MONASH University

Spatial and temporal analysis of qualia as evidence of conscious perception

Matthew James Davidson

Bachelor of Psychology (Hons 1)

A thesis submitted for the degree of Doctor of Philosophy (PhD)

Monash University, 2019

Department of Psychological Sciences

Copyright notice

In any field, find the strangest thing and then explore it.

John Archibald Wheeler.

ABSTRACT

Although seemingly effortless, our conscious experience is formed from a complex stream of incoming sensory stimulation. The consciousness that we are each intimately familiar with becomes realized, due to processes that *lie outside* of our conscious awareness. Not only are these processes resistant to introspection, but current opinion is also divided over whether we should, and how best we could, empirically dissociate the neural correlates of consciousness, from the overlapping focus of our selective attention. The primary aim of this thesis was to aid these issues by investigating how selective attention contributes to our conscious experience, using the application of a technique known as frequency-tagging. During frequency-tagging, sensory stimuli are rapidly flickered at a particular frequency in order to entrain brain responses recorded in the electroencephalogram (EEG). By leveraging the empirical advantages of frequency-tagging, this thesis explores how consciousness emerges from the interplay between sensory stimulation and selective attention, along two complementary axes. In the first axis I apply frequency-tagging to binocular rivalry, during which two visual stimuli compete for perceptual dominance. I continue by showing that the contents of visual awareness during binocular rivalry are biased by auditory and tactile stimulation, and critically depend upon flicker frequency as well as selective attention. Investigating this interaction further reveals that the neural correlates of sustained attention are inherently rhythmic, with novel implications for the targeted relationship between consciousness and attention. In the second empirical axis I characterize changes in perception during two novel visual illusions, each tailored to enable the disappearance of multiple visual stimuli from conscious awareness. To begin I capture the graded changes in neural activity which occur when multiple flickering stimuli disappear from the visual periphery. I then refine this paradigm and provide a neural marker to distinguish between the effects of attention and current contents of conscious perception using frequency-tagging. Taken together, this thesis represents a snapshot of the benefits which are gained through the combination of frequency-tagging with investigations into the dissociation between attention and conscious perception.

Publications, presentations and awards during candidature

The empirical chapters included in this thesis have been published (**Chapter 4, Appendix 1**), submitted (**Chapter 5**), or accepted as abstracts for conference presentations after peer-review (contributions to **Chapters 3, 4, 5, 6**). All research output has been slightly adapted to increase consistency and readability between chapters.

Manuscripts

Paper	Status	Chapter
Davidson, M., Alais, D., van Boxtel, J., Tsuchiya, N. (2018). Attention periodically samples competing stimuli during binocular rivalry, *eLIFE, 7,* doi.org/10.7554/eLife.40868	Published	4
Davidson, M., Graafsma, I., Tsuchiya, N., & van Boxtel., J.J.A. (preprint). Frequency-tagging visual background information enables multi-target perceptual filling-in to be distinguished from phenomenally matched replay. *bioRxiv,* doi.org/10.1101/499517	Submitted	5
Thomas, V., **Davidson, M**., Zakavi, P., Tsuchiya, N., & van Boxtel, J. (2017). Simulated forward and backward self-motion, based on realistic parameters, causes motion induced blindness. *Scientific reports, 7*(1), 9767,	Published*	Appendix 1

* Denotes work that was initiated and undertaken during enrolment, with time taken to prepare the manuscript for publication. This manuscript has not been included in this thesis due to the conditions of the award (1st author V. Thomas submitted this manuscript for her Honours thesis at Monash University).

Peer-reviewed conference presentations

1. **Davidson, M.**, Graafsma, I., Tsuchiya, N., & van Boxtel, J.J.A. (2018). *Investigating perceptual filling-in using steady-state visually evoked potentials,* Invited presentation for the Australasian Cognitive Neuroscience Society (ACNS). November 2018, Melbourne, Australia.
2. **Davidson, M.**, Graafsma, I., van Boxtel, J.J.A., & Tsuchiya, N. (2018). *Sensation and attention differentiated by first and second harmonics during perceptual filling-in,* Invited presentation for the Association for the Scientific Study of Consciousness (ASSC), June 2018, Krakow, Poland.
3. **Davidson, M**., Alais, D., van Boxtel, J., & Tsuchiya, N. (2017). *The neural correlates of cross-modally facilitated transitions from nonconscious to conscious percepts during binocular rivalry.* Invited presentation for the Association for the Scientific Study of Consciousness (ASSC). June 2017, Beijing, China.
4. Graafsma, I., **Davidson, M.,** Tsuchiya, N., van Boxtel, J. (2016). *Neural correlates of filling-in using steady state visually evoked potentials (SSVEPs)*, Poster for Australasian Cognitive Neuroscience Society (ACNS), November 2016, Newcastle, Australia. Abstract in Karayanidis, F. & Todd, J. (Eds). 6th Australasian Cognitive Neuroscience Society (ACNS) Conference Proceedings. Published by The University of Newcastle, Australia. ISBN: 978-0-648-00036-5.
5. Thomas, V., **Davidson, M.,** van Boxtel, J., Tsuchiya, N. (2015). *Simulated forward and backward self-motion, based on realistic parameters, causes motion induced blindness,* December 2015, Melbourne, Australia. Poster for the Australasian Society for Experimental Psychology (EPC), Melbourne, March 2016. Abstract published online in conference proceedings.

Awards

1. **28/02/2019:** Monash University Post-Publication Award. Financial support for high-achieving PhD graduates to prepare their research for publication while awaiting the result of their examination *(AUD $4728)*
2. **09/03/2015 – 30/11/2018:** Australian postgraduate award (APA). Annual stipend to continue research training throughout the candidature of Doctor of Philosophy, Monash University. *(AUD $27,082 per annum)*

3. **23/03/2018:** Marie-Curie Master Class at Aarhus University, Health Department. Travel and accommodation grant to attend Aarhus University for the purposes of developing and refining a grant application *(AUD $3000)*
4. **04/05/2017:** Monash University Research Travel Grant. Awarded to present research at the annual Association for the Scientific Study of Consciousness (ASSC), June, Beijing, China *(AUD $1200)*
5. **15/06/2017:** Monash University Research Travel Grant. Awarded to present research at the triennial International Conference for Cognitive Neuroscience (ICON), August, Amsterdam, Netherlands. *(AUD $1200)*

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes **1** original paper published in a peer reviewed journal and **1** submitted publication. The core theme of the thesis is the ***neuroscience of consciousness***. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the ***School of Psychological Sciences*** under the supervision of ***Jeroen van Boxtel*** and ***Naotsugu Tsuchiya***.

(The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.)

In the case of all chapters my contribution to the work involved the following:

4	Attention periodically samples competing stimuli during binocular rivalry	Published	1,2,3,4,5,6, 7,8,9. **70%**	David Alais: 1,5,8,9,10, **(10%)** Jeroen van Boxtel and Naotsugu Tsuchiya: 1,5,6,10, 11,12, **(20%)**	No
5	Measuring graded changes in consciousness through multi-target filling-in	Submitted	1,3,4,5,6,7, 8,9,11 **70%**	Irene Graafsma: 1,2,4,8, **(10%)**: Jeroen van Boxtel and Naotsugu Tsuchiya: 1,5,6,10, 11,12, (**20%)**:	No
Appendix	Simulated forward and backward motion, based on realistic self-parameters, causes motion induced blindness	Published	1,8,9,11 **5%**	Victoria Thomas: 1,2,3,4,5,6,7,8, **(70%)** Parisa Zakavi: 4,8, **(5%)** Jeroen van Boxtel and Naotsugu Tsuchiya:1,5,6,10, 11,12 **(20%)**	Yes (V.T.)

1- Conceptualization; 2- Data curation; 3- Formal analysis; 4- Investigation; 5- Visualization; 6- Methodology; 7- Writing original draft; 8- Writing review and editing; 9- Software; 10- Resources, 11- Supervision, 12- Funding acquisition.

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: **REDACTED** **Date:** 27/02/2019

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author, I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: **REDACTED** **Date:** 26/02/2019

Acknowledgements

It has been my privilege to have Jeroen van Boxtel and Naotsugu Tsuchiya as my thesis supervisors. Throughout my candidature they have given me enormous freedom to pursue my ideas, within an extraordinarily warm and encouraging working environment.

Jeroen, the worth of your exceptionally kind approach to teaching and supervision cannot be overstated. Your support was unwavering, giving me the confidence that ideas should never go unspoken, and peace of mind to know that any shortcomings we would tackle together.

Nao, your humour and erudite knack for analyses were a joy to see in action. Your creative insights over a scribbled whiteboard will inspire me for years to come. You've taught me there is always another approach to a difficult problem, and your infectious enthusiasm for consciousness research is legendary.

I thank you both sincerely for all the guidance you have provided, and for all the intellectual contributions that were critical to this work. Thank you for putting your trust in a stranger and for shaping me into a better scientist, I hope our collaborations continue for years to come.

To my collaborators, David Alais, Irene Graafsma, and Will Mithen. Thanks for all the help in proofreading, sanity-checking, and the development of these ideas. it has been a pleasure working with each of you.

To my MoNoC/t-lab/MBI family, I would not have been able to get through this project without your solidarity, good humour, and willingness to partake in procrastination. The camping trips, soccer, and table-tennis tournaments were the perfect antidote to long stretches in the depths of my analyses.

To the members of my internal milestone and external review panels, thank you for taking the time to provide valuable feedback and recommendations for the completion of this work.

To my family, thank you for excusing my absence from the dinner table, from communication in general, and for sustaining me with love, confidence and affection throughout every step of my education. To my parents, thank you for believing in me, and for instilling a joy of knowledge and optimism in my ability for as long as I can remember.

To my beautiful wife, Jane. Thank you for moving across the world to support me throughout this endeavour. These pages would be blank without your warmth, companionship, endless love and understanding. I dedicate this thesis to you.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship.

Table of contents

List of Figures

Figure 1.1: Our unconscious mind - Tim Bower (2014).

