanisms*. OXFORD UNIV PR.

Howard, I. P. (2012b). *Perceiving in Depth, Volume 2: Stereoscopic Vision*. Oxford University Press, USA.

Howard, I. P. (2012c). *Perceiving in Depth, Volume 3: Other Mechanisms of Depth Perception*. OXFORD UNIV PR.

Kandel, E., Koester, J. D., Mack, S. H., & Siegelbaum, S. (2021). *Principles of Neural Science, Sixth Edition*. McGraw-Hill Interamericana de España S.L.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Purves, D., Cabeza, R., Huettel, S., LaBar, K., Platt, M. L., & Woldorff, M. (2013). *Principles of Cognitive Neuroscience*. Sinauer.

Regan, D. (1989). *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Elsevier.

Schomer, D. L., & Silva, F. H. L. da (Eds.). (2018). *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Oxford University Press.

Walsh, V., & Kulikowski, J. (1998). *Perceptual Constancy: Why Things Look as They Do*. Cambridge University Press.

References

Adam, K. C. S., Chang, L., Rangan, N., & Serences, J. T. (2021). Steady-state visually evoked potentials and feature-based attention: Pre-registered null results and a focused review of methodological considerations. *Journal of Cognitive Neuroscience*, *33*(4), 695–724. https://doi.org/10.1162/jocn_a_01665

Adamian, N., & Andersen, S. K. (2024). Attentional Modulation in Early Visual Cortex: A Focused Reanalysis of Steady-state Visual Evoked Potential Studies. *Journal of Cognitive Neuroscience*, *36*(1), 46–70. https://doi.org/10.1162/jocn_a_02070

Adamian, N., Andersen, S. K., & Hillyard, S. A. (2020). Parallel attentional facilitation of features and objects in early visual cortex. *Psychophysiology*, *57*(3), e13498. https://doi.org/10.1111/psyp.13498

Adams, D. L., & Horton, J. C. (2003). A Precise Retinotopic Map of Primate Striate Cortex Generated from the Representation of Angioscotomas. *Journal of Neuroscience*, *23*(9), 3771–3789. https://doi.org/10.1523/JNEUROSCI.23-09-03771.2003

Adrian, E. D., & Matthews, B. H. C. (1934). The Berger rhythm: Potential changes from the occipital lobes in man. *Brain: A Journal of Neurology*, *57*, 355–385. https://doi.org/10.1093/brain/57.4.355

Aitken, F., Menelaou, G., Warrington, O., Koolschijn, R. S., Corbin, N., Callaghan, M. F., & Kok, P. (2020). Prior expectations evoke stimulus-specific activity in the deep layers of the primary visual cortex. *PLOS Biology*, *18*(12), e3001023. https://doi.org/10.1371/journal.pbio.3001023

Allen, D., Norcia, A. M., & Tyler, C. W. (1986). Comparative Study of Electrophysiological and Psychophysical Measurement of the Contrast Sensitivity Function in Humans. *Optometry and Vision Science*, *63*(6), 442.

Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local–global comparisons in visual neurons. *Annual Review of Neuroscience*, *8*, 407–430. https://doi.org/10.1146/annurev.ne.08.030185.002203

Andersen, S. K., Fuchs, S., & Müller, M. M. (2011). Effects of feature-selective and spatial attention at different stages of visual processing. *Journal of Cognitive Neuroscience*, *23*(1), 238–246. https://doi.org/10.1162/jocn.2009.21328

Azzopardi, P., & Cowey, A. (1993). Preferential representation of the fovea in the primary visual cortex. *Nature*, *361*(6414), 719–721. https://doi.org/10.1038/361719a0

Baker, D. H., Vilidaite, G., & Wade, A. R. (2021). Steady-state measures of visual suppression. *PLOS Computational Biology*, *17*(10), e1009507. https://doi.org/10.1371/journal.pcbi.1009507

Ban, H., & Welchman, A. E. (2015). fMRI Analysis-by-Synthesis Reveals a Dorsal Hierarchy That Extracts Surface Slant. *Journal of Neuroscience*, *35*(27), 9823–9835. https://doi.org/10.1523/JNEUROSCI.1255-15.2015

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., De Weerd, P., Kennedy, H., & Fries, P. (2015). Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron*, *85*(2), 390–401. https://doi.org/10.1016/j.neuron.2014.12.018

Belmonte, M. (1998). Shifts of visual spatial attention modulate a steady-state visual evoked potential. *Cognitive Brain Research*, *6*(4), 295–307. https://doi.org/10.1016/S0926-6410(98)00007-X

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benwell, C. S. Y., London, R. E., Tagliabue, C. F., Veniero, D., Gross, J., Keitel, C., & Thut, G. (2019). Frequency and power of human alpha oscillations drift systematically with time-on-task. *NeuroImage*, *192*, 101–114. https://doi.org/10.1016/j.neuroimage.2019.02.067

Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, *87*(1), 527–570. https://doi.org/10.1007/BF01797193

Bergmann, J., Petro, L. S., Abbatecola, C., Li, M. S., Morgan, A. T., & Muckli, L. (2024). Cortical depth profiles in primary visual cortex for illusory and imaginary experiences. *Nature Communications*, *15*(1), 1002. https://doi.org/10.1038/s41467-024-45065-w

Blake, A., Bülthoff, H. H., & Sheinberg, D. (1993). Shape from texture: Ideal observers and human psychophysics. *Vision Research*, *33*(12), 1723–1737. https://doi.org/10.1016/0042-6989(93)90037-W

Blom, J. D. (2016). Alice in Wonderland syndrome. *Neurology: Clinical Practice*, *6*(3), 259–270. https://doi.org/10.1212/CPJ.0000000000000251

Bodis-Wollner, I., Hendley, C. D., & Kulikowski, J. J. (1972). Electrophysiological and Psychophysical Responses to Modulation of Contrast of a Grating Pattern. *Perception*, *1*(3), 341–349. https://doi.org/10.1068/p010341

Borjesson, E., & von Hofsten, C. (1973). Visual perception of motion in depth: Application of a vector model to three-dot motion patterns. *Perception & Psychophysics*, *13*(2), 169–179. https://doi.org/10.3758/BF03214123

BrainVision Analyzer (Version 1.21). (2019). [Computer software]. Brain Products GmbH.

Bricolo, A., Turazzi, S., Faccioli, F., Odorizzi, F., Sciarretta, G., & Erculiani, P. (1978). Clinical application of compressed spectral array in long-term EEG monitoring of comatose patients. *Electroencephalography and Clinical Neurophysiology*, *45*(2), 211–225. https://doi.org/10.1016/0013-4694(78)90005-6

Brosgole, L. (1993). Dynamic size perception as a function of target location in egocentric space. *Bulletin of the Psychonomic Society*, *31*(4), 282–284. https://doi.org/10.3758/BF03334929

Brosgole, L., McNichol, D. G., Doyle, J., & Neylon, A. (1976). Dynamic size constancy. *Bulletin of the Psychonomic Society*, *7*(1), 12–14. https://doi.org/10.3758/BF03337106

Busch, N. A., Debener, S., Kranczioch, C., Engel, A. K., & Herrmann, C. S. (2004). Size matters: Effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clinical Neurophysiology*, *115*(8), 1810–1820. https://doi.org/10.1016/j.clinph.2004.03.015

Busse, L., Cardin, J. A., Chiappe, M. E., Halassa, M. M., McGinley, M. J., Yamashita, T., & Saleem, A. B. (2017). Sensation during Active Behaviors. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *37*(45), 10826–10834. https://doi.org/10.1523/JNEUROSCI.1828-17.2017

Campbell, F. W., & Gubisch, R. W. (1966). Optical quality of the human eye. *The Journal of Physiology*, *186*(3), 558–578. https://doi.org/10.1113/jphysiol.1966.sp008056

Campbell, F. W., & Kulikowski, J. J. (1972). The visual evoked potential as a function of contrast of a grating pattern. *The Journal of Physiology*, *222*(2), 345–356. https://doi.org/10.1113/jphysiol.1972.sp009801

Campbell, F. W., & Maffei, L. (1970). Electrophysiological evidence for the existence of orientation and size detectors in the human visual system. *The Journal of Physiology*, *207*(3), 635–652. https://doi.org/10.1113/jphysiol.1970.sp009085

Cao, T., Wan, F., Wong, C. M., da Cruz, J. N., & Hu, Y. (2014). Objective evaluation of fatigue by EEG spectral analysis in steady-state visual evoked potential-based brain-computer interfaces. *BioMedical Engineering OnLine*, *13*(1), 28. https://doi.org/10.1186/1475-925X-13-28

Cavallo, F., Brubaker, B., Bruckner, E., & Castro, S. (2023). A Preliminary Study Investigating the Acquisition of valid qEEG Data While Wearing a Virtual Reality (VR) Headset. *NeuroRegulation*, *10*, 170–178. https://doi.org/10.15540/nr.10.3.170

Cavina-Pratesi, C., Goodale, M. A., & Culham, J. C. (2007). FMRI Reveals a Dissociation between Grasping and Perceiving the Size of Real 3D Objects. *PLOS ONE*, *2*(5), e424. https://doi.org/10.1371/journal.pone.0000424

Cecotti, H. (2010). Effect of the Stimulus Duty Cycle on Steady-State Visual Evoked Potential detection. *Neurocomp 2010 - Cinquième Conférence Plénière Française de Neurosciences Computationnelles - 5th French Conference on Computational Neuroscience / Cinquième Conférence Plénière Française de Neurosciences Computationnelles*, 205–209. https://hal.science/hal-00536118

Celesia, G. G., Meredith, J. T., & Pluff, K. (1983). Perimetry, visual evoked potentials and visual evoked spectrum array in homonymous hemianopsia. *Electroencephalography and Clinical Neurophysiology*, *56*(1), 16–30. https://doi.org/10.1016/0013-4694(83)90003-2

Cepeda-Freyre, H. A., Garcia-Aguilar, G., Eguibar, J. R., & Cortes, C. (2020). Brain Processing of Complex Geometric Forms in a Visual Memory Task Increases P2 Amplitude. *Brain Sciences*, *10*(2), Article 2. https://doi.org/10.3390/brainsci10020114

Chang, M. H., Baek, H. J., Lee, S. M., & Park, K. S. (2014). An amplitude-modulated visual stimulation for reducing eye fatigue in SSVEP-based brain–computer interfaces. *Clinical Neurophysiology*, *125*(7), 1380–1391. https://doi.org/10.1016/j.clinph.2013.11.016