Introduction

The problem(s) with consciousness

At once intimately familiar and frustratingly enigmatic, the study of consciousness has been hailed as one of the hardest problems of our time. Perhaps as a consequence of this difficulty, allegory and metaphor pervade public understanding of how the mind works: one hears of left-and-right hemispheric individuals using only 10% of their brains' capacity, while unconscious motives drive personal behaviour. While some of these claims can be rebuked, as a scientific frontier, the mechanisms that separate consciousness from other higher cognitive functions, as well as unconscious processes, remain elusive. Critically, there is an unmet empirical need for methods that probe the differences between consciousness and selective attention.

This unmet empirical need is the driving focus of the present thesis, and several chapters are devoted to disentangling the contents of consciousness from the overlapping focus of selective attention. It should be noted that this focus is distinct from philosophical enquiries concerning the easy and hard problems of consciousness research (Chalmers, 1995, 1996). While Chalmers and others have outlined empirical limits to explaining the phenomenal sense or 'what it is like' of a subjective state (Block, 1995, 2008, 2011; Chalmers, 1996; Nagel, 1974), here we are concerned with how best to distinguish the contents of consciousness from other overlapping phenomena.

The term 'consciousness' here will be used to define the contents of phenomenal experience rather than the level or state change that occurs during periods of sleep or anaesthesia (Laureys, 2005; Owen, 2008; Storm et al., 2017). The study of consciousness is now a hotbed of intellectual pursuit, in large part due to a renewed focus upon establishing what are regarded as the neural correlates of consciousness (NCCs; Crick & Koch, 2003; Koch, Massimini, Boly, & Tononi, 2016; Tononi & Koch, 2015). Defined as the minimal set of neural substrates that give rise to a conscious experience, the NCCs were introduced in a series of influential publications that rejuvenated the study of consciousness in the 1990s (Crick & Koch, 1990b, 1990a, 1995, 2003). The most elementary units of phenomenal experience, known as qualia, are the focus of these investigations. Qualia are defined as the

character or ‘what it is like’ of a subjective experience (Kanai & Tsuchiya, 2012; Lamme, 2010; Loorits, 2014; Tononi, 2008). For example the singular quale of red, or ‘what it is like’ to see red, is argued to have separate neural correlates from the quale of green (Kanai & Tsuchiya, 2012; Loorits, 2014). As such, contrasting neural activity between periods of experiencing one quale over another are used to investigate the NCC.

To contrast the neural activity evoked by changes in qualia, researchers have relied upon paradigms which alter perceptual experience without external change, known as multistable stimuli (Blake, Brascamp, & Heeger, 2014; Blake & Logothetis, 2002; Kim & Blake, 2005). Multistable stimuli are so named because an unchanged physical stimulus can be perceived in one of two, competing perceptual interpretations. These stimuli are used to investigate the endogenous changes in perception by asking participants to behaviourally report on the content of their subjective experience. Contrasting the neural activity evoked by these reported changes in subjective state can then localize the NCCs, as differences in brain activity must result from within, since the physical input stimulus itself remains unchanged.

Opponents to the use of multistable stimuli to investigate the NCCs cite a number of limitations which must be overcome (Bayne & Chalmers, 2003; De Graaf, Hsieh, & Sack, 2012; Hohwy, 2009; Miller, 2014). Chief among these are the necessity to detect true neural correlates of consciousness, rather than the overlapping correlates of our focused attention (Aru & Bachmann, 2015; Jack & Hacker, 2014; Koch & Tsuchiya, 2007; Tsuchiya, Wilke, Frassle, & Lamme, 2015; van Boxtel & Tsuchiya, 2014). This thesis contributes novel experimental paradigms to dissociate perceptual changes from attention, using scalp-level recordings of electroencephalographic (EEG) data.

Outline of the thesis

Chapter 1 of this thesis reviews the background and current state of the literature regarding the use of multistable paradigms to investigate the NCC. It is shown that these investigations predominantly focus on the visual system, before highlighting the contributions that multisensory investigations can make to our understanding of consciousness. Particular focus is also placed on the methodological

advantage of combining multistable paradigms with a technique known as frequency-tagging, which is used as a recurrent technique in the empirical chapters of this thesis.

These empirical chapters form two complementary axes, the first of which investigates multisensory contributions to visual experience, the second of which develops a novel multistable paradigm to leverage the advantages of frequency-tagging to dissociate neural markers of selective attention from the contents of consciousness.

These empirical chapters begin with **Chapter 3**, which builds upon the premise that consciousness is never unisensory, and that the integration of multisensory input may be one functional role that consciousness serves. The consequent investigation explores the effect that non-visual stimuli have on contents of visual consciousness. **Chapter 4** then provides evidence that attention can rhythmically sample away from a conscious visual image, in support of a distinction between the focus of attention and current contents of consciousness. Together, **Chapters 3 and 4** demonstrate the benefit of considering multisensory paradigms to investigate the NCCs, and provide evidence for the independent nature of consciousness and attention.

The second empirical axis begins with **Chapter 5**. By combining frequency-tagging with perceptual-filling in (PFI), a novel paradigm is developed to dissociate the correlates of attention and conscious contents. **Chapter 6** then refines this paradigm, and shows neural evidence in support of the notion that the contents of consciousness can be dissociated from the allocation of attention. **Chapter 7** then summarizes the empirical contributions of this thesis and contains speculations for future research.

Chapter 1: Literature Review

A primer on perception as unconscious inference

How can the quality of subjective experience be accounted for by the physical substrates of the human brain? This question has resulted in a renewed focus at the intersection of psychology and neuroscience, and forms the basis for the empirical investigations of this thesis. To understand the empirical motivations of this thesis, it is worth briefly reviewing the philosophical background that forms the backbone of contrastive analysis.

In this thesis, consciousness is defined as the contents of phenomenal experience, rather than 'what it is like' to experience a subjective state (Nagel, 1974). Although we each have an intuitive understanding of what this end result feels like, exactly how our brain performs the task of transmuting incoming sensory stimulation into the vivid experiences of our everyday lives is obscure, and a process which for the most part occurs outside of our conscious awareness. The importance and obscurity of this process has been recognized for almost 1000 years (cf. al-Haitham, 1989), though it is Von Helmholtz (1867) who is credited with its most recent incarnation. Specifically, he proposed that a process of 'unconscious inference' must take place to understand how neural activity relates to perception and how incoming sensory stimulation identifies an object in the outside world. This inference casts perception as an active process, whereby unconscious decisions have been made to determine the most likely state of the outside world on the basis of other information (Brascamp, Sterzer, Blake, & Knapen, 2018; C.-Y. Kim & Blake, 2005; Kleinschmidt, Sterzer, & Rees, 2012).

For the study of consciousness, adopting the philosophy that perception results from an active, inferential mechanism enables its building blocks to be examined. This is because although descriptions for consciousness abound, the majority of contemporary models assume that differences in perception correlate with differences in neural function or structure. As such, contrasting the spatial and temporal patterns of neural activity that occur with changes in perception can serve to localize the end-

product of this unconscious inference: the NCCs (Crick & Koch, 1990b, 2003; Koch et al., 2016).

1.1. Multistable stimuli: when the brain changes its mind

Ambiguous or multistable stimuli are unique, in that an unchanged or invariant physical stimulus is results in more than one perceptual outcome (Kim & Blake, 2005). A key feature of these perceptual outcomes is that they are competitive, and temporarily stable, resulting in ongoing fluctuations in conscious perception. Multistable perception was critical to von Helmholtz accounts of the inferential process of perception, revealing the spontaneous and constructive process of conscious experience. Today, multistable stimuli are recognized as a valuable tool for penetrating into the features that mediate the content of consciousness (Blake & Logothetis, 2002; Maier, Panagiotaropoulos, Tsuchiya, & Keliris, 2012; Panagiotaropoulos, Kapoor, & Logothetis, 2014). In particular, contrasting the neural activity which arises during each percept has been critical to studies of consciousness, as since the physical stimulus is unchanged, changes in consciousness are believed to be endogenously generated, with differences in neural activity representative of differences in consciousness itself (Miller, 2013; Sterzer, Kleinschmidt, & Rees, 2009). Popular examples of multistable stimuli include the Necker-cube, which oscillates in depth (Necker, 1832; Figure 1.2a), or examples of figure-ground ambiguity (Boring, 1930; Rubin, 1980; Figure 1.2b). Examples of multistable stimuli also occur for moving stimuli, resulting in ambiguous motion (Doner, Lappin, & Perfetto, 1984) as well as in the olfactory (Stevenson & Mahmut, 2013; Zhou & Chen, 2009), auditory (Almonte, Jirsa, Large, & Tuller, 2005; Dykstra, Cariani, & Gutschalk, 2017; Kondo et al., 2012), and tactile sensory modalities (Carter, Konkle, Wang, Hayward, & Moore, 2008). Despite the existence of multistable stimuli outside of vision, the majority of investigations have occurred in vision, and frame the scope of this review.

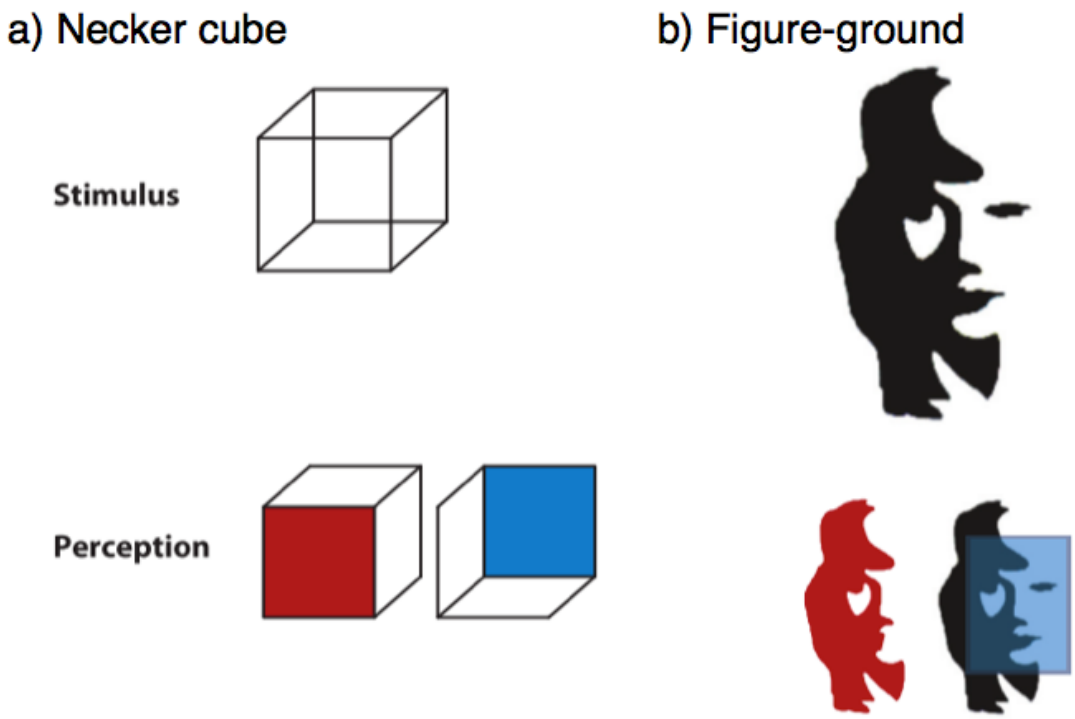

Figure 1.2: Two popular examples of multistable stimuli.

a) The Necker cube, a single visual stimulus from which two competing perceptual interpretations regarding the orientation of the leading face/depth are possible. b) A type of figure-ground illusion, during which ambiguities in figure-ground assignment can result in the same features being seen as a man playing the saxophone or shading to the face of a young woman. From Brascamp et al., 2018.