Chen, J., McManus, M., Valsecchi, M., Harris, L. R., & Gegenfurtner, K. R. (2019). Steady-state visually evoked potentials reveal partial size constancy in early visual cortex. *Journal of Vision*, *19*(6), 8. https://doi.org/10.1167/19.6.8

Chen, J., Sperandio, I., Henry, M. J., & Goodale, M. A. (2019). Changing the Real Viewing Distance Reveals the Temporal Evolution of Size Constancy in Visual Cortex. *Current Biology*, *29*(13), 2237-2243.e4. https://doi.org/10.1016/j.cub.2019.05.069

Chen, L., Wu, B., Yu, H., & Sperandio, I. (2024). Network dynamics underlying alterations in apparent object size. *Brain Communications*, *6*(1), fcae006. https://doi.org/10.1093/braincomms/fcae006

Chen, S., Weidner, R., Zeng, H., Fink, G. R., Müller, H. J., & Conci, M. (2021). Feedback from lateral occipital cortex to V1/V2 triggers object completion: Evidence from functional magnetic resonance imaging and dynamic causal modeling. *Human Brain Mapping*, *42*(17), 5581–5594. https://doi.org/10.1002/hbm.25637

Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.-P., & Gao, S. (2015). High-speed spelling with a noninvasive brain–computer interface. *Proceedings of the National Academy of Sciences*, *112*(44), E6058–E6067. https://doi.org/10.1073/pnas.1508080112

Chen, X., Wang, Y., Zhang, S., Xu, S., & Gao, X. (2019). Effects of stimulation frequency and stimulation waveform on steady-state visual evoked potentials using a computer monitor. *Journal of Neural Engineering*, *16*(6), 066007. https://doi.org/10.1088/1741-2552/ab2b7d

Clavagnier, S., Falchier, A., & Kennedy, H. (2004). Long-distance feedback projections to area V1: Implications for multisensory integration, spatial awareness, and visual consciousness. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 117–126. https://doi.org/10.3758/CABN.4.2.117

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. The MIT Press. https://doi.org/10.7551/mitpress/9609.001.0001

Cohen, M. X., & Gulbinaite, R. (2017). Rhythmic entrainment source separation: Optimizing analyses of neural responses to rhythmic sensory stimulation. *NeuroImage*, *147*, 43–56. https://doi.org/10.1016/j.neuroimage.2016.11.036

Combe, E., & Wexler, M. (2010). Observer movement and size constancy. *Psychological Science*, *21*(5), 667–675. https://doi.org/10.1177/0956797610367753

Cottereau, B., Durand, J.-B., Vayssiere, N., & Trotter, Y. (2014, August). *The influence of eye vergence on retinotopic organization in human early visual cortex* [Poster].

European Conference on Visual Perception, Belgrade, Serbia. https://journals.sagepub.com/doi/10.1177/03010066140430S101

Cumming, B. G., & DeAngelis, G. C. (2001). The Physiology of Stereopsis. *Annual Review of Neuroscience*, *24*(Volume 24, 2001), 203–238. https://doi.org/10.1146/annurev.neuro.24.1.203

Cumming, B. G., & Parker, A. J. (2000). Local Disparity Not Perceived Depth Is Signaled by Binocular Neurons in Cortical Area V1 of the Macaque. *Journal of Neuroscience*, *20*(12), 4758–4767. https://doi.org/10.1523/JNEUROSCI.20-12-04758.2000

Davidson, M. J., Mithen, W., Hogendoorn, H., van Boxtel, J. J., & Tsuchiya, N. (2020). The SSVEP tracks attention, not consciousness, during perceptual filling-in. *eLife*, *9*, e60031. https://doi.org/10.7554/eLife.60031

Day, R. H., & McKenzie, B. E. (1981). Infant perception of the invariant size of approaching and receding objects. *Developmental Psychology*, *17*(5), 670–677. https://doi.org/10.1037/0012-1649.17.5.670

de Hollander, G., van der Zwaag, W., Qian, C., Zhang, P., & Knapen, T. (2021). Ultra-high field fMRI reveals origins of feedforward and feedback activity within laminae of human ocular dominance columns. *NeuroImage*, *228*, 117683. https://doi.org/10.1016/j.neuroimage.2020.117683

de Lange, H. (1958). Research into the Dynamic Nature of the Human Fovea-Cortex Systems with Intermittent and Modulated Light. I. Attenuation Characteristics with White and Colored Light. *JOSA*, *48*(11), 777–784. https://doi.org/10.1364/JOSA.48.000777

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Di Russo, F., Pitzalis, S., Aprile, T., Spitoni, G., Patria, F., Stella, A., Spinelli, D., & Hillyard, S. A. (2007). Spatiotemporal analysis of the cortical sources of the steady-state visual evoked potential. *Human Brain Mapping*, *28*(4), 323–334. https://doi.org/10.1002/hbm.20276

Ding, J., Sperling, G., & Srinivasan, R. (2006). Attentional modulation of SSVEP power depends on the network tagged by the flicker frequency. *Cerebral Cortex (New York, N.Y. : 1991)*, *16*(7), 1016–1029. https://doi.org/10.1093/cercor/bhj044

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision*, *3*(10), 1. https://doi.org/10.1167/3.10.1

Duart, X., Quiles, E., Suay, F., Chio, N., García, E., & Morant, F. (2020). Evaluating the Effect of Stimuli Color and Frequency on SSVEP. *Sensors (Basel, Switzerland)*, *21*(1), 117. https://doi.org/10.3390/s21010117

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron*, *38*(4), 659–671. https://doi.org/10.1016/s0896-6273(03)00265-4

Duszyk, A., Bierzyńska, M., Radzikowska, Z., Milanowski, P., Kuś, R., Suffczyński, P., Michalska, M., Łabęcki, M., Zwoliński, P., & Durka, P. (2014). Towards an Optimization of Stimulus Parameters for Brain-Computer Interfaces Based on Steady State Visual Evoked Potentials. *PLOS ONE*, *9*(11), e112099. https://doi.org/10.1371/journal.pone.0112099

Emmert, E. (n.d.). Grössenverhältnisse der Nachbilder. *Klinische Monatsblätter Für Augenheilkunde*, *19*, 443–450.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex (New York, N.Y.: 1991)*, *7*(2), 181–192. https://doi.org/10.1093/cercor/7.2.181

Epstein, W., Park, J., & Casey, A. (1961). The current status of the size-distance hypotheses. *Psychological Bulletin*, *58*(6), 491–514. https://doi.org/10.1037/h0042260

Fang, F., Boyaci, H., Kersten, D., & Murray, S. O. (2008). Attention-dependent representation of a size illusion in human V1. *Current Biology: CB*, *18*(21), 1707–1712. https://doi.org/10.1016/j.cub.2008.09.025

Fişek, M., Herrmann, D., Egea-Weiss, A., Cloves, M., Bauer, L., Lee, T.-Y., Russell, L. E., & Häusser, M. (2023). Cortico-cortical feedback engages active dendrites in visual cortex. *Nature*, *617*(7962), 769–776. https://doi.org/10.1038/s41586-023-06007-6

Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, *65*(7), 999–1010. https://doi.org/10.3758/BF03194829

Fraser, L. E., Makooie, B., & Harris, L. R. (2015). The Subjective Visual Vertical and the Subjective Haptic Vertical Access Different Gravity Estimates. *PloS One*, *10*(12), e0145528. https://doi.org/10.1371/journal.pone.0145528

Friedrich, M. U., Baughan, E. C., Kletenik, I., Younger, E., Zhao, C. W., Howard, C., Ferguson, M. A., Schaper, F. L. W. V. J., Chen, A., Zeller, D., Piervincenzi, C., Tommasin, S., Pantano, P., Blanke, O., Prasad, S., Nielsen, J. A., & Fox, M. D. (2024). Lesions Causing Alice in Wonderland Syndrome Map to a Common Brain Network Linking Body and Size Perception. *Annals of Neurology*, *96*(4), 662–674. https://doi.org/10.1002/ana.27015

Gable, P. A., & Miller, M. W. (2022). Introduction: Methods for collecting EEG data for frequency analyses in humans. In P. A. Gable, M. W. Miller, & E. M. Bernat (Eds), *The Oxford handbook of EEG frequency*. Oxford University Press.

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nature Reviews Neuroscience*, *4*(7), 563–572. https://doi.org/10.1038/nrn1138

Georgieva, S. S., Todd, J. T., Peeters, R., & Orban, G. A. (2008). The Extraction of 3D Shape from Texture and Shading in the Human Brain. *Cerebral Cortex*, *18*(10), 2416–2438. https://doi.org/10.1093/cercor/bhn002

Giangrande, A., Botter, A., Piitulainen, H., & Cerone, G. L. (2024). Motion Artifacts in Dynamic EEG Recordings: Experimental Observations, Electrical Modelling, and

Design Considerations. *Sensors*, *24*(19), Article 19. https://doi.org/10.3390/s24196363

Gilinsky, A. S. (1955). The Effect of Attitude upon the Perception of Size. *The American Journal of Psychology*, *68*(2), 173–192. https://doi.org/10.2307/1418890

Gittinger, J. W., & Sokol, S. (1982). The visual-evoked potential in the diagnosis of congenital ocular motor apraxia. *American Journal of Ophthalmology*, *93*(6), 700–703. https://doi.org/10.1016/0002-9394(82)90462-7

Gogel, W. C. (1969). The sensing of retinal size. *Vision Research*, *9*(9), 1079–1094. https://doi.org/10.1016/0042-6989(69)90049-2

Gogel, W. C. (1998). An analysis of perceptions from changes in optical size. *Perception & Psychophysics*, *60*(5), 805–820. https://doi.org/10.3758/BF03206064

Gogel, W. C., & Da Silva, J. A. (1987a). A two-process theory of the response to size and distance. *Perception & Psychophysics*, *41*(3), 220–238. https://doi.org/10.3758/bf03208221

Gogel, W. C., & Da Silva, J. A. (1987b). Familiar size and the theory of off-sized perceptions. *Perception & Psychophysics*, *41*(4), 318–328. https://doi.org/10.3758/BF03208233

Gogel, W. C., & Eby, D. W. (1997). Measures of perceived linear size, sagittal motion, and visual angle from optical expansions and contractions. *Perception & Psychophysics*, *59*(5), 783–806. https://doi.org/10.3758/BF03206024

Gogel, W. C., & Tietz, J. D. (1992). Determinants ofthe perception of sagittal motion. *Perception & Psychophysics*, *52*(1), 75–96. https://doi.org/10.3758/BF03206761

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25. https://doi.org/10.1016/0166-2236(92)90344-8