1.1.1. Introduction to binocular rivalry

In the first empirical axis of this thesis I investigate the interaction between attention and vision during one of the most thoroughly investigated multistable phenomena, known as binocular rivalry (Levelt, 1965; Wheatstone, 1838). During binocular rivalry incompatible monocular images are presented simultaneously to each eye (Alais, 2012; Blake, 1989; Blake & Logothetis, 2002; Levelt, 1965; Maier et al., 2012). If the images are sufficiently incongruent, then each image dominates conscious awareness for several seconds at a time, before awareness alternates to the competing, previously suppressed rival image (Wheatstone, 1838; Figure 1.3). Binocular rivalry benefits from a rich investigative history with a number of comprehensive reviews available on various aspects of the phenomena (Alais, 2012; Alais & Blake, 2005; Blake, 2001; Blake & Logothetis, 2002; Paffen & Alais, 2011). A brief review is here presented to frame the empirical contributions of this thesis.

1.1.2. Eye-based or stimulus-based competition

Disentangling the stage in visual processing at which binocular rivalry emerges has been an attractive prospect, and forecast to provide a benchmark for NCCs in the human brain (Crick & Koch, 1990b, 1995, 2003). As a consequence, debate has centred around whether the perceptual alternations experienced during binocular rivalry are the result of conflict resolution at relatively low or high levels of the visual hierarchy (Alais, 2012; Blake, 1989; Blake & Logothetis, 2002; Tong, Meng, & Blake, 2006).

In support of the early-stage model, conventionally summarized as eye-based rivalry (Blake, 2001), a range of basic stimulus features have been shown to affect the dynamics of perceptual alternations (reviewed in Blake, 1989; 2001). For example, an increase in stimulus contrast (Whittle, 1965), or luminance (Kaplan & Metlay, 1964) promotes the relative dominance of a visual stimulus, implicating early feature-tuned cortical interactions in the process of visual rivalry. Strong evidence for eye-based rivalry has also been obtained through the use of monocular probes, which superimpose a small target over a region of the suppressed visual image (Blake & Fox, 1974; Fox & Check, 1968; Wales & Fox, 1970). Probes presented to suppressed eye regions are reported with a loss in contrast sensitivity of approximately 0.3 log units (Alais, 2012; Blake 1989; Fox & Check, 1968; Wales & Fox, 1970), indicating the presence of early inter-ocular inhibition between monocular channels in primary visual cortex (Blake, 1989). More recently, Blake, Tadin, Sobel, Raissian, and Chong (2006) were also able to demonstrate a reduction in the strength of motion and pattern after-effects, depending on the depth of suppression during rivalry. These authors interpreted their results as evidence of early suppression mechanisms, capable of mediating the strength of neural adaptation that is traditionally thought to transpire within V1.

In contrast to these early, eye-based accounts, psychophysical evidence in support of later-stage or stimulus-based rivalry has also increased (Logothetis, Leopold, & Sheinberg, 1996; Ooi & He, 2003). Strong evidence has been provided by rapidly exchanging rivalry stimuli, such that the eye-location of each stimulus is swapped many times a second between the two eyes (Alais, 2012; Logothetis et al., 1996). While conventional eye-based models of rivalry would predict that these

alternations would be visible during the dominance of one monocular stream of information, stable percepts were capable of lasting several swaps of eye-based information (Logothetis et al., 1996). Similarly, binocular rivalry can persist when the composite features of a complete-stimulus are distributed between each eye (Kovacs, Papathomas, Yang, & Fehér, 1996; Figure 1.3c). For this type of rivalry to occur, a form of gestalt, inter-ocular grouping must have occurred to combine the composite rivalry images in support of later-stage visual suppression. Notably however, a restricted range of stimulus parameters have been shown to enable this type of inter-ocular grouping (Lee & Blake, 1999, 2004), which has catalysed support for a hybrid model of binocular rivalry that incorporates both early and late stages of visual suppression (Alais, 2012; Blake, 2001; Blake & Logothetis, 2002).For hybrid models, a two-stage pattern of early eye-based conflict detection is suggested to be followed by the recruitment of higher-level stimulus based processes (Alais & Melcher, 2007; Alais & Parker, 2006).

As a complement to these rather discrete two-stage models, recent discussions have suggested that a more dynamic interplay between internal models and stimulus attributes may give rise to binocular rivalry (Hohwy, Roepstorff, & Friston, 2008; Kanai, Carmel, Bahrami, & Rees, 2011). Hohwy et al., (2008) have applied the framework of predictive error minimization to explain binocular rivalry. In their view, incoming sensory input is compared to hypotheses of the environment that are generated by higher-level regions, which can lead to prediction errors when sensory input mismatches with internal predictions. When visual stimuli are sufficiently ambiguous, as is the case with binocular rivalry, perceptual dominance between visual alternatives in a continuous effort toward prediction error minimization. Empirical support for this dynamic interplay was also provided by Kanai et al., (2011), who observed changes to perceptual dominance durations after repetitive transcranial magnetic stimulation (r-TMS) was applied to the superior parietal cortex. Strikingly, when r-TMS was applied to anterior regions of the Superior Parietal Lobule (SPL), switch rates increased relative to pre-TMS periods. This increase was in contrast to a decrease in switch rate which followed TMS to the posterior SPL. This functional fractionation based on cortical hierarchy was interpreted in support of the predictive coding account: anterior SPL may be involved in generating environmental predictions, which when disrupted via TMS lead to weaker predictions concerning the current visual input, and more frequent perceptual switching. To forecast, **Chapters 3**

and 4 demonstrate support for such a hybrid two-stage pattern between lower-level stimulus attributes and internal models, as the presence (or absence) of stimulus conflict during binocular rivalry is shown to interact with the allocation of attention.

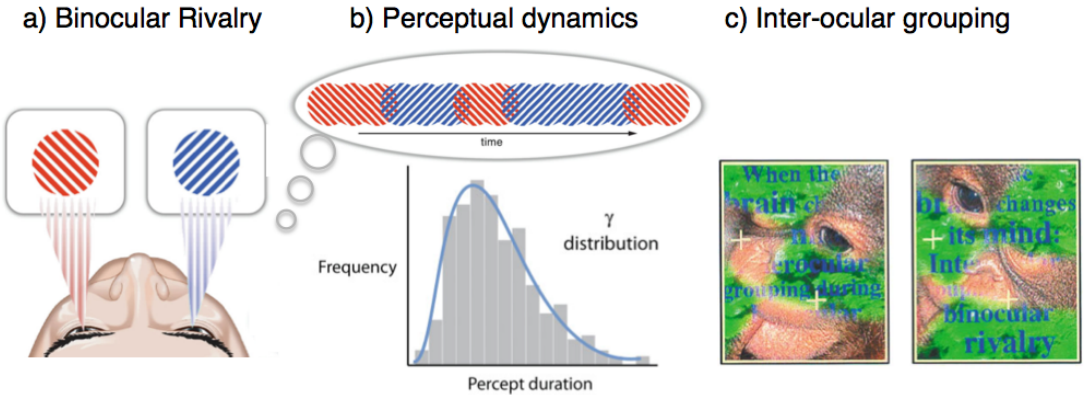

Figure 1.3: Introduction to Binocular Rivalry.

a) Example binocular rivalry input. Conflicting images are presented simultaneously, one to each eye. b) Over time, the reported dominance of each image fluctuates, with the distribution of dominance durations following a gamma distribution (a-b adapted from Dieter & Tadin, 2011; Alais, 2012). c) Example stimulus of the type used by Kovacs et al., (1996). Against rivalry occurring at only early stages of visual processing, these authors reported perceptual alternations between completed images of an ape's face, and text. Such alternations are only possible if inter-ocular grouping had occurred, in support of late-stage binocular rivalry (c adapted from Kovacs et al., 1996).

1.1.3. Attention modestly impacts binocular rivalry

Although strong effects of attention might complement claims regarding the importance of higher cortical regions to the dynamics of binocular rivalry, such effects have been difficult to establish. On balance, evidence suggests that attention is only modestly capable of affecting rivalry (Dieter, Brascamp, Tadin, & Blake, 2016; Meng & Tong, 2004; Paffen & Alais, 2011). In this thesis, a key objective was to empirically dissociate the behavioural and neural signatures of attention and contents of conscious awareness. As a result, the role of attention in the maintenance or suppression of visual imagery during binocular rivalry is critical to review.

Attention, here defined as a selective mechanism for determining the fate of stimulus processing among alternatives, is unable to completely halt or stabilize stimulus alternations during binocular rivalry (Blake, 1989; Lack, 1978; reviewed in

Paffen & Alais, 2011). The selective attention of well-trained observers to features of dominant stimuli can however slightly extend dominance durations, as well as bias the initial dominance of a stimulus via pre-cueing (Lack, 1978; Mitchell, Stoner, & Reynolds, 2004). Endogenous cueing has also been shown to bias dominance durations in favour of a cued visual image (Chong, Tadin, & Blake, 2005). Critically however, this endogenous allocation of attention is generally not strong enough to override the occurrence of subsequent perceptual alternations (Klink et al., 2008; Chong, Tadin, & Blake, 2005)[1].