Gordon, N., Koenig-Robert, R., Tsuchiya, N., van Boxtel, J. J., & Hohwy, J. (n.d.). Neural markers of predictive coding under perceptual uncertainty revealed with Hierarchical Frequency Tagging. *eLife*, *6*, e22749. https://doi.org/10.7554/eLife.22749

Gordon, N., Tsuchiya, N., Koenig-Robert, R., & Hohwy, J. (2019). Expectation and attention increase the integration of top-down and bottom-up signals in perception through different pathways. *PLOS Biology*, *17*(4), e3000233. https://doi.org/10.1371/journal.pbio.3000233

Gray, M., Kemp, A. H., Silberstein, R. B., & Nathan, P. J. (2003). Cortical neurophysiology of anticipatory anxiety: An investigation utilizing steady state probe topography (SSPT). *NeuroImage*, *20*(2), 975–986. https://doi.org/10.1016/S1053-8119(03)00401-4

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10), 1409–1422. https://doi.org/10.1016/S0042-6989(01)00073-6

Guo, M., Yue, K., Hu, H., Lu, K., Han, Y., Chen, S., & Liu, Y. (2022). Neural Research on Depth Perception and Stereoscopic Visual Fatigue in Virtual Reality. *Brain Sciences*, *12*(9), 1231. https://doi.org/10.3390/brainsci12091231

Harvey, B. M., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences*, *112*(44), 13525–13530. https://doi.org/10.1073/pnas.1515414112

He, D., Mo, C., Wang, Y., & Fang, F. (2015). Position shifts of fMRI-based population receptive fields in human visual cortex induced by Ponzo illusion. *Experimental Brain Research*, *233*(12), 3535–3541. https://doi.org/10.1007/s00221-015-4425-3

Hedjar, L., Martinovic, J., Andersen, S. K., & Shapiro, A. G. (2025). Separation of luminance and contrast modulation in steady-state visual evoked potentials. *Vision Research*, *230*, 108567. https://doi.org/10.1016/j.visres.2025.108567

Helmholtz, H. von (with University of California Libraries). (1867). *Handbuch der physiologischen Optik*. Leipzig : Leopold Voss. http://archive.org/details/handbuchderphysi00helm

Herrmann, C. S. (2001). Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental Brain Research*, *137*(3), 346–353. https://doi.org/10.1007/s002210100682

Hershenson, M. (1982). Moon illusion and spiral aftereffect: Illusions due to the loom-zoom system? *Journal of Experimental Psychology: General*, *111*(4), 423–440. https://doi.org/10.1037/0096-3445.111.4.423

Hershenson, M. (1992). Size-distance invariance: Kinetic invariance is different from static invariance. *Perception & Psychophysics*, *51*(6), 541–548. https://doi.org/10.3758/BF03211651

Hershenson, M. (1993). Structural constraints: Further evidence from apparent motion in depth. *Perception*, *22*(3), 323–334. https://doi.org/10.1068/p220323

Hervais-Adelman, A., Legrand, L. B., Zhan, M., Tamietto, M., de Gelder, B., & Pegna, A. J. (2015). Looming sensitive cortical regions without V1 input: Evidence from a patient with bilateral cortical blindness. *Frontiers in Integrative Neuroscience*, *9*. https://doi.org/10.3389/fnint.2015.00051

Ho, M.-L., & Schwarzkopf, D. S. (2022). The human primary visual cortex (V1) encodes the perceived position of static but not moving objects. *Communications Biology*, *5*(1), 181. https://doi.org/10.1038/s42003-022-03136-y

Holway, A. H., & Boring, E. G. (1941). Determinants of apparent visual size with distance variant. *The American Journal of Psychology*, *54*, 21–37. https://doi.org/10.2307/1417790

Hornsey, R. L., Hibbard, P. B., & Scarfe, P. (2020). Size and shape constancy in consumer virtual reality. *Behavior Research Methods*, *52*(4), 1587–1598. https://doi.org/10.3758/s13428-019-01336-9

Inouye, T. (1909). *Die Sehstorungen bei Schussverletzungen der kortikalen Sehsphare. Nach Beobachtungen an Verwundeten der letzten japanischen Kriege*. Leipzig : Engelmann. http://archive.org/details/b21271999

Interrante, V., Ries, B., & Anderson, L. (2006). Distance Perception in Immersive Virtual Environments, Revisited. *IEEE Virtual Reality Conference (VR 2006)*, 3–10. https://doi.org/10.1109/VR.2006.52

Itaguchi, Y. (2021). Size Perception Bias and Reach-to-Grasp Kinematics: An Exploratory Study on the Virtual Hand With a Consumer Immersive Virtual-Reality Device. *Frontiers in Virtual Reality*, *2*. https://doi.org/10.3389/frvir.2021.712378

Ittelson, W. H. (1961). Size as a cue to distance. In F. P. Kilpatrick (Ed.), *Explorations in transactional psychology.* (pp. 69–98). New York University Press. https://doi.org/10.1037/11303-004

Johansson, G. (1964). PERCEPTION OF MOTION AND CHANGING FORM: A study of visual perception from continuous transformations of a solid angle of light at the eye. *Scandinavian Journal of Psychology*, *5*(1), 181–208. https://doi.org/10.1111/j.1467-9450.1964.tb01425.x

Johansson, G. (1975). Visual Motion Perception. *Scientific American*, *232*(6), 76–89.

Kayser, J., & Tenke, C. E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *117*(2), 348–368. https://doi.org/10.1016/j.clinph.2005.08.034

Keller, G. B., Bonhoeffer, T., & Hübener, M. (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. *Neuron*, *74*(5), 809–815. https://doi.org/10.1016/j.neuron.2012.03.040

Kelly, J., Cherep, L., & Siegel, Z. (2017). Perceived space in the HTC Vive. *ACM Transactions on Applied Perception*, *15*, 1–16. https://doi.org/10.1145/3106155

Kelly, J. W. (2023). Distance Perception in Virtual Reality: A Meta-Analysis of the Effect of Head-Mounted Display Characteristics. *IEEE Transactions on Visualization and Computer Graphics*, *29*(12), 4978–4989. https://doi.org/10.1109/TVCG.2022.3196606

Kelly, J. W., Doty, T. A., Ambourn, M., & Cherep, L. A. (2022). Distance Perception in the Oculus Quest and Oculus Quest 2. *Frontiers in Virtual Reality*, *3*. https://doi.org/10.3389/frvir.2022.850471

Kemp, J. T., Cesanek, E., & Domini, F. (2023). Perceiving depth from texture and disparity cues: Evidence for a non-probabilistic account of cue integration. *Journal of Vision*, *23*(7), 13. https://doi.org/10.1167/jov.23.7.13

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304. https://doi.org/10.1146/annurev.psych.55.090902.142005

Kilpatrick, F. P., & Ittelson, W. H. (1953). The size-distance invariance hypothesis. *Psychological Review*, *60*(4), 223–231. https://doi.org/10.1037/h0060882

Kim, N.-G. (2018). Independence of Size and Distance in Binocular Vision. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00988

Kim, Y. C., & Joo, S. J. (2024). Size constancy in Virtual Reality. *Journal of Vision*, *24*(10), 1391. https://doi.org/10.1167/jov.24.10.1391

Klatzky, R. L., Lederman, S. J., & Metzger, V. A. (1985). Identifying objects by touch: An ‘expert system.’ *Perception & Psychophysics*, *37*(4), 299–302. https://doi.org/10.3758/BF03211351

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What’s new in psychtoolbox-3. *Perception*, *36*(14), 1–16.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*(12), 606–617. https://doi.org/10.1016/j.tics.2012.10.007

Klimesch, W., Schack, B., Schabus, M., Doppelmayr, M., Gruber, W., & Sauseng, P. (2004). Phase-locked alpha and theta oscillations generate the P1–N1 complex and are related to memory performance. *Cognitive Brain Research*, *19*(3), 302–316. https://doi.org/10.1016/j.cogbrainres.2003.11.016

Kok, P., & de Lange, F. P. (2014). Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. *Current Biology: CB*, *24*(13), 1531–1535. https://doi.org/10.1016/j.cub.2014.05.042

Konkle, T., & Oliva, A. (2012). A Real-World Size Organization of Object Responses in Occipitotemporal Cortex. *Neuron*, *74*(6), 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036

Kristensen, S., Fracasso, A., Dumoulin, S. O., Almeida, J., & Harvey, B. M. (2021). Size constancy affects the perception and parietal neural representation of object size. *NeuroImage*, *232*, 117909. https://doi.org/10.1016/j.neuroimage.2021.117909

Kunz, B. R., Wouters, L., Smith, D., Thompson, W. B., & Creem-Regehr, S. H. (2009). Revisiting the effect of quality of graphics on distance judgments in virtual environments: A comparison of verbal reports and blind walking. *Attention, Perception, & Psychophysics*, *71*(6), 1284–1293. https://doi.org/10.3758/APP.71.6.1284

Labecki, M., Nowicka, M. M., & Suffczynski, P. (2019). Temporal Modulation of Steady-State Visual Evoked Potentials. *International Journal of Neural Systems*, *29*(03), 1850050. https://doi.org/10.1142/S0129065718500508

Łabęcki, M., Nowicka, M. M., Wróbel, A., & Suffczynski, P. (2024). Frequency-dependent dynamics of steady-state visual evoked potentials under sustained flicker stimulation. *Scientific Reports*, *14*(1), 9281. https://doi.org/10.1038/s41598-024-59770-5

Ladouce, S., Darmet, L., Torre Tresols, J. J., Velut, S., Ferraro, G., & Dehais, F. (2022). Improving user experience of SSVEP BCI through low amplitude depth and high frequency stimuli design. *Scientific Reports*, *12*(1), 8865. https://doi.org/10.1038/s41598-022-12733-0

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–579. https://doi.org/10.1016/S0166-2236(00)01657-X

Lawrence, S. J., Norris, D. G., & de Lange, F. P. (2019). Dissociable laminar profiles of concurrent bottom-up and top-down modulation in the human visual cortex. *eLife*, *8*, e44422. https://doi.org/10.7554/eLife.44422

Lawson, R. (1999). Achieving visual object constancy across plane rotation and depth rotation. *Acta Psychologica*, *102*(2–3), 221–245. https://doi.org/10.1016/s0001-6918(98)00052-3

Lee, S., Kim, M., & Ahn, M. (2024). Evaluation of consumer-grade wireless EEG systems for brain-computer interface applications. *Biomedical Engineering Letters*, *14*(6), 1433–1443. https://doi.org/10.1007/s13534-024-00416-w

Li, H., Shin, H., Zhang, M., Yu, A., Huh, H., Kwon, G., Riveira, N., Kim, S., Gangopadahyay, S., Peng, J., Li, Z., Rao, Y., Sentis, L., Millán, J. del R., & Lu, N. (2023). Hair-compatible sponge electrodes integrated on VR headset for electroencephalography. *Soft Science*, *3*(3), N/A-N/A. https://doi.org/10.20517/ss.2023.11

Lichten, W., & Lurie, S. (1950). A New Technique for the Study of Perceived Size. *The American Journal of Psychology*, *63*(2), 280–282. https://doi.org/10.2307/1418936

Lissa, P. de, Caldara, R., Nicholls, V., & Miellet, S. (2020). In pursuit of visual attention: SSVEP frequency-tagging moving targets. *PLOS ONE*, *15*(8), e0236967. https://doi.org/10.1371/journal.pone.0236967

Liu, H., Wang, Z., Li, R., Zhao, X., Xu, T., Zhou, T., & Hu, H. (2024). A comparative study of stereo-dependent SSVEP targets and their impact on VR-BCI performance. *Frontiers in Neuroscience*, *18*. https://doi.org/10.3389/fnins.2024.1367932

Loomis, J., & Knapp, J. (2003). Visual perception of egocentric distance in real and virtual environments. In *Virtual and adaptive environments: Applications, implications, and human performance issues* (pp. 21–46). Lawrence Erlbaum Associates Publishers. https://doi.org/10.1201/9781410608888.pt1

Luo, X., Kenyon, R. V., Kamper, D. G., Sandin, D. J., & DeFanti, T. A. (2009). On the Determinants of Size-Constancy in a Virtual Environment. *International Journal of Virtual Reality*, *8*(1), Article 1. https://doi.org/10.20870/IJVR.2009.8.1.2712

MacEvoy, S. P., & Fitzpatrick, D. (2006). Visual Physiology: Perceived Size Looms Large. *Current Biology*, *16*(9), R330–R332. https://doi.org/10.1016/j.cub.2006.03.076

MacEvoy, S. P., & Paradiso, M. A. (2001). Lightness constancy in primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(15), 8827–8831. https://doi.org/10.1073/pnas.161280398

Marg, E., & Adams, J. E. (1970). Evidence for a neurological zoom system in vision from angular changes in some receptive fields of single neurons with changes in fixation distance in the human visual cortex. *Experientia*, *26*(3), 270–271. https://doi.org/10.1007/BF01900088

MathWorks. (2022a). *Instrument Control Toolbox* (Version R2022a) [MATLAB]. The MathWorks, Inc.

MathWorks. (2022b). *MATLAB* (Version R2022a) [MATLAB]. The MathWorks, Inc.

McKeefry, D. J., Russell, M. H. A., Murray, I. J., & Kulikowski, J. J. (1996). Amplitude and phase variations of harmonic components in human achromatic and chromatic visual evoked potentials. *Visual Neuroscience*, *13*(4), 639–653. https://doi.org/10.1017/S0952523800008543

McKenzie, B. E., Tootell, H. E., & Day, R. H. (1980). Development of visual size constancy during the 1st year of human infancy. *Developmental Psychology*, *16*(3), 163–174. https://doi.org/10.1037/0012-1649.16.3.163

Meigen, T., & Bach, M. (1999). On the statistical significance of electrophysiological steady-state responses. *Documenta Ophthalmologica*, *98*(3), 207–232. https://doi.org/10.1023/A:1002097208337

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H., & Fries, P. (2016). Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward

Influences among Human Visual Cortical Areas. *Neuron*, *89*(2), 384–397. https://doi.org/10.1016/j.neuron.2015.12.018

Middendorf, M., McMillan, G., Calhoun, G., & Jones, K. S. (2000). Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions on Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, *8*(2), 211–214. https://doi.org/10.1109/86.847819

Mohr, K. S., Geuzebroek, A. C., & Kelly, S. P. (2024). Visual cortical area contributions to the transient, multifocal and steady-state VEP: A forward model-informed analysis. *Imaging Neuroscience*, *2*, 1–26. https://doi.org/10.1162/imag_a_00152

Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(10), 4770–4774. https://doi.org/10.1073/pnas.93.10.4770

Moutsiana, C., de Haas, B., Papageorgiou, A., van Dijk, J. A., Balraj, A., Greenwood, J. A., & Schwarzkopf, D. S. (2016). Cortical idiosyncrasies predict the perception of object size. *Nature Communications*, *7*(1), 12110. https://doi.org/10.1038/ncomms12110

Mruczek, R. E. B., Blair, C. D., Cullen, K., & Caplovitz, G. P. (2020). Opposite effects of motion dynamics on the Ebbinghaus and corridor illusions. *Attention, Perception, & Psychophysics*, *82*(4), 1912–1927. https://doi.org/10.3758/s13414-019-01927-w

Mruczek, R. E. B., Blair, C. D., Strother, L., & Caplovitz, G. P. (2015). The Dynamic Ebbinghaus: Motion dynamics greatly enhance the classic contextual size illusion. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00077

Mruczek, R. E. B., Fanelli, M., Kelly, S., & Caplovitz, G. P. (2022). The combination of target motion and dynamic changes in context greatly enhance visual size illusions. *Frontiers in Human Neuroscience*, *16*. https://doi.org/10.3389/fnhum.2022.959367

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., Goebel, R., & Yacoub, E. (2015). Contextual Feedback to Superficial Layers of V1. *Current Biology*, *25*(20), 2690–2695. https://doi.org/10.1016/j.cub.2015.08.057

Müller, M. M., Teder-Sälejärvi, W., & Hillyard, S. A. (1998). The time course of cortical facilitation during cued shifts of spatial attention. *Nature Neuroscience*, *1*(7), 631–634. https://doi.org/10.1038/2865

Murgia, A., & Sharkey, P. M. (2009). *Estimation of Distances in Virtual Environments Using Size Constancy*. *8*, 67–74. https://doi.org/10.20870/IJVR.2009.8.1.2714

Murray, S., Boyaci, H., & Kersten, D. (2006). The Representation of Perceived Angular Size in Human Primary Visual Cortex. *Nature Neuroscience*, *9*, 429–434. https://doi.org/10.1038/nn1641

Nefs, H. T., O'Hare, L., & Harris, J. M. (2010). Two Independent Mechanisms for Motion-In-Depth Perception: Evidence from Individual Differences. *Frontiers in Psychology*, *1*, 155. https://doi.org/10.3389/fpsyg.2010.00155

Ng, K. B., Bradley, A. P., & Cunnington, R. (2012). Stimulus specificity of a steady-state visual-evoked potential-based brain–computer interface. *Journal of Neural Engineering*, *9*(3), 036008. https://doi.org/10.1088/1741-2560/9/3/036008

Ni, A. M., Murray, S. O., & Horwitz, G. D. (2014). Object-centered shifts of receptive field positions in monkey primary visual cortex. *Current Biology: CB*, *24*(14), 1653–1658. https://doi.org/10.1016/j.cub.2014.06.003

Noguchi, K., & Taya, K. (1981). A neglected problem: Kinetic size constancy. *Acta Psychologica*, *48*(1), 187–194. https://doi.org/10.1016/0001-6918(81)90060-3

Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottereau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, *15*(6), 4. https://doi.org/10.1167/15.6.4

Olman, C. A., Harel, N., Feinberg, D. A., He, S., Zhang, P., Ugurbil, K., & Yacoub, E. (2012). Layer-Specific fMRI Reflects Different Neuronal Computations at Different

Depths in Human V1. *PLOS ONE*, *7*(3), e32536. https://doi.org/10.1371/journal.pone.0032536

Pai, D., Rolin, R., Fooken, J., & Spering, M. (2017). Time to Contact Estimation in Virtual Reality. *Journal of Vision*, *17*(10), 417. https://doi.org/10.1167/17.10.417

Panitz, C., Gundlach, C., Boylan, M. R., Keil, A., & Müller, M. M. (2023). Higher amplitudes in steady-state visual evoked potentials driven by square-wave versus sine-wave contrast modulation – A dual-laboratory study. *Psychophysiology*, *60*(8), e14287. https://doi.org/10.1111/psyp.14287

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, *8*(5), 379–391. https://doi.org/10.1038/nrn2131

Parker, P. R. L., Abe, E. T. T., Leonard, E. S. P., Martins, D. M., & Niell, C. M. (2022). Joint coding of visual input and eye/head position in V1 of freely moving mice. *Neuron*, *110*(23), 3897-3906.e5. https://doi.org/10.1016/j.neuron.2022.08.029

Pei, F., Pettet, M. W., & Norcia, A. M. (2002). Neural correlates of object-based attention. *Journal of Vision*, *2*(9), 1. https://doi.org/10.1167/2.9.1

Pennartz, C. M. A., Dora, S., Muckli, L., & Lorteije, J. A. M. (2019). Towards a Unified View on Pathways and Functions of Neural Recurrent Processing. *Trends in Neurosciences*, *42*(9), 589–603. https://doi.org/10.1016/j.tins.2019.07.005

Pereverzeva, M., & Murray, S. O. (2008). Neural activity in human V1 correlates with dynamic lightness induction. *Journal of Vision*, *8*(15), 8. https://doi.org/10.1167/8.15.8

Pfabigan, D. M., Sailer, U., & Lamm, C. (2015). Size does matter! Perceptual stimulus properties affect event-related potentials during feedback processing. *Psychophysiology*, *52*(9), 1238–1247. https://doi.org/10.1111/psyp.12458

Polat, U., & Norcia, A. M. (1996). Neurophysiological Evidence for Contrast Dependent Long-range Facilitation and Suppression in the Human Visual Cortex. *Vision Research*, *36*(14), 2099–2109. https://doi.org/10.1016/0042-6989(95)00281-2

Pooresmaeili, A., Arrighi, R., Biagi, L., & Morrone, M. C. (2013). Blood oxygen level-dependent activation of the primary visual cortex predicts size adaptation illusion. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *33*(40), 15999–16008. https://doi.org/10.1523/JNEUROSCI.1770-13.2013

Prado-Gutiérrez, P., Otero, M., Martínez-Montes, E., Weinstein, A., Escobar, M.-J., El-Deredy, W., & Zañartu, M. (2019). A Method for Tracking the Time Evolution of Steady-State Evoked Potentials. *Journal of Visualized Experiments : JoVE*, *147*, 10.3791/59898. https://doi.org/10.3791/59898

Purkyně, J. E. (1818). *Beiträge zur Kenntniss des Sehens in subjectiver Hinsicht*. In Commission bei Johann Gottfried Calve.