While voluntary attention may extend a period of dominance, the use of spatial cues to direct attention can also cause a suppressed image to reach visual awareness (Dieter, Melnick, & Tadin, 2015; Ooi & He, 1999). This spatial cueing suggests that the spatiotemporal allocation of attention may modulate rivalry dynamics (Ooi & He, 1999; Paffen & Alais, 2011), and is consistent with reports that attending to either left-or-right hemifields increases the rate of rivalry at those locations (Paffen, Alais, & Verstraten, 2006; Paffen & Van der Stigchel, 2010). Dividing attention through the use of a simultaneous dual-task has also been shown to decrease the duration of perceptual dominance at rivalry onset (Ooi & He, 1999), as well as the perceptual alternation rate during continuously reported binocular rivalry (Alais, van Boxtel, Parker, & van Ee, 2010; Brascamp & Blake, 2012; Paffen et al., 2006).

Taken together, the reviewed evidence shows that attention is not solely responsible for perceptual change during binocular rivalry, yet remains capable of affecting the dominance durations and switch rates of a visual stimulus (cf. Paffen & Alais, 2011). This partial control of rivalry dynamics via attention suggests interplay between low- and high-stage models of visual rivalry, as mentioned above (Dieter et al., 2016; Paffen & Alais, 2011). **Chapters 3 and 4** demonstrate that rivalry dynamics are susceptible to both low-level stimulus conflict and attention, by varying stimulus conflict based on crossmodal information. We return to the effects of crossmodal stimulation on binocular rivalry in **Chapter 1.4**.

[1] Though recently, 5 participants (3 naive) when trained for ~6 hours were able to increase the strength of feature-based attention during rivalry, succeeding in extending dominance durations for minutes (Dieter, Melnick, et al., 2016).

KEY POINTS

- The content of conscious perception at a particular moment is the end result of unconscious processes that occur outside of conscious awareness.
- Multistable stimuli highlight the constructive nature of this process, as a single invariant physical stimulus can give rise to more than one conscious percept.
- Binocular rivalry is a widely studied multistable stimulus, in which two monocular presented images compete for perceptual dominance.
- During binocular rivalry, perceptual alternations are believed to result from a hybrid of both low- and high-level processing of visual conflict.
- Attention can modestly alter the dynamics of binocular rivalry.

1.2. Contemporary accounts of the NCCs

Binocular rivalry is one popular paradigm used to understand how visual content may gain access to consciousness. While the neural basis for binocular rivalry continues to be investigated (Blake et al., 2014; Carmel, Walsh, Lavie, & Rees, 2010; Kanai et al., 2011; Li, Rankin, Rinzel, Carrasco, & Heeger, 2017; Maier et al., 2008; Panagiotaropoulos et al., 2014; Tong et al., 2006; Zaretskaya, Thielscher, Logothetis, & Bartels, 2010), the neural correlates of visual awareness are also popular due to the well characterized architecture and accessibility of the visual system (Felleman & Van, 1991; Klink, Self, Lamme, & Roelfsema, 2015; Figure 1.4) Before reviewing the empirical literature investigating the NCCs, it is worth briefly summarizing prominent theoretical accounts for the neural basis of consciousness, and how they relate to attention.

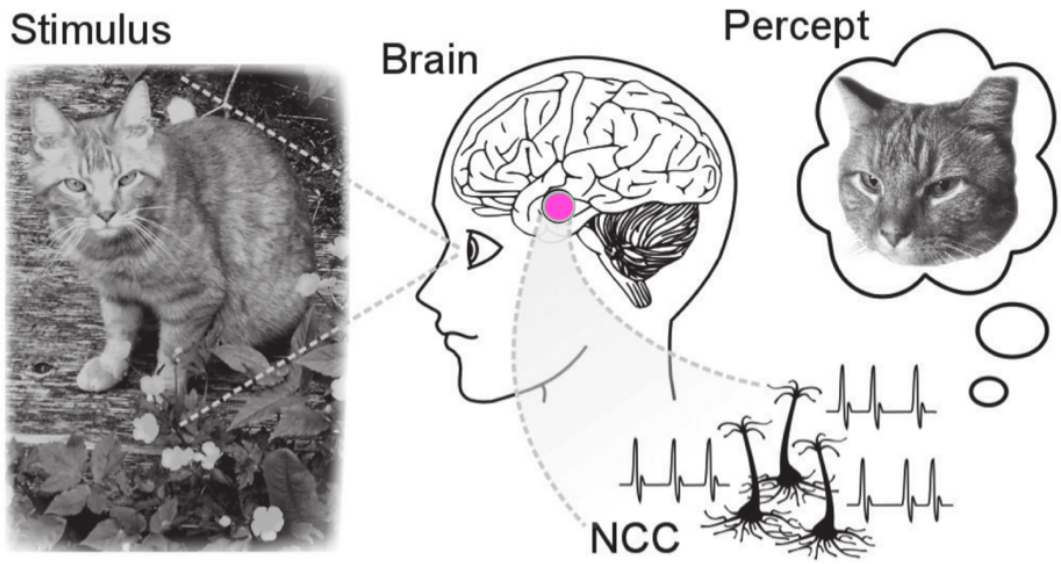

Figure 1.4: A pathway to the Neural Correlates of Consciousness (NCC).

The NCCs are defined as the minimal set of neural substrates that give rise to a conscious percept (Koch, 2004). In this example, viewing a cat corresponds to spiking activity in temporal cortex, which correlates with the experienced percept, "I have seen a cat" (adapted from Klink et al., 2015).

1.2.1. Current theoretical accounts for the neural basis of consciousness

The majority of contemporary theories of consciousness support the idea that consciousness emerges from distributed neural processes, rather than a specific cortical locus within the brain (for specific loci see, for example; Goodale & Milner, 1992; Posner, 1994; Zeki, 2003). Though the current landscape of neurobiological theories of consciousness is too large to be comprehensively reviewed here (reviewed in Seth, 2007), key features of prominent theories will be revised to frame the *where*, and *when,* that NCCs can be expected to arise.

Frequency-specific oscillations

To begin, Crick and Koch (1990a, 1990b) revitalized the empirical study of consciousness by proposing that gamma activity (40-70 Hz range) may mediate conscious access, whereby synchronous gamma firing in distributed cortical areas acts to bind different features into a unified percept (Tallon-Baudry & Bertrand, 1999).

Similarly, Llinás and colleagues have suggested that cortico-thalamic loops of synchronous 40 Hz activity could constitute the basis of conscious perception (Llinás & Paré, 1991; Llinás, Ribary, Contreras, & Pedroarena, 1998). The functional role of neural synchrony has also been championed in visual perception (Fries, Neuenschwander, Engel, Goebel, & Singer, 2001; Fries, Roelfsema, Engel, König, & Singer, 1997), more specifically in the 30-80 Hz range (Tallon-Baudry, 2009; Tallon-Baudry & Bertrand, 1999). While these accounts offered an important focus for empirical investigation, the role of gamma oscillations recorded at scalp level has been countered, by showing such high-frequency activity may reflect micro-saccades upon object exploration, as opposed to the binding of distributed stimulus features into a unified percept (Yuval-Greenberg & Deouell, 2009; Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell, 2008). Outside of these frequency-specific accounts, exactly how distributed regions interact has been a point of contention among the leading theoretical accounts of consciousness (Klink et al., 2015).

A shared global workspace

For example, global workspace theory (GWT; Baars, 1993, 1997), proposes that a single shared neuronal resource maintains selected information, with the contents of consciousness determined by competition in a winner-takes-all fashion. In this account, peripheral cortical regions compete for access to this global resource, with a contrast between conscious and unconscious processing demarcated by access to the workspace. Extending on GWT, Dehaene, Naccache, and Changeux have proposed a global neuronal workspace model (Dehaene & Changeux, 2005, 2011; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Dehaene, Kerszberg, & Changeux, 1998; Dehaene & Naccache, 2001; Sergent & Naccache, 2012), suggesting that the central workspace may be promoted by pyramidal neurons with long-range axons in frontal parietal, and temporal cortices. A key feature of this model is that conscious access, or updating of the global workspace is accompanied by an ignition in the form of widespread neural activity to peripheral regions (for recent review see Dehaene & Changeux, 2011). Prior to ignition, peripheral sensory regions completing specialized processing lie largely outside the scope of consciousness, with the sensory content therein believed to be subliminal (never reaching consciousness), or preconscious (with the potential to be broadcast via ignition).

Recurrent processing

In a similar vein, recurrent processing theories of consciousness also make a distinction between unconscious processing in localized nodes and the necessary activation of a broader network (Lamme, 2003, 2006). In these accounts, sensory stimulation initiates a feedforward sweep, that is subsequently followed by a second stage of activity known as recurrent processing (Lamme & Roelfsema, 2000). The recurrent theory is usefully applied to visual consciousness, with specific roles for feedforward, feedback, and recurrent neural processing of visual information along the cortical hierarchy (Klink et al., 2015; Lamme, 2006, 2010). The theory posits that primarily feedforward processing follows the presentation of visual stimulus to the retina, with different cortical regions extracting information along this feedforward sweep. For example, in successive stages, features such as colour, shape, or orientation are progressively extracted from the visual scene, before this feedforward information flow is output to motor and prefrontal areas of the brain where it is maintained (Lamme & Roelfsema, 2000). This output results in additional lateral, non-feedforward or 'recurrent' processing, which is viewed as necessary for consciousness to emerge. In other words, if a visual stimulus is not strong or salient enough to successively recruit recurrent processing in higher cortical areas via top-down signals or horizontal connections, then the stimulus will not be consciously perceived (Lamme 2003, 2010). Notably this theory does not predict the specific involvement of attention-related networks, nor that high-order prefrontal regions will be necessary prerequisites to consciousness.

A Dynamic core and re-entrant, integrated information

Another way in which distinct features of the environment may be bound together into a unified percept is via re-entrant connections between the cortex and nuclei within the thalamus (Edelman, 1989, 1993; Seth, Izhikevich, Reeke, & Edelman, 2006; Sporns, Gally, Reeke, & Edelman, 1989). This thalamo-cortical circuit has been proposed to define a functionally significant, relatively stable, 'dynamic core' (Tononi & Edelman, 1998; Tononi, Edelman, & Sporns, 1998). A crucial feature of this dynamic core is that for specific conscious experiences, or qualia, the properties of this dynamic core become highly differentiated (representing a single percept), as well as highly integrated (Tononi, 2004). While a plastic, re-

entrant dynamic core is reminiscent of GWT (Edelman, Gally, & Baars, 2011), dynamic cores are suggested to emerge as life progresses, through a process of neural Darwinism which prunes and reinforces synaptic connections that improve coordinated brain function (Edelman, 1987; Edelman & Mountcastle, 1978).

Tononi (2004; 2008) has recently extended this account into a mathematical formalism for consciousness known as Integrated Information Theory (IIT). Building upon the notion that differentiation and integration are the hallmarks of moment-to-moment experiences, IIT quantifies consciousness based on a system's capacity for integrated information. This measure ('phi') is used to quantify the integrated information in complex systems in a computational framework, based on the mutual information and complexity of the system. One of the most striking predictions of IIT is that non-zero phi, and hence non-zero consciousness, emerges in any complex system with sufficiently integrated information processing, which may include non-biological networks. At present however, calculating phi is limited to smaller, simpler networks, as the computational demands significantly increase with an increasing number of nodes in the network, and the number of their connectivity patterns (Oizumi, Albantakis, & Tononi, 2014; Tononi & Sporns, 2003).

While Tononi's IIT may not be empirically testable in current laboratory settings, the remaining prominent theories of consciousness predict that consciousness is supported by a distributed network of patterned neural activity: the NCCs may reside in a shared global workspace connecting sensory regions, within re-entrant activation along the cortical hierarchy, or through the activation of a dynamic, thalamo-cortical loop that changes with moment-to-moment experiences (Figure 1.5).

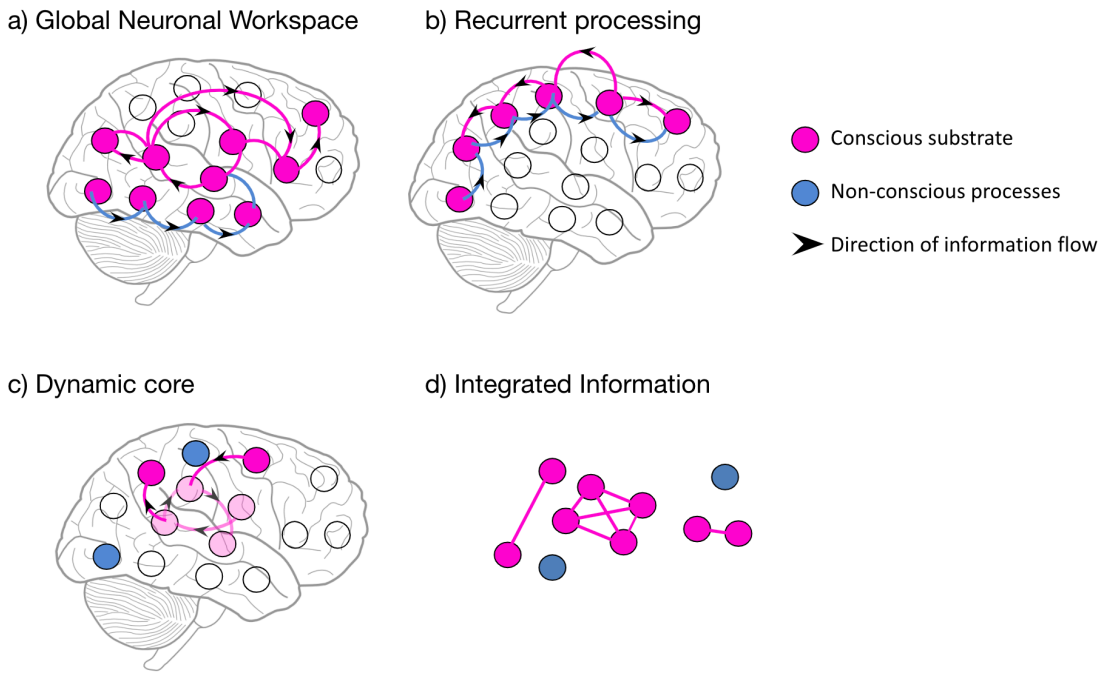

Figure 1.5: Distinct neural mechanisms are proposed to support the neural bases of consciousness.

a) According to Dehaene and Naccache (2001), neural processes compete for access to a shared global neuronal workspace, the contents of which are determined by top-down feedback in a winner take all fashion. b) Lamme and colleagues (2003, 2010 have proposed that recurrent feedback to earlier sensory regions is necessary for the maintenance of conscious perception. c) Alternatively, feedback between cortical areas and the thalamus may constitute a dynamic core, whereby neural integration corresponds to different conscious states (Edelman & Tononi, 2000). d) Integrated information has also been proposed as an index of consciousness (Tononi, 2004).

1.2.2. Where are the NCCs? Insights from fMRI

Having briefly reviewed theoretical accounts for the neurobiological basis of consciousness, I will now introduce empirical efforts to localize the NCCs. Particular focus is paid toward potential overlap between NCCs and other cognitive functions, including attention and report.

When exploring for the NCCs, contrasts are typically made either between the neural activity which accompanies alternative, stable periods of perceptual dominance (e.g. Tononi, Srinivasan, Russell, & Edelman, 1998), or the transition periods when changes in consciousness are being endogenously generated (e.g. Lumer, Friston, & Rees, 1998). NCCs have primarily been localized using techniques with relatively fine spatial resolution, such as fMRI, focusing on transient changes to blood-oxygen level dependent (BOLD) activity during the perception of multistable images.

Among the first of these accounts, Lumer et al. (1998) recorded BOLD signals during binocular rivalry, and in their critical manipulation, also during physical replay conditions designed to mimic the alternating states of perceptual dominance. During these replay states, instead of having the contents of consciousness switch between each eye's competing images naturally, Lumer et al., physically presented two stimuli in alternation. During the endogenously generated, spontaneous changes in conscious perception, a right-lateralized fronto-parietal region was shown to increase in BOLD activity when compared to the physically alternating replay condition. This activity was taken as a neural correlate of true changes in consciousness, and since this early influential work, a right fronto-parietal region been supported using a variety of paradigms (for review Brascamp et al., 2018; Sterzer et al., 2009).

In the most recent review on this topic, Brascamp et al. (2018) synthesized available neuroimaging evidence to highlight all the brain regions associated with changes in visual awareness during multistable perception. Their results of a meta-analysis collated the empirical research inspired by the early work of Lumer et al., (1998), and thus focused on brain regions implicated during perceptual reversals as captured by fMRI-BOLD activity. They also included in their review sites that alter the rate of perceptual changes when targeted with neurostimulation methods such as transcranial magnetic stimulation (TMS). As outlined above, the manipulation of neural states via TMS has implicated right parietal cortical regions in the dynamics of perceptual multistability (Carmel et al., 2010; Kanai et al., 2011; Zaretskaya et al., 2010). Overall, the meta-analysis of Brascamp et al., (2018) confirmed that a right-lateralized network of frontal and parietal cortical regions has routinely been identified using contrastive analysis and TMS during multistable perception. Figure 1.6a displays their key figure, and highlights the regions identified as putative NCCs based on perceptual multistability.

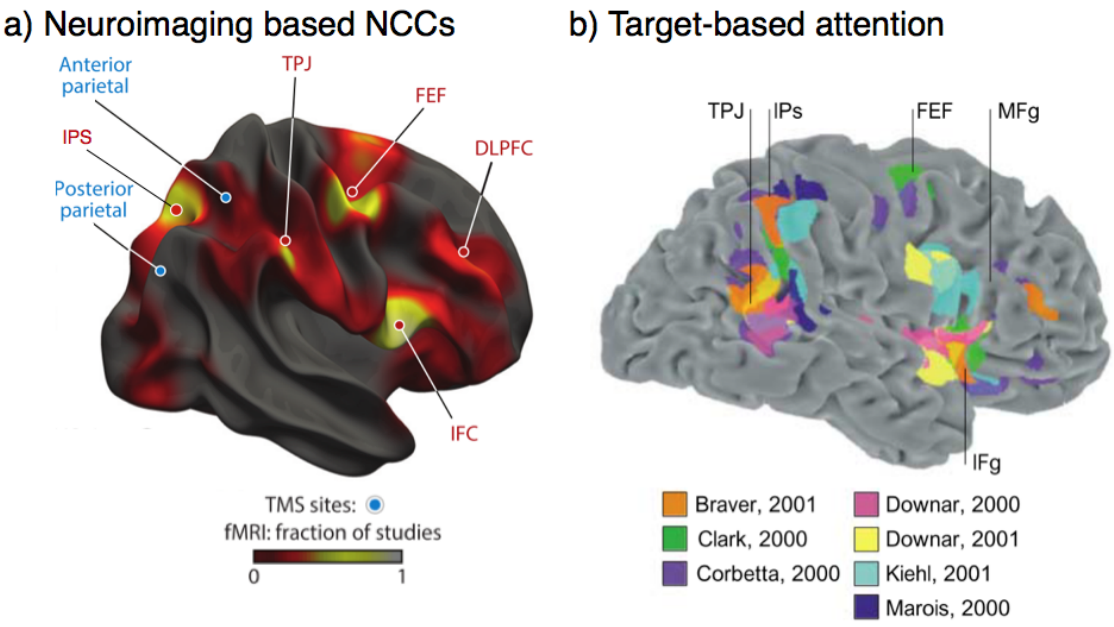

Figure 1.6: A functional overlap between NCCs and attention-related fronto-parietal networks.

a) Results of an fMRI and TMS meta-analysis, showing the locations where recorded fMRI (BOLD) activation was greater during perceptual reversals than during replay, or physical stimulus alternations using multistable stimuli (adapted from Brascamp et al., 2018). b) The results of a meta-analysis of imaging studies capturing goal-directed attention to changes in the visual environment. The different colours represent separate studies (see figure legend; adapted from Corbetta et al., 2002). TPJ = temporoparietal junction, IPs= intraparietal sulcus, IFC= inferior frontal cortex, IFg= inferior frontal gyrus, FEF= frontal eye field, MFg = middle frontal gyrus, DLPFC = dorsolateral prefrontal cortex; IFC = inferior frontal cortex.