Radtke, E. L., Martens, U., & Gruber, T. (2021). The steady-state visual evoked potential (SSVEP) reflects the activation of cortical object representations: Evidence from semantic stimulus repetition. *Experimental Brain Research*, *239*(2), 545–555. https://doi.org/10.1007/s00221-020-05992-8

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Ray, W. J., & Cole, H. W. (1985). EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science (New York, N.Y.)*, *228*(4700), 750–752. https://doi.org/10.1126/science.3992243

Regan, D. (1966). Some characteristics of average steady-state and transient responses evoked by modulated light. *Electroencephalography and Clinical Neurophysiology*, *20*(3), 238–248. https://doi.org/10.1016/0013-4694(66)90088-5

Regan, D. (1973). Evoked potentials specific to spatial patterns of luminance and colour. *Vision Research*, *13*(12), 2381–2402. https://doi.org/10.1016/0042-6989(73)90237-X

Regan, D. (1975). Recent advances in electrical recording from the human brain. *Nature*, *253*(5491), 401–407. https://doi.org/10.1038/253401a0

Regan, D. (1977). Steady-state evoked potentials. *JOSA*, *67*(11), 1475–1489. https://doi.org/10.1364/JOSA.67.001475

Regan, D. (1989). *Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine*. Elsevier.

Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments—A review. *ACM Comput. Surv.*, *46*(2), 23:1-23:40. https://doi.org/10.1145/2543581.2543590

Rock, I., Shallo, J., & Schwartz, F. (1978). Pictorial depth and related constancy effects as a function of recognition. *Perception*, *7*(1), 3–19. https://doi.org/10.1068/p070003

Rossini, J. C. (2014). Looming motion and visual attention. *Psychology & Neuroscience*, *7*, 425–431. https://doi.org/10.3922/j.psns.2014.042

Rzepka, A. M., Hussey, K. J., Maltz, M. V., Babin, K., Wilcox, L. M., & Culham, J. C. (2022). Familiar size affects perception differently in virtual reality and the real world. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1869), 20210464. https://doi.org/10.1098/rstb.2021.0464

Saleem, A. B., Ayaz, A., Jeffery, K. J., Harris, K. D., & Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. *Nature Neuroscience*, *16*(12), 1864–1869. https://doi.org/10.1038/nn.3567

Salelkar, S., & Ray, S. (2020). Interaction between steady-state visually evoked potentials at nearby flicker frequencies. *Scientific Reports*, *10*(1), 5344. https://doi.org/10.1038/s41598-020-62180-y

Schellekens, W., van Wezel, R. J. A., Petridou, N., Ramsey, N. F., & Raemaekers, M. (2016). Predictive coding for motion stimuli in human early visual cortex. *Brain Structure and Function*, *221*(2), 879–890. https://doi.org/10.1007/s00429-014-0942-2

Schira, M. M., Tyler, C. W., Breakspear, M., & Spehar, B. (2009). The Foveal Confluence in Human Visual Cortex. *Journal of Neuroscience*, *29*(28), 9050–9058. https://doi.org/10.1523/JNEUROSCI.1760-09.2009

Schütt, H., Harmeling, S., Macke, J., & Wichmann, F. (2015). Psignifit 4: Pain-free Bayesian Inference for Psychometric Functions. *Journal of Vision*, *15*(12), 474. https://doi.org/10.1167/15.12.474

Schwarzkopf, D. S., & Rees, G. (2013). Subjective size perception depends on central visual cortical magnification in human v1. *PloS One*, *8*(3), e60550. https://doi.org/10.1371/journal.pone.0060550

Schwarzkopf, D. S., Song, C., & Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nature Neuroscience*, *14*(1), 28–30. https://doi.org/10.1038/nn.2706

Semedo, J. D., Jasper, A. I., Zandvakili, A., Krishna, A., Aschner, A., Machens, C. K., Kohn, A., & Yu, B. M. (2022). Feedforward and feedback interactions between visual cortical areas use different population activity patterns. *Nature Communications*, *13*(1), 1099. https://doi.org/10.1038/s41467-022-28552-w

Shishavan, H. H., Roy, R., Golzari, K., Singla, A., Zalozhin, D., Lohan, D., Farooq, M., Dede, E. M., & Kim, I. (2024). Optimization of stimulus properties for SSVEP-based BMI system with a heads-up display to control in-vehicle features. *PLOS ONE*, *19*(9), e0308506. https://doi.org/10.1371/journal.pone.0308506

Siegel, Z. D., Kelly, J. W., & Cherep, L. A. (2017). Rescaling of perceived space transfers across virtual environments. *Journal of Experimental Psychology. Human Perception and Performance*, *43*(10), 1805–1814. https://doi.org/10.1037/xhp0000401

Silberstein, R. B., Schier, M. A., Pipingas, A., Ciorciari, J., Wood, S. R., & Simpson, D. G. (1990). Steady-State Visually Evoked Potential topography associated with a visual vigilance task. *Brain Topography*, *3*(2), 337–347. https://doi.org/10.1007/BF01135443

Smith, A. M. (1996). *Ptolemy's Theory of Visual Perception: An English Translation of the Optics*. American Philosophical Society.

Smith, K. C., & Abrams, R. A. (2018). Motion onset really does capture attention. *Attention, Perception, & Psychophysics*, *80*(7), 1775–1784. https://doi.org/10.3758/s13414-018-1548-1

Spekreijse, H. (1966). *Analysis of e. E. G. Responses in man evoked by sine wave modulated light* [Thesis]. Universiteit van Amsterdam.

Spekreijse, H., & Reits, D. (1982). Sequential analysis of the visual evoked potential system in man: Nonlinear analysis of a sandwich system. *Annals of the New York Academy of Sciences*, *388*, 72–97. https://doi.org/10.1111/j.1749-6632.1982.tb50785.x

Spekreijse, H., Van der Tweel, L. H., & Zuidema, Th. (1973). Contrast evoked responses in man. *Vision Research*, *13*(8), 1577–1601. https://doi.org/10.1016/0042-6989(73)90016-3

Sperandio, I., & Chouinard, P. A. (2015). *The Mechanisms of Size Constancy*. https://doi.org/10.1163/22134808-00002483

Sperandio, I., Chouinard, P. A., & Goodale, M. A. (2012). Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. *Nature Neuroscience*, *15*(4), 540–542. https://doi.org/10.1038/nn.3069

Srinivasan, R., Bibi, F. A., & Nunez, P. L. (2006). Steady-State Visual Evoked Potentials: Distributed Local Sources and Wave-Like Dynamics Are Sensitive to Flicker Frequency. *Brain Topography*, *18*(3), 167–187. https://doi.org/10.1007/s10548-006-0267-4

Stefanucci, J. K., Creem-Regehr, S. H., Thompson, W. B., Lessard, D. A., & Geuss, M. N. (2015). Evaluating the accuracy of size perception on screen-based displays: Displayed objects appear smaller than real objects. *Journal of Experimental Psychology. Applied*, *21*(3), 215–223. https://doi.org/10.1037/xap0000051

Suchow, J. W., & Alvarez, G. A. (2011). Motion Silences Awareness of Visual Change. *Current Biology*, *21*(2), 140–143. https://doi.org/10.1016/j.cub.2010.12.019

Sugarman, A., McGlinchey, R. E., & Fortenbaugh, F. C. (2021). Attentional Capture from Looming Alters Perception. *Visual Cognition*, *29*(2), 118–124. https://doi.org/10.1080/13506285.2021.1874583

Tang, Z., Liu, X., Huo, H., Tang, M., Liu, T., Wu, Z., Qiao, X., Chen, D., An, R., Dong, Y., Fan, L., Wang, J., Du, X., & Fan, Y. (2022). The role of low-frequency oscillations in three-dimensional perception with depth cues in virtual reality. *NeuroImage*, *257*, 119328. https://doi.org/10.1016/j.neuroimage.2022.119328

Tass, P. (1995). Cortical pattern formation during visual hallucinations. *Journal of Biological Physics*, *21*(3), 177–210. https://doi.org/10.1007/BF00712345

Tauscher, J.-P., Schottky, F. W., Grogorick, S., Bittner, P. M., Mustafa, M., & Magnor, M. (2019). Immersive EEG: Evaluating Electroencephalography in Virtual Reality. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 1794–1800. https://doi.org/10.1109/VR.2019.8797858

Taylor, M., & McCulloch, D. (1992). Visual Evoked Potentials in Infants and Children. *Journal of Clinical Neurophysiology : Official Publication of the American Electroencephalographic Society*, *9*, 357–372. https://doi.org/10.1097/00004691-199207010-00004

Teghtsoonian, M. (1974). The Doubtful Phenomenon of Over-Constancy. In H. R. Moskowitz, B. Scharf, & J. C. Stevens (Eds), *Sensation and Measurement: Papers in Honor of S. S. Stevens* (pp. 411–420). Springer Netherlands. https://doi.org/10.1007/978-94-010-2245-3_40

Teng, F., Chen, Y., Choong, A. M., Gustafson, S., Reichley, C., Lawhead, P., & Waddell, D. (2011). Square or Sine: Finding a Waveform with High Success Rate of Eliciting SSVEP. *Computational Intelligence and Neuroscience*, *2011*(1), 364385. https://doi.org/10.1155/2011/364385

Thieu, M. K., Ayzenberg, V., Lourenco, S. F., & Kragel, P. A. (2024). Visual looming is a primitive for human emotion. *bioRxiv*, 2023.08.29.555380. https://doi.org/10.1101/2023.08.29.555380

Thompson, W., Willemsen, P., Gooch, A., Creem-Regehr, S., Loomis, J., & Beall, A. (2004). Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence*, *13*, 560–571. https://doi.org/10.1162/1054746042545292

Tian, P., Xu, G., Han, C., Du, C., Li, H., Chen, R., Xie, J., Wang, J., Jiang, H., Guo, X., Zhang, S., & Wu, Q. (2024). A subjective and objective fusion visual fatigue assessment system for different hardware and software parameters in SSVEP-based BCI applications. *Scientific Reports*, *14*(1), 27872. https://doi.org/10.1038/s41598-024-79401-3

Tyler, C. W., Apkarian, P., Levi, D. M., & Nakayama, K. (1979). Rapid assessment of visual function: An electronic sweep technique for the pattern visual evoked potential. *Investigative Ophthalmology & Visual Science*, *18*(7), 703–713.

Unreal Engine 4 (Version 4.24). (2019). [Computer software]. Epic Games.