1.2.3. NCCs are distributed, and overlap with attention-related networks.

The first obvious feature to emerge from the synthesis of Brascamp et al., (2018), is that endogenous changes in consciousness have been associated with a wide variety of networks, rather than a single locus (Figure 1.6a). Given the highly integrated nature of the cortex, this should come as little surprise, but it is the shared functional organization of these regions that has spurred the most consistent debate within the literature[2].

[2] Most recently, the relative importance of pre-frontal and parietal regions has come to dominate the discourse. While both sides agree that consciousness is reliant upon a broad network of

For example, although it is tempting to implicate these areas with purely perceptual processes, the regions identified support a wide range of cognitive functions. As these regions were activated and identified by time-locking BOLD activity to a reported change in perception, whether these activations also reflect the demands of a subjective report, or other cognitive processes, must be identified.

This sentiment was echoed in the earliest works on this topic, as Lumer et al. (1998) were quick to acknowledge that there may be a close association between the mechanisms that govern multistable perception, and thus consciousness, with the mechanisms that govern the allocation of attention. To investigate this claim Sterzer and Kleinschmidt (2007) analysed changes in chronometric BOLD activity over time, using a temporal resolution (step-size) of 2s intervals. These authors reasoned that if increased activation in BOLD activity represented a consequence of reporting on a change, rather than heralding an endogenously generated change, then the time-course of BOLD activity could differentiate whether a putative NCC served a causal, or ancillary role to changes in perception. Using this chronometric analysis of BOLD activity, only the right inferior frontal cortex (IFC) displayed an increase in BOLD signal *before* the report of an endogenous change in perception, which the authors interpreted as a neural correlate for the brain selecting among competing perceptual alternatives (Sterzer & Kleinschmidt, 2007). Critically however, these increases in local BOLD activity cannot easily be differentiated from other higher cognitive processes, including attention and the preparation of report. The extent of this potential overlap is clear when comparing the NCCs identified by Brascamp et al., with another review on the networks serving attention (Corbetta & Shulman, 2002; Figure 1.6b).

In summary, when viewing multistable images, fMRI reveals increased BOLD activity over a wide range of cortical regions. The use of physical replay conditions has been critical to these estimates, as well as efforts to distinguish the true neural correlates of perceptual change from the correlates of attending to stimulus reversal. However, significant overlap remains between the implicated NCCs, and cortical regions that increase in activity with goal-directed attention. This overlap emphasizes

---

cortical regions, the *necessity* of intact parietal and prefrontal cortical structures has been re-examined on the basis of clinical neurophysiology, especially lesion studies (Boly et al., 2017; Odegaard et al., 2017).

the need to improve methods capable of disambiguating these phenomena, and is a key motivation for the contributions of this thesis.

1.2.4. When are the NCCs present? Insights from M/EEG

Owing to their superior spatial resolution, hemodynamic based neuroimaging methods such as fMRI have been primarily used to localize NCCs. It's important to note however that BOLD activity provides only an indirect proxy for increased neural activity, and the exact nature of the relationship between increased oxygen uptake and neural firing is unclear (Logothetis, 2008). Furthermore, as indexed by chronometric BOLD activity, methods with an increased temporal resolution may also contribute valuably to our understanding of the correlates of consciousness.

To supplement fMRI analyses, magnetoencephalography and electroencephalography (M/EEG) have been used to record time-varying changes in neural activity from scalp level sensors, at millisecond temporal resolution. Here I will focus on the most common methods for inferring about consciousness via M/EEG data; event-related potential (ERP) and endogenous oscillatory activity analyses (reviewed in Gallotto, Sack, Schuhmann, & de Graaf, 2017; Kornmeier & Bach, 2012).

1.2.5. ERPs do not capture the pure contents of perception

Over the long history of EEG analysis (Berger, 1929; Haas, 2003) the P300 (~300 ms positivity, widespread over midline central and parietal scalp electrodes) has been closely associated with the perceived presence of a stimulus (Davis, 1964; Hillyard, Squires, Bauer, & Lindsay, 1971; Spong, Haider, & Lindsley, 1965). Since its identification, the P300 has been redefined to include two components, an earlier frontal component (P3a; Polich, 2007; Sutton, Braren, Zubin, & John, 1965), and a slightly later fronto-parietal component (P3b; Dehaene & Changeux, 2011; Squires, Hillyard, & Lindsay, 1973a, 1973b). While the P3b has indexed the conscious detection of stimuli with a clearly defined onset (e.g. Del Cul, Baillet, & Dehaene, 2007; Sekar, Findley, Poeppel, & Llinás, 2013; Sergent, Baillet, & Dehaene, 2005), in the context of multistable perception, such a specific interpretation is difficult to obtain. As a non-exhaustive list, the P300 component has been implicated in the

conscious recognition of a perceptual reversal (Britz & Pitts, 2011; Strüber & Herrmann, 2002), as well as attention during (O'Donnell, Hendler, & Squires, 1988) and after reversal-completion (Isoglu-Alkaç et al., 2000).

Outside of the P300, a number of alternate ERP signatures have been implicated as neural correlates of endogenously generated perceptual reversals (Figure 1.7). To identify these components, multistable images are typically presented with short inter-stimulus intervals, with onset-related ERP activity averaged to capture the neural correlates of perceptual reversals occurring shortly after stimulus onset (Kornmeier & Bach, 2012; O'Donnell et al., 1988; Orbach, Zucker, & Olson, 1966). When comparing trials with perceptual reversals to those without, an occipital reversal positivity (RP) has been observed approximately 130 ms after the presentation of Necker cubes (Britz et al., 2009; Kornmeier & Bach, 2005, 2006), as well as binocular rivalry stimuli (Britz & Pitts, 2011). As the RP occurs early following stimulus onset, and for different varieties of multistable stimuli, it has been argued to reflect the detection of visual ambiguity, most likely generated from occipital sources (Kornmeier & Bach, 2012). As a putative NCC however, it has also been criticized, as it is heavily reliant on the number of trials collated in the averaging procedure, with a low signal-to-noise ratio typically reaching <1 microvolt in amplitude (Britz et al., 2009; Kornmeier & Bach, 2005, 2006). Reduced alpha power over occipital electrodes has also been identified using the onset-paradigm, consistent with the notion that reduced alpha-power indexes cortical excitability (Hanslmayr, Gross, Klimesch, & Shapiro, 2011). As such, alpha power has been interpreted to gate the top-down selection of one perceptual alternative over another, in the case of multistable visual stimuli (Kornmeir & Bach, 2012).

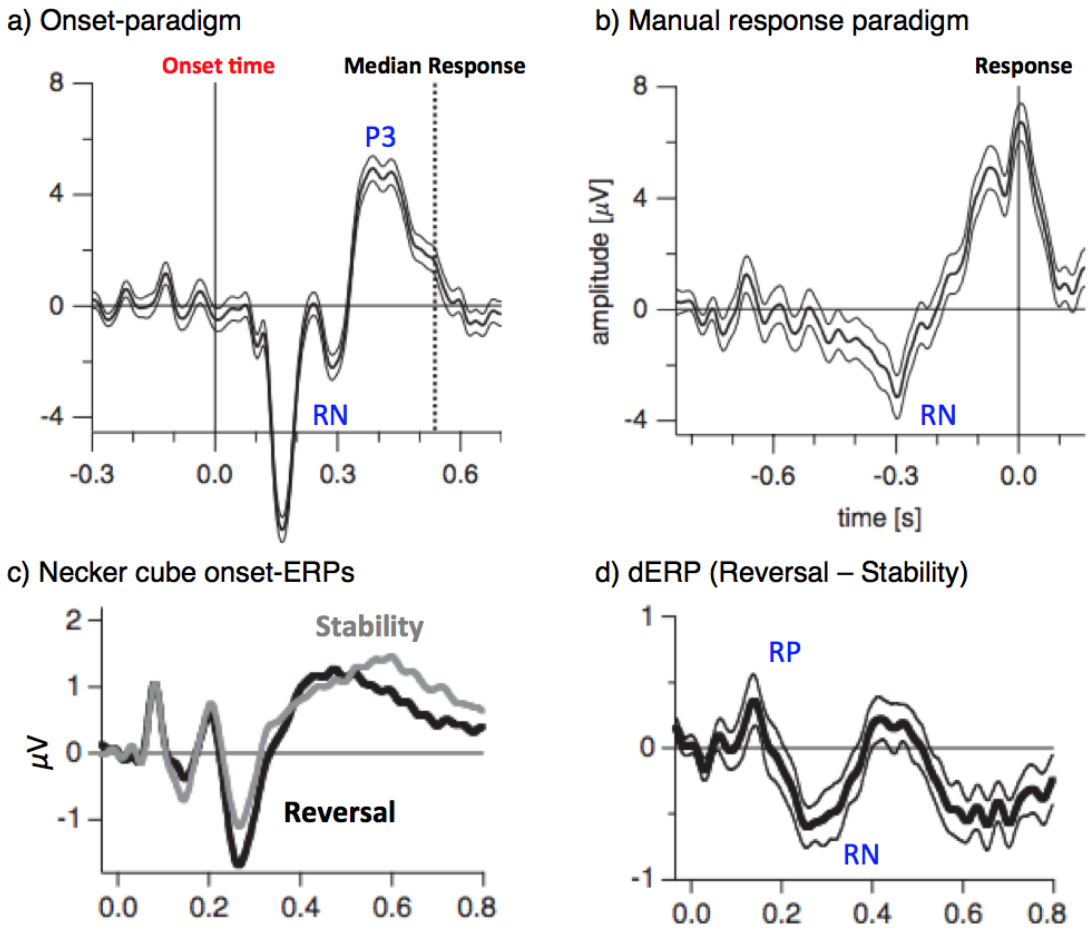

Figure 1.7. ERP correlates of perceptual reversals when viewing a Necker cube, recorded from occipital electrode location Oz.

a) When examining the temporal correlates of perceptual reversals, ERPs can be aligned at the onset of a stimulus (solid vertical line). Response times will vary (broken vertical line). A reversal negativity (RN) and later positive component (P3) are displayed. b) Alternatively, ERPs can be aligned to response times, though the precise timing and generator of ERP components is unknown. c) Using the onset-paradigm, trials containing perceptual reversals (in grey) and without (black) show similar ERPs. d) The difference (dERP) between these two (in c) indicates a reversal positivity and negativity (RP; RN), which is enhanced for perceptual reversals. Note however, the low signal-to-noise ratio of these measures. Adapted from Kornmeier & Bach, 2004).