Vagnoni, E., Lourenco, S. F., & Longo, M. R. (2015). Threat modulates neural responses to looming visual stimuli. *The European Journal of Neuroscience*, *42*(5), 2190–2202. https://doi.org/10.1111/ejn.12998

van der Tweel, L. H., & Spekreijse, H. (1966). Visual Evoked Responses. *The Clinical Value of Electroretinography*, 83–94.

van der Tweel, L. H., & Verduyn Lunel, H. F. E. (1965). Human visual responses to sinusoidally modulated light. *Electroencephalography and Clinical Neurophysiology*, *18*(6), 587–598. https://doi.org/10.1016/0013-4694(65)90076-3

Vanegas, M. I., Blangero, A., & Kelly, S. P. (2013). Exploiting individual primary visual cortex geometry to boost steady state visual evoked potentials. *Journal of Neural Engineering*, *10*(3), 036003. https://doi.org/10.1088/1741-2560/10/3/036003

Vanegas, M. I., Blangero, A., & Kelly, S. P. (2015). Electrophysiological indices of surround suppression in humans. *Journal of Neurophysiology*, *113*(4), 1100–1109. https://doi.org/10.1152/jn.00774.2014

Vaziri, K., Liu, P., Aseeri, S., & Interrante, V. (2017). Impact of visual and experiential realism on distance perception in VR using a custom video see-through system. *Proceedings of the ACM Symposium on Applied Perception*, 1–8. https://doi.org/10.1145/3119881.3119892

Vighneshvel, T., & Arun, S. P. (2015). Coding of relative size in monkey inferotemporal cortex. *Journal of Neurophysiology*, *113*(7), 2173–2179. https://doi.org/10.1152/jn.00907.2014

Wacker, M., Galicki, M., Putsche, P., Milde, T., Schwab, K., Haueisen, J., Ligges, C., & Witte, H. (2011). A Time-Variant Processing Approach for the Analysis of Alpha and Gamma MEG Oscillations During Flicker Stimulus Generated Entrainment. *IEEE Transactions on Biomedical Engineering*, *58*(11), 3069–3077. IEEE Transactions on Biomedical Engineering. https://doi.org/10.1109/TBME.2011.2160640

Wagner, M. (2012). Sensory and cognitive explanations for a century of size constancy research. In *Visual experience: Sensation, cognition, and constancy* (pp. 63–86). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199597277.003.0004

Walter, W. G., Dovey, V. J., & Shipton, H. (1946). Analysis of the Electrical Response of the Human Cortex to Photic Stimulation. *Nature*, *158*(4016), 540–541. https://doi.org/10.1038/158540a0

Wang, A., Chen, L., & Jiang, Y. (2021). Anodal Occipital Transcranial Direct Current Stimulation Enhances Perceived Visual Size Illusions. *Journal of Cognitive Neuroscience*, *33*(3), 528–535. https://doi.org/10.1162/jocn_a_01664

Weber, D., Hertweck, S., Alwanni, H., Fiederer, L. D. J., Wang, X., Unruh, F., Fischbach, M., Latoschik, M. E., & Ball, T. (2021). A Structured Approach to Test the Signal Quality of Electroencephalography Measurements During Use of Head-Mounted

Displays for Virtual Reality Applications. *Frontiers in Neuroscience*, *15*, 733673. https://doi.org/10.3389/fnins.2021.733673

Weidner, R., Boers, F., Mathiak, K., Dammers, J., & Fink, G. R. (2010). The Temporal Dynamics of the Müller-Lyer Illusion. *Cerebral Cortex*, *20*(7), 1586–1595. https://doi.org/10.1093/cercor/bhp217

Weidner, R., & Fink, G. R. (2007). The neural mechanisms underlying the Müller-Lyer illusion and its interaction with visuospatial judgments. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(4), 878–884. https://doi.org/10.1093/cercor/bhk042

Wertheimer, M. (1955). Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B., & Wapner, S. Personality Through Perception: An Experimental and Clinical Study. (Book Review). *Journal of General Psychology*, *52*, 181–184.

Whitney, D., Goltz, H. C., Thomas, C. G., Gati, J. S., Menon, R. S., & Goodale, M. A. (2003). Flexible Retinotopy: Motion-Dependent Position Coding in the Visual Cortex. *Science*, *302*(5646), 878–881. https://doi.org/10.1126/science.1087839

Witmer, B. G., & Sadowski, W. J. (1998). Nonvisually Guided Locomotion to a Previously Viewed Target in Real and Virtual Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(3), 478–488. https://doi.org/10.1518/001872098779591340

Wittevrongel, B., Khachatryan, E., Fahimi Hnazaee, M., Carrette, E., De Taeye, L., Meurs, A., Boon, P., Van Roost, D., & Van Hulle, M. M. (2018). Representation of steady-state visual evoked potentials elicited by luminance flicker in human occipital cortex: An electrocorticography study. *NeuroImage*, *175*, 315–326. https://doi.org/10.1016/j.neuroimage.2018.04.006

Wu, Z., & Su, S. (2014). A dynamic selection method for reference electrode in SSVEP-based BCI. *PloS One*, *9*(8), e104248. https://doi.org/10.1371/journal.pone.0104248

Yeh, L.-C., Gayet, S., Kaiser, D., & Peelen, M. V. (2024). *The neural time course of size constancy in natural scenes*. Neuroscience. https://doi.org/10.1101/2024.09.26.615104

Zemon, V., & Ratliff, F. (1984). Intermodulation components of the visual evoked potential: Responses to lateral and superimposed stimuli. *Biological Cybernetics*, *50*(6), 401–408. https://doi.org/10.1007/BF00335197

Zeng, H., Fink, G. R., & Weidner, R. (2020). Visual Size Processing in Early Visual Cortex Follows Lateral Occipital Cortex Involvement. *Journal of Neuroscience*, *40*(22), 4410–4417. https://doi.org/10.1523/JNEUROSCI.2437-19.2020

Zhang, S., & Chen, X. (2022). Effect of background luminance of visual stimulus on elicited steady-state visual evoked potentials. *Brain Science Advances*, *8*(1), 50–56. https://doi.org/10.26599/BSA.2022.9050006

Zhang, Y., Valsecchi, M., Gegenfurtner, K. R., & Chen, J. (2023a). Laplacian reference is optimal for steady-state visual-evoked potentials. *Journal of Neurophysiology*, *130*(3), 557–568. https://doi.org/10.1152/jn.00469.2022

Zhang, Y., Valsecchi, M., Gegenfurtner, K. R., & Chen, J. (2023b). The time course of chromatic adaptation in human early visual cortex revealed by SSVEPs. *Journal of Vision*, *23*(5), 17. https://doi.org/10.1167/jov.23.5.17

Zhang, Y., Zhang, L., Hua, H., Jin, J., Zhu, L., Shu, L., Xu, X., Kuang, F., & Liu, Y. (2021). Relaxation Degree Analysis Using Frontal Electroencephalogram Under Virtual Reality Relaxation Scenes. *Frontiers in Neuroscience*, *15*. https://doi.org/10.3389/fnins.2021.719869

Zhu, S., Yang, J., Ding, P., Wang, F., Gong, A., & Fu, Y. (2023). Optimization of SSVEP-BCI Virtual Reality Stereo Stimulation Parameters Based on Knowledge Graph. *Brain Sciences*, *13*(5), 710. https://doi.org/10.3390/brainsci13050710

Appendix A: Time-Frequency Optimization for rapidly changing SSVEPs

The heart of the SSVEP technique lies in extracting the stimulation frequency from the EEG data. EEG data is a collection of continuously sampled electrical potentials at the site of the EEG electrodes, resulting in a set of signals that varies in amplitude over time. The data in this form is known as being in the time domain, which is also the form that the data takes when used for studies into event related potentials or many clinical applications. For the study of SSVEPs, to examine the relative strength of the stimulation frequency in this overall signal, the data need to be transformed from the time domain to the frequency domain. This is most commonly done using a Fast Fourier transformation (FFT). Since the work presented in this thesis aims to investigate SSVEP signals evoked by dynamic stimuli, we will be put in the uncomfortable position of making tradeoffs between having a fine resolution in the time domain and having a fine resolution in the frequency domain. This section discusses some of the considerations involved in this tradeoff.

An inherent quality of this transformation is that the resolution of the resultant frequency spectrum is a function of the length of the time window. More data in the time domain leads to more data in the frequency domain. This can be expressed mathematically by the formula $df = \frac{1}{T}$, where *df* is the frequency resolution, and *T* is the length of the time window. Finer frequency resolutions (as expressed by a smaller *df*) provide more fine grained information of the frequency components that make up a signal. A frequency spectrum based on a window of length T will contain bins of all frequencies between 0 and the Nyquist frequency in steps of *df*.

Even though the length of one of these windows can be scaled up on a continuous scale to increase frequency resolution in theory, for its use with SSVEPs it may be preferable to use discrete steps. Since SSVEPs rely on stimulation at one or several specific frequencies, the time window should ideally be of such a length that these specific frequencies are represented in the resultant spectrum. For instance, when using a stimulation frequency of 7.5Hz, the shortest time window that, when transformed to the frequency domain, would contain this stimulation frequency is 133ms, which is equal to the length of a single 7.5Hz cycle. The resultant *df* ($\frac{1}{0.133}$) would be approximately 7.5Hz and the frequency spectrum would therefore contain the stimulation frequency and all its consequent harmonics. A window of 150ms, although longer and therefore having a smaller *df*, would contain bins from

0 to the Nyquist frequency in steps of 6.66Hz, thereby 'missing' the stimulation frequency and its harmonics. It should be noted that due to spectral leakage in these short time windows, effects of the stimulation frequency may be observed in the 6.66Hz bin when using a 150ms time window, though this would not be an ideal approach.

We can illustrate this phenomenon by simulating some data. We shall imitate EEG data with an inverse fourier transformation of a 1/f pink noise function, and combine this with a simulated stimulation frequency of 7.5Hz. The signal is 10 seconds long and is sampled with a rate of 500hz. In these data, the stimulation frequency varies in amplitude over time to mimic an SSVEP that varies in strength over time. We will further add harmonics to this imitated SSVEP to simulate nonlinear brain responses. These harmonics are commonly included in the extracted SSVEP value (Regan, 1989). These data in both time and frequency domain are represented in figure A1.