Aside from the RP and reduced alpha power, like fMRI analyses, M/EEG components elicited by endogenous reversals and their physical replay are often difficult to dissociate. For example, both endogenous and exogenous perceptual reversals emit reversal negativities (RNs, ~220-260ms; Britz et al., 2009; Intaitė, Koivisto, Rukšėnas, & Revonsuo, 2010; Kornmeier, Ehm, Bigalke, & Bach, 2007; Kornmeier, Pfäffle, & Bach, 2011) a fronto-parietal positivity similar to P300

components (Britz & Pitts, 2011; Pitts, Gavin, & Nerger, 2008), increased gamma (40-65 Hz) and beta activity (14-26 Hz; Kornmeir & Bach, 2012).

In addition to time-locking ERP responses to a stimulus onset, time-locking to participant responses has also been adopted with distinct implications for the interpretation of ERP data (Figure 1.7b). Similar to onset-paradigms, an increased P300-like positivity has been identified and linked to the recognition of a perceptual reversal (~250 ms pre key-press; Mathes, Strüber, Stadler, & Basar-Eroglu, 2006; Strüber, Baar-Eroglu, Miener, & Stadler, 2001; Strüber & Herrmann, 2002). Reduced occipital alpha power (~300-200ms pre key-press; Isoglu-Alkaç et al., 2000; Strüber & Herrmann, 2002) and increased gamma activity (~1000ms pre key-press; Başar-Eroglu, Strüber, Kruse, Başar, & Stadler, 1996) can also be observed. These components indicate that endogenous perceptual reversals may take place up to 1000ms before participant responses (if identified by gamma increase), yet crucially, owing to the inherent variability in participant response times, the precise timing and neural correlates of endogenous perceptual reversals is difficult to isolate.

In summary, while NCCs based on fMRI overlap with attention, M/EEG recordings also suffer in their ability to distinguish genuine perceptual reversals from the stimulus matched physical replay. Paradigms are required to not only distinguish true NCCs from their physical replay counterparts, but also to distinguish between the overlapping foci of consciousness and attention.

KEY POINTS

- Using fMRI to *spatially* localize NCCs reveals considerable overlap with attention related networks.
- How NCCs vary over time is informative, though fMRI lacks the temporal resolution to study rapid changes in perception.
- Using M/EEG analyses to *temporally* localize NCCs implicate both early and late ERP components.
- Distinguishing ERP correlates of genuine perceptual reversals from physical replay is of key concern.

1.3. Differentiating the NCCs from attention and conscious report

Research regarding the NCCs has been inherently reliant upon participant introspection and subjective report to be aware of the timing of internal events. As such, to move forward it is of particular importance to ascertain whether attention and consciousness are inextricably linked, or whether they are dissociable in some way.

1.3.1. Attention and consciousness

The question of whether attention and consciousness are functionally distinct, and the relationship between them remains an ongoing topic of debate (Dehaene et al., 2006; Koch & Tsuchiya, 2007; van Boxtel & Tsuchiya, 2014; van Boxtel, Tsuchiya, & Koch, 2010a). Attention in this context refers to selective, directed attention, and thus would be seen to increase upon introspection during subjective report. As the focus of attention can define what is consciously perceived, previous studies have argued for the strong functional connection between attention and consciousness (Baars, 2002; Merikle & Joordens, 1997; Posner, 1994). Others have proposed that attention is a requirement for consciousness, or even the same thing (Cohen, Cavanagh, Chun, & Nakayama, 2012; Posner, 1994, 2012).

Although the functional overlap between NCCs and attention related networks is undisputed, it is now an open question whether the fronto-parietal NCCs represent changes in top-down control of visual perception (e.g. Leopold & Logothetis, 1999), or the end result of cortical changes occurring in earlier levels of visual processing (cf. Paffen & Alais, 2011). Toward distinguishing these alternatives, novel paradigms have begun to reveal an independence between attention and particular aspects of conscious experience (Chica & Bartolomeo, 2012; Koch & Tsuchiya, 2007; van Boxtel, Tsuchiya, & Koch, 2010b; van Boxtel et al., 2010a). For example, van boxtel et al., (2010) demonstrated a double dissociation between the effects of attention and awareness, by showing opposite effects on the length of visual afterimages (Figure 1.8). In their paradigm, an after-image inducer was presented off fixation, and either consciously visible, or masked from awareness through the use of continuous flash suppression (CFS; Tsuchiya & Koch, 2005). Subjects then either attended to the location of the after-image inducer or ignored that location, while the

visibility of this induced was either suppressed (or not) from awareness. The duration of afterimages formed by this inducer were shortened by attention, though increased with visibility. These opposing effects offered compelling evidence that attention and consciousness are distinct processes with the potential for their neural correlates to be disambiguated (yet see Cohen et al., 2012).

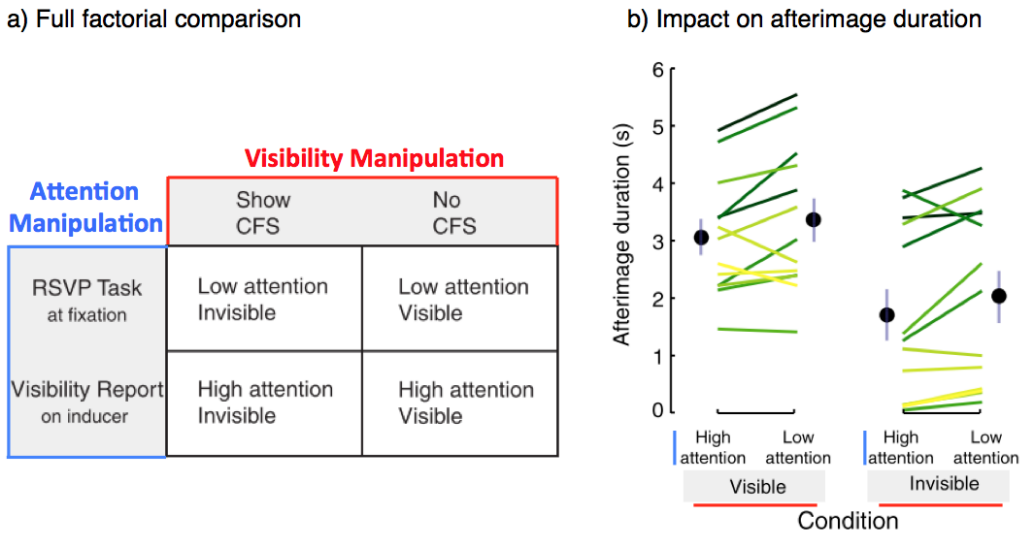

Figure 1.8: Experimental manipulation to dissociate attention and visibility.

a) van Boxtel et al., (2010) compared the effects of attention and visibility on the duration of afterimages using a 2 x 2 factorial design. b) The visibility of an after-image inducer positively increased after-image duration, while increased attention reduced the duration of after-images. These results support the potential to dissociate the neural correlates of conscious visibility from attention, as here their effects are distinct.

Neural evidence for the dissociation between attention and consciousness has been less forthcoming (Chica et al., 2012). In one study using a similar variant of CFS and a 2 x 2 factorial design (dissociating attention from awareness), Watanabe et al., (2011) examined whether fMRI-BOLD activity in V1 was modulated by attention or awareness of monocular targets. In their study, when targets were visible but relatively unattended (by allocating attention instead to fixation), fMRI-BOLD signals in retinotopic areas corresponding to the target location went relatively unchanged compared to pre-stimulus periods. By contrast, when attending to the location of either visible or invisible targets, the percentage change in fMRI-BOLD activity compared to pre-stimulus periods was greatest. The authors interpreted this attention dependent signal change as evidence for a dissociation between attention and

awareness, implicating the correlates of visual awareness outside of these early areas of the visual hierarchy (note however, that this study has since been argued to be underpowered; Sterzer, Stein, Ludwig, Rothkirch, & Hesselmann, 2014; Yuval-Greenberg & Heeger, 2013). Another fMRI study by Bahrami, Lavie, and Rees (2007), demonstrated increased BOLD activity in V1 with attentional allocation toward invisible objects, suggesting attention rather than visibility mediated early cortical activity. These authors followed this fMRI study with related psychophysical evidence to show that attentional resources can be attributed to stimuli outside of conscious awareness (Bahrami, Carmel, Walsh, Rees, & Lavie, 2008b, 2008a). In a complement to these dissociations, activity beyond V1 in dorsal regions also persists when invisible visual objects are suppressed during CFS (Fang & He, 2005), demonstrating that the locus of invisible visual information is not confined solely to early visual areas. Together, attention rather than visibility may mediate early cortical responses (Bahrami et al., 2007; Watanabe et al., 2011), yet invisible information also may penetrate higher up the cortical hierarchy (Fang & He, 2003).

On important caveat is that these behavioural and fMRI studies have relied upon inter-ocular suppression to remove an object from awareness (Sterzer et al., 2009). While these techniques are powerful, they have generated new controversies, particularly as the depth of suppression may wane, resulting in temporary breakthroughs of previously suppressed stimuli into conscious awareness (reviewed in Sterzer et al., 2009; Blake et al., 2014). In contrast, Wyart and Tallon-Baudry (2008) have also demonstrated a dissociation between attention-related and consciousness related neural activity by thresholding the visibility of targets on a trial-by-trial basis. They used MEG to compare oscillatory activity when subjects attended to stimuli which were seen/unseen, and stimuli which were visible when attended/unattended. An attention dependent high-frequency gamma band correlate, and visibility dependent low-frequency gamma correlate were identified, the latter of which was later localized to the posterior lateral occipital cortex (Wyart & Tallon-Baudry, 2009).

Although a nascent literature, these studies point to correlates of consciousness that are independent to attention, located either outside of (e.g. Bahrami et al., 2007; Watanabe et al., 2011), or within (e.g. Wyart & Tallon-Baudry, 2008, 2009) early visual areas. Many open questions remain. As attention is not a unitary concept, but comprising both spatial, temporal, and feature-based dimensions, the manipulation of these dimensions may offer new insights into the dissociation

between attention and consciousness. This spatiotemporal allocation of attention is addressed in **Chapters 3 and 4**. Further, if attention can act independently to conscious visibility, how might the attentional selection of an invisible stimulus or object be mediated? A novel feature of this interaction is revealed in **Chapter 4**.