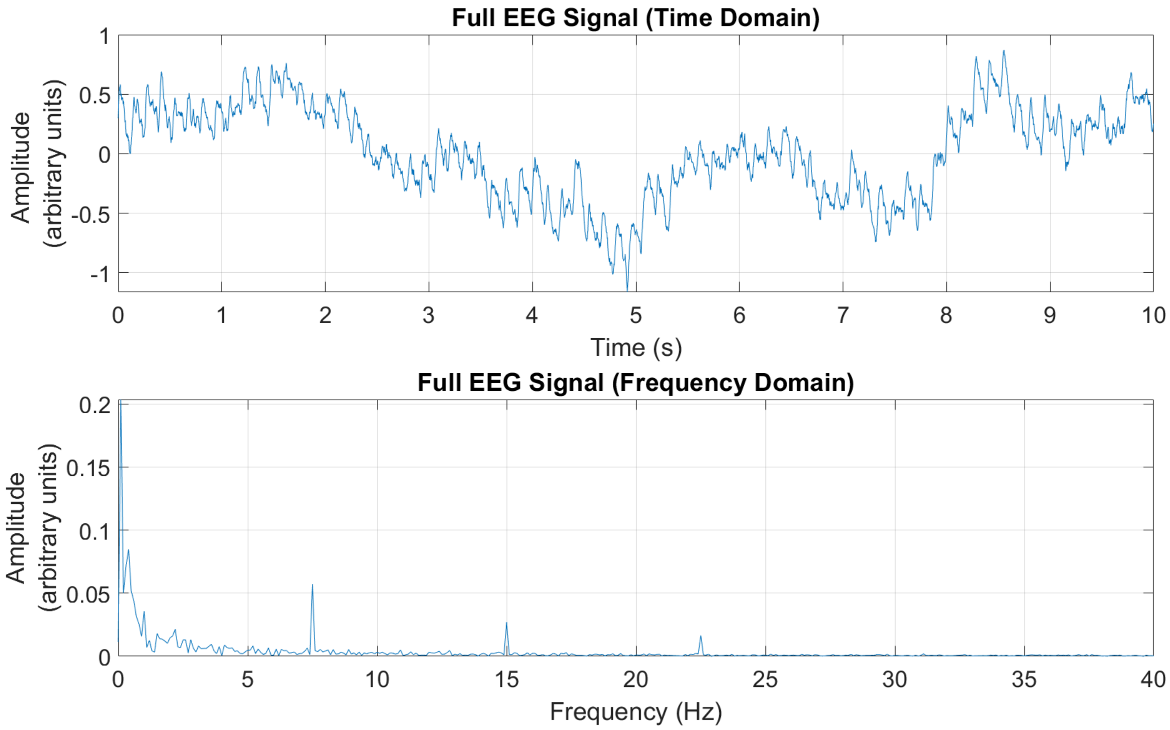


Figure A1: Simulated data in time and frequency domains

If we then treat this data like we would actual EEG data and attempt to extract the stimulation frequency by means of a moving-window FFT, we can observe the effect of the window width (fig. A2). The moving-window technique involves taking a discrete amount of time-domain data, converting this to the frequency domain, extracting the stimulation frequency from it and consequently moving this window along by one step. This allows us to track the amplitude of a certain frequency in a complex signal over time. In Figure A2, the amplitude of the stimulation frequency of our simulated data is pictured in red, and takes the shape of a triangular wave. The value of the extracted signal is pictured in blue, and is noticeably noisier than the actual amplitude of the signal as it is present in the original data. Using a 100ms window (fig A2.A), the extracted signal is extremely noisy, and the signal-to-noise ratio (SNR) is noticeably improved when the window length is increased to 133ms (fig A2.B). However, when the window is further increased to 150ms, the SNR decreases again and the signal is visibly noisier again (fig A2.C). N.B. The stimulation frequency in the simulated data is more pronounced than might be expected in real data for illustration purposes, making the overall amplitude modulation clear to see in all window length conditions. Figure A2.D shows boxplots of SNRs derived from multiple (50) simulated SSVEP extractions using windows of various lengths. This shows the relation between SNR of the extracted signal and time window length more generally, and illustrates the value of using window lengths that are multiples of the cycle, and when transformed to the frequency domain contain the exact stimulation frequency. A window length of 133ms as well any integer multiplication of this length (e.g. 266 as shown in figure A2.D) will contain the stimulation frequency exactly. Although it is not generally recommended to use different parameters in the EEG data analysis of a single experiment, in experiments that use two or more stimulation frequencies concurrently, analyses may also use different time windows to analyse the resultant SSVEPs (Müller, Teder-Sälejärvi & Hillyard, 1998).

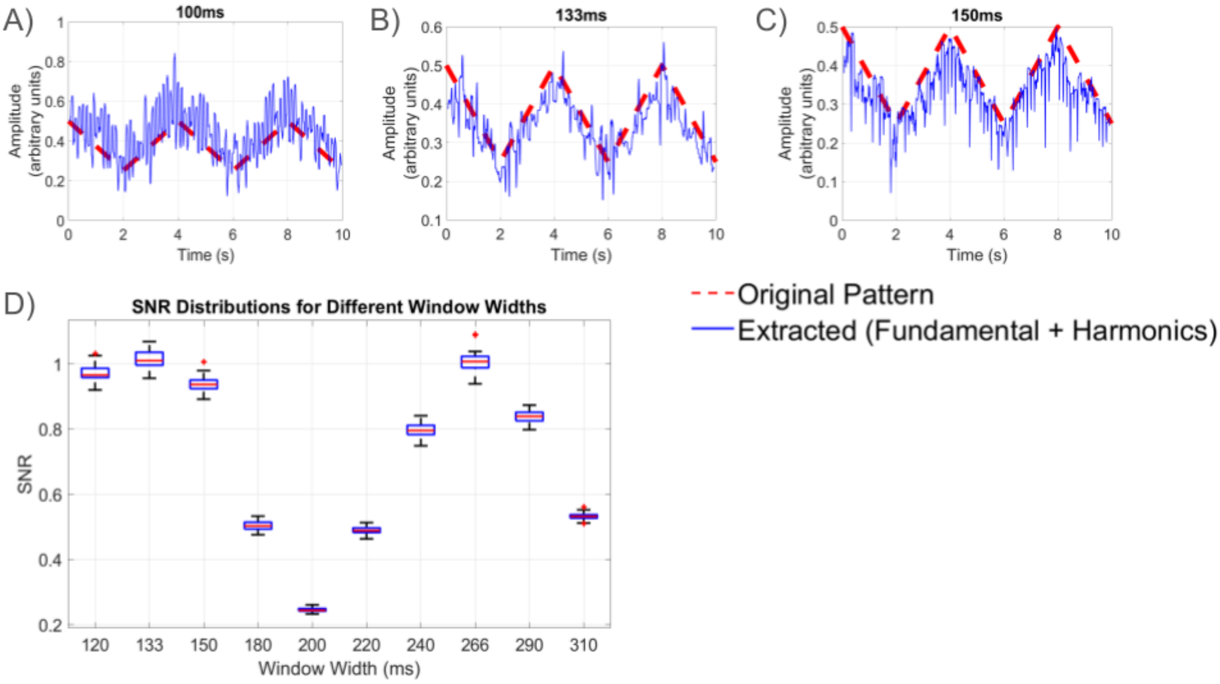


Figure A2: Extracted SSVEP from simulated data using varying time windows. A-C show window widths of 100ms, 133ms and 150ms respectively. Amplitude on the y axis is displayed in arbitrary units. D Shows signal-to-noise ratios (y-axis) for various window widths (x-axis)

The use of short time windows and the moving window analysis technique more generally may well be required to successfully study the neural dynamics of size constancy in dynamic environments. Previous studies have demonstrated the very rapid nature with which size constancy seems to be established at a neural level. Chen, Sperandio, Henry & Goodale (2019) used EEG to show that the integration of retinal size and depth cues can be observed at 150ms after stimulus onset. Yeh, Gayet, Kaiser & Peelen (2024) further show that distance modulation of size information of an object that is already present in the visual field can be decoded from EEG data 120ms after retinal size can be decoded. Yeh et al.'s findings are more indicative of the time course that is relevant for our purposes, because they refer to the time it takes for visual information already present in the visual system to be modulated by novel size and or distance information. More circumstantially, research using MEG has shown the effect of certain size illusions to emerge as early as 80ms after stimulus onset (Weidner, Boers, Mathiak, Dammers & Fink, 2010). This suggests that some degree of size perception that is not dependent on retinal size alone can occur as quickly as 80ms post stimulus onset. These findings suggest that the adjustment of the neural signature of the size of an object occurs exceedingly rapidly. Even though there has not been any experimental examination of the neural representation of size when an object actively moves in a depth cue rich environment, the above mentioned findings suggest that adjustments may occur within 100ms. If we wish to capture not just the overall size representation before or after a movement, but the nature of this change, we will need to analyse the data in such a way as to obtain as fine a temporal resolution as possible.

Two parameters of importance in capturing fast changes in stimulation frequency amplitude are: window width, the size of the temporal window used to extract frequency components from, and step size, the distance between the first data point of window n, and the first data point of window n+1 (figure A3). Window width, like discussed above, indicates the amount of time domain data used to transform into frequency domain data. An SSVEP value will be extracted from each window, and this value indicates the degree to which periodic components of the stimulation frequency make up the time domain data in this window. This value does not indicate how much, if at all, this frequency’s contribution to the overall signal has changed over the course of the window. Analyses using wider windows are therefore less sensitive to rapid changes in stimulation frequency amplitude. This effect can be partially counteracted by employing short step sizes. As illustrated in figure A3, short step sizes lead to a large amount of overlap between two adjacent analysis windows. Due to this overlap, resultant SSVEPs will exhibit smoother transitions over time, reflective of changes in the underlying stimulation frequency amplitude.

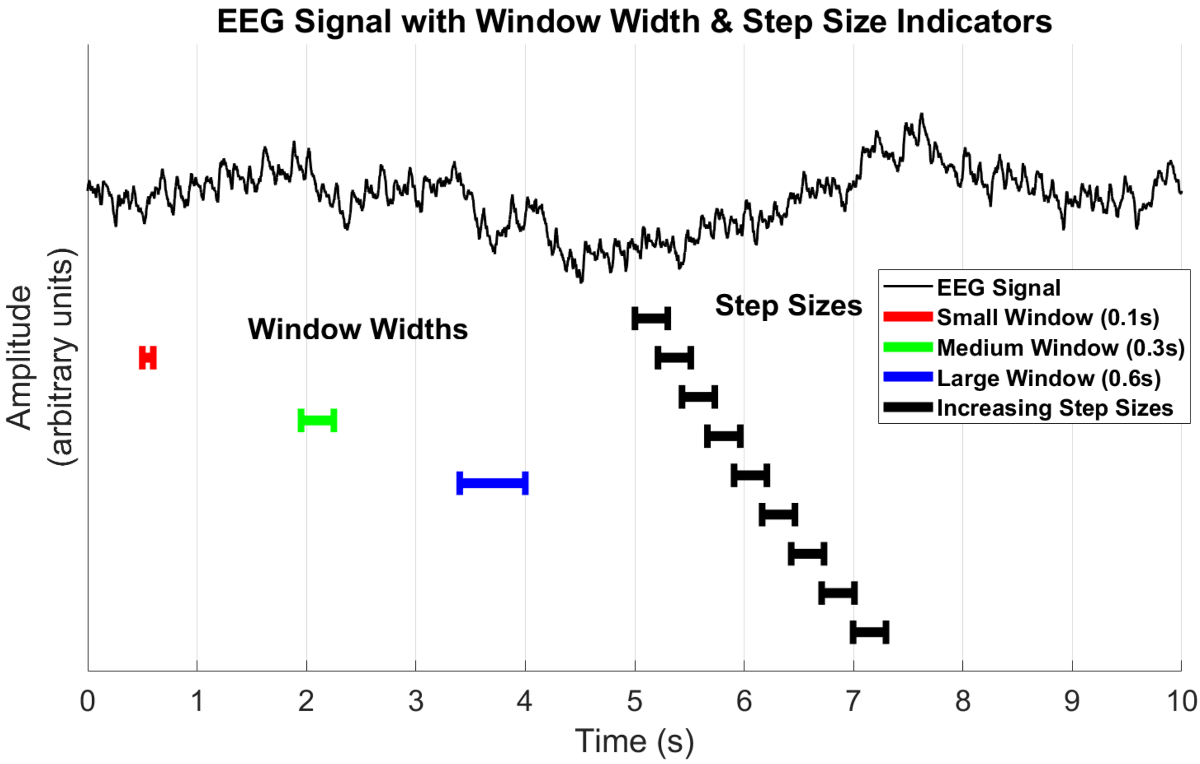


Figure A3: Model EEG data with analysis windows of varying widths and varying step sizes.

We can illustrate what parameter values would work best for examining quick SSVEP strength changes by again turning to modeled data. We shall again create simulated EEG data with the same parameters as before, but change the amplitude of the stimulation frequency. Instead of the stimulation frequency changing its amplitude in a triangle wave as before, it will have a low amplitude during the first half of the signal, and a higher amplitude in the second half. During the transition from low to high, an additional quick enhancement occurs, mimicking something along the lines of a rapid ‘neurological zoom’ in early visual cortex (Marg & Adams, 1970). We will have this quick change occur over the course of 100ms.

Figure A4 illustrates the effect of window width on the pattern of the extracted SSVEP. In figure A4.A, each data point from the extracted pattern represents the stimulation frequency extracted from a 133ms (1 stimulus cycle) window, which is then moved forward in steps of 10ms. This results in a generally noisy signal, as indicated by the variance in the extracted signal during periods when the actual amplitude of the stimulation frequency is constant. Despite this variance, a shorter window is able to capture events in a very short time scale. Longer windows on the other hand, like ones of 5 cycles (fig A4.B) or 10 cycles (fig A4.C) may well fail to properly reflect sudden, sharp changes in signal amplitude. Wider windows will further have the tendency to ‘smear out’ changes in amplitude over time, potentially creating the impression that shifts in neural activation occur slower than they really do. This decrease in temporal resolution is of course counterbalanced by an increase in frequency resolution, and will result in cleaner data when the SSVEP is in fact constant.

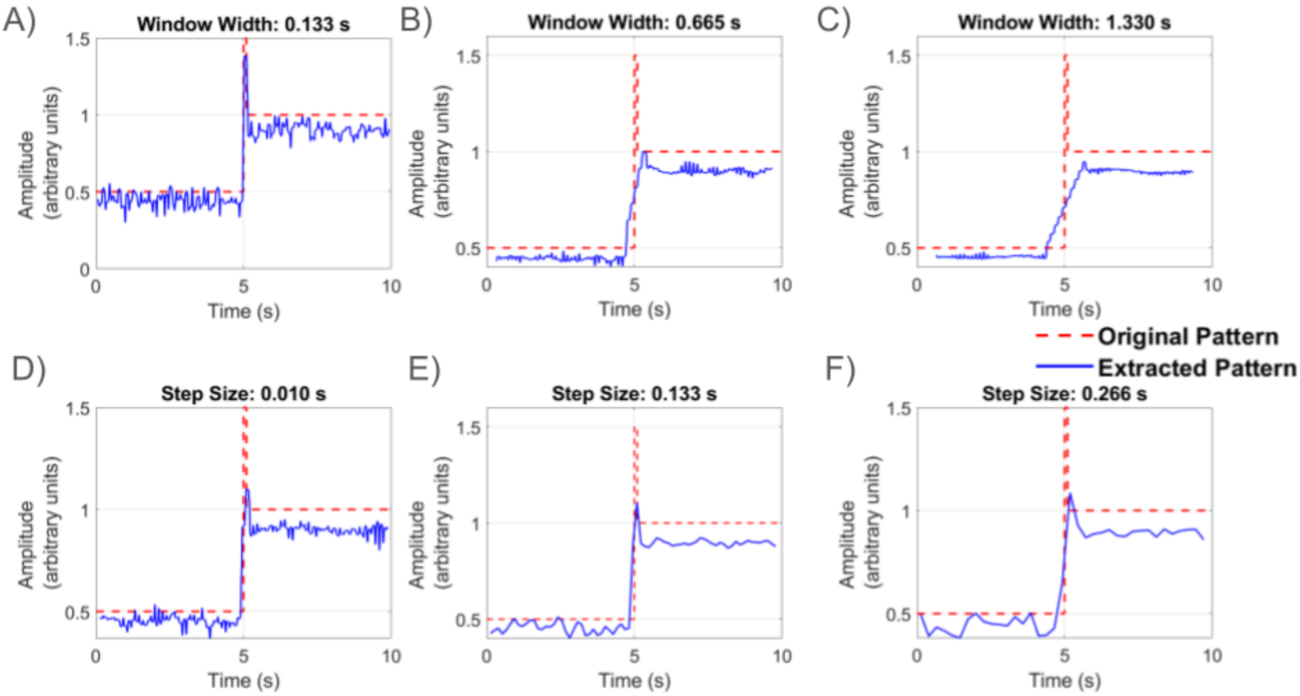


Figure A4: illustrations of the effect of window width and step size on extracted signals

The bottom row of graphs in figure A4 shows the effect of varying step sizes. Figure A4.D shows the SSVEP being extracted using a time window of 266ms and the smallest step size possible for data recorded at 500Hz; 2ms. This means that each consecutive time window from which an SSVEP is calculated only differs by one data point, resulting in the most fine grained timeline, temporally speaking. It also results in a signal that contains a great deal of redundancy, with adjacent SSVEP values being extracted from near-identical data. If we instead use a step size of 266ms, which in this case is identical to the window width and would mean no overlap of analysis windows at all, we again see some degree of temporal smearing, which make sudden, fast changes in SSVEPs harder to spot, and creating the impression that changes in SSVEP strength occur slower than they really do. Step sizes can, in theory, exceed the length of the time window, though the use of these is never recommended. Using a step size larger than the time window will result in EEG data that is not used at all in the analysis of the SSVEP, which results in loss of information from the original signal.

Overall, for the detection of rapid changes in SSVEP strength, narrow windows and small step sizes are recommended. It bears repeating that narrow windows are more susceptible to noise, and a certain level of data quality is required for an analysis like this. This is due to spectral leakage from surrounding frequencies being more pronounced in low resolution frequency bins, as well as the decreased effectiveness of correcting the stimulation frequency using the surrounding frequencies. Still, experiments where rapid changes in SSVEP strength are relevant have used window widths of a single cycle and step sizes of a single data point (e.g. Belmonte, 1998). There are fewer downsides to having smaller step sizes, as these do not inherently increase the noise present in the data, but just make it more salient. It is possible to employ step sizes of a single data point even if the window sizes are 10 cycles or even 10 seconds (Silberstein, Schier, Pipingas, Ciorciari, Wood, & Simpson, 1990; Gray, Kemp, & Silberstein, 2003). However, if the goal is to describe the time course of SSVEPs over the course of several minutes, it may just not be worthwhile to use step sizes less than several seconds (Zhang, Valsecchi, Gegenfurtner & Chen, 2023).

One way of enhancing the SNR in SSVEP data is to adjust the spectral power of the stimulation frequency bin by subtracting the power of the surrounding frequency bins from it (Meigen & Bach, 1999). The rationale behind this is that the extracted SSVEP is reflective of spectral power that is directly induced by the stimulus, as well as endogenous brain activity that happens to occur at this same frequency. There are large physiological correlations between adjacent frequencies in EEG data, often expressed in terms of frequency bands (Gable & Miller, 2022). It is therefore reasonable to assume that subtracting close neighboring frequencies will remove the endogenous brain activity present at the stimulation frequency, and leave just the added activity of the SSVEP. Although useful when using large window widths, this technique is of limited use when using shorter time windows. For the shortest usable window width of one cycle, the neighbouring frequencies will be integer multiplications of the stimulation frequency (harmonics of the stimulation frequency), and therefore contain valuable information of the SSVEP strength. Despite this window width being something of an exception, the principle that subtracting surrounding frequency bins from the stimulation frequency is more useful for longer window widths holds true in general. In short time windows the frequency resolution will be low, leading to larger steps in frequency between adjacent bins. For instance, in a window of 266ms for a stimulation frequency of 7.5Hz, the frequency resolution will be 3.75Hz. The neighboring frequencies to the stimulation frequency will therefore be 3.75Hz and 11.25Hz, which are both in a different frequency band to the stimulation frequency. The stimulation frequency falls in the theta band, while its neighbors fall in the delta and alpha bands respectively. Subtracting these neighboring frequencies will therefore not greatly improve the SNR, since their activity will only be tangentially related to the endogenous activity at the stimulation frequency. Correcting SSVEPs for neighboring frequencies is therefore only recommended for sufficiently large time windows, in the range of several seconds.

Overall, the choice of window width and step size plays a crucial role in determining the effectiveness of an SSVEP analysis, particularly when studying rapid changes in stimulation frequency amplitude. Shorter windows provide finer temporal resolution but are more susceptible to noise, while longer windows improve frequency resolution at the cost of temporal precision. The right balance between these factors depends on the specific goals of the experiment. Particularly in the first series of experiments performed in this project, the emphasis was placed on studying the SSVEP with as fine grained a temporal resolution as possible, and very narrow windows combined with small step sizes were mainly used.