1.3.2. The impact of subjective report and replay conditions

While the need to distinguish attention from consciousness is increasingly recognized, necessary concerns have also been raised regarding the validity of replay conditions when attempting to mimic multistable perception (Frassle, Sommer, Jansen, Naber, & Einhauser, 2014; Knapen, Brascamp, Pearson, van Ee, & Blake, 2011; Tsuchiya, Wilke, Frässle, & Lamme, 2015; Weilnhammer, Ludwig, Hesselmann, & Sterzer, 2013). Binocular rivalry in particular is challenging to emulate (Frassle et al., 2014; Knapen et al., 2011). Transitions between stable dominance periods are often perceived as a complex mix of both eyes images (Alais et al., 2012), often with changes moving as a travelling wave across the perceived image (Wilson, Blake, & Lee, 2001), or as a patchwork of the two (Kovacs et al., 1996). In an early study to assess the effect that replay features have on the NCCs, Knapen et al., (2011) focused on the time taken to perform these induced replays. Knapen and colleagues mimicked replay with extended transition periods, rather than the quasi-instantaneous physical reversals that had been used in previous investigations of binocular-rivalry replay (e.g. Lumer et al., 1998). By indicating both the onset and offset of a perceptual transition, physical replay conditions could be custom tailored to individualized changes in perception. During genuine perceptual reversals, this study supported a greater transition-related BOLD signal within right fronto-parietal networks (Lumer et al., 1998; Sterzer & Kleinschmidt, 2007). For their unique, duration matched physical replay conditions, however, fMRI-BOLD activity was indistinguishable from spontaneous reversals. As a result, fronto-parietal regions cannot solely be responsible for the initiation of perceptual transitions during multistable perception (Knapen et al., 2011).

As an alternative to more closely approximating transitions with replay, paradigms have also reduced the salience of perceptual reversals (Brascamp, Blake, & Knapen, 2015a; Zou, He, & Zhang, 2016), or removed the need for report entirely by tracking signals from the eyes (e.g. Fox, Todd, & Bettinger, 1975; Frassle et al., 2014)

or brain directly (Brouwer & van Ee, 2007; Brown & Norcia, 1997; Haynes, Driver, & Rees, 2005; see Figure 1.9). These results have also shown that fronto-parietal activity is strongly related to task-demands, including the initiation of a report. For example, in an effort to remove task-demands during binocular rivalry, Brascamp et al., (2015) cleverly rendered perceptual transitions imperceptible, by matching the images being presented to both eyes. To track changes in consciousness, short probes were introduced intermittently which would be perceived differently by the participants depending on whichever eye was dominant at that moment. They then contrasted fronto-parietal BOLD activation while participants searched for these probes, and found that when rivalry went unnoticed, transition-related frontoparietal activity was minimal. When perceptual alternations are tracked during binocular rivalry based on reflexive ocular-motor data, this reduced role of fronto-parietal activity persists, again in the absence of overt responses (Frassle et al., 2014). By presenting drifting gratings to the two eyes, the contents of perception were inferred from whether micro saccades were following a leftward or rightward drift. In this absence of active report, these authors demonstrated that fronto-parietal BOLD activity was weakened, in support of an overlap between previously obtained NCCs and networks responsive to task-demands (Figure 1.9).

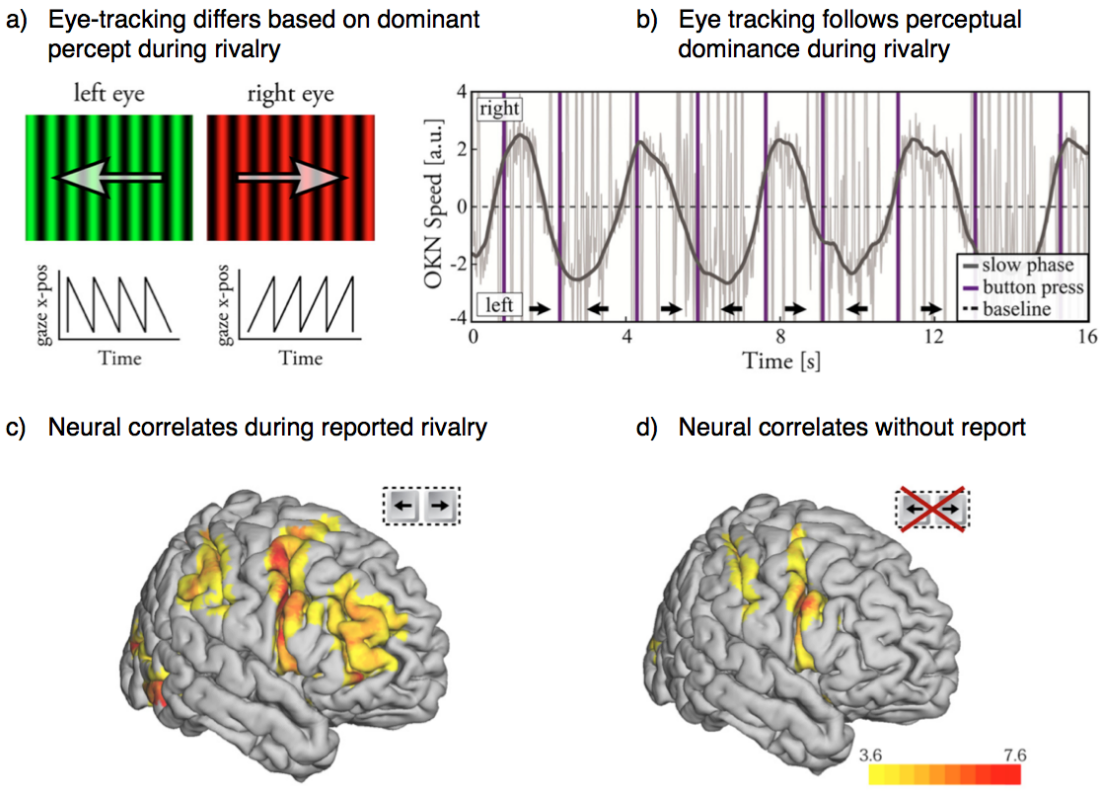

Figure 1.9: The effects of active report on the NCCs during binocular rivalry.

a) Frassle et al., (2014) used leftward or rightward drifting gratings to manipulate reflexive eye movements during binocular rivalry. B) The speed and direction of small eye movements (optokinetic nystagmus; OKN) tracked perceptual reversals. c) When reporting on perceptual reversals using button-press, a traditional right lateralized fronto-parietal network of NCCs increased in BOLD activity. D) In the absence of report, these NCCs differed significantly. Colour scale represents z-scores comparing rivalry to replay conditions.

For the BOLD activity recorded during fMRI, and subsequently used to localise NCCs, the aforementioned timing differences (Knapen et al., 2011), salience of replay (Brascamp et al., 2015), and task-demands (Frassle et al., 2014) represent key challenges moving forward (Brascamp et al., 2018; Tsuchiya, Wilke, Frassle, et al., 2015). The second empirical axis of this thesis is motivated toward addressing these concerns. Specifically, as any salient distinction between a physical replay condition and endogenous reversal may contribute to the evoked NCCs, **Chapters 5 and 6** introduce relatively simple and covert examples of physical stimulus replay, in a novel multistable paradigm. I will now introduce the methodological advantages of these empirical contributions.

KEY POINTS

- Though attention and consciousness can be dissociated behaviourally, neural evidence has been less forthcoming.
- Open questions concern how the link between attention and visible/invisible stimuli may be mediated.
- Through reliance on binocular rivalry and inter-ocular suppression, it is difficult to mimic genuine perceptual reversals with physical alternatives.
- An open concern regards the subjective saliency of physical replay compared to genuine perceptual reversals, and impact on measures of the NCCs.

1.4. Conceptual advantages of crossmodal rivalry

Thus far I have reviewed contemporary accounts for the location of NCCs, which rely on inter-ocular suppression and increases in fMRI based BOLD activity. These efforts have resulted in debate concerning the relative importance of frontal and parietal regions (see also Boly et al., 2017; Odegaard, Knight, & Lau, 2017), as well as the nature of any interaction between attention and conscious awareness (Koch & Tsuchiya, 2007; Tallon-Baudry, 2012; van Boxtel & Tsuchiya, 2014). While I have focused on the overlap between attention and the NCCs, a second point to address is the dominance of binocular rivalry in the literature that is available, and consequent visual-only bias of current investigations (Dykstra et al., 2017; Faivre, Arzi, Lunghi, & Salomon, 2017).

To illustrate this point, of the 10 studies included by Brascamp and colleagues (2018) review comparing reversals to replay, all involved the perception of visual stimuli. Six were variants of binocular rivalry (Brascamp, Blake, & Knapen, 2015b; Frassle et al., 2014; Knapen et al., 2011; Lumer et al., 1998; Lumer & Rees, 1999; Zaretskaya et al., 2010) with the remainder composed of ambiguous figures (Kleinschmidt, Buchel, Zeki, & Frackowiak, 1998), or reversals in the direction of ambiguous motion (Megumi, Bahrami, Kanai, & Rees, 2015a; Sterzer & Kleinschmidt, 2007; Weilnhammer et al., 2013; Zaretskaya et al., 2010). In their recent opinion paper, Faivre et al., (2017) also reviewed the number of peer-reviewed articles investigating consciousness or awareness, and separated their results by the sensory modality of interest (Figure 1.10). While the explosion in consciousness

research after the 1990's is clear, the second most striking feature is the omission of research investigating consciousness from a non-visual perspective.

Critically however, to be conscious, requires more than being able to see. The seamless integration across multiple senses is what gives rise to our conscious experience (Alais, Newell, & Mamassian, 2010; Deroy, Chen, & Spence, 2014; Faivre et al., 2017; Nagel, 1974). Indeed, both the early (Baars, 2002), and most recent theoretical accounts of consciousness (Mudrik, Faivre, & Koch, 2014; Tononi, 2012) have stressed the multisensory nature of our everyday experience.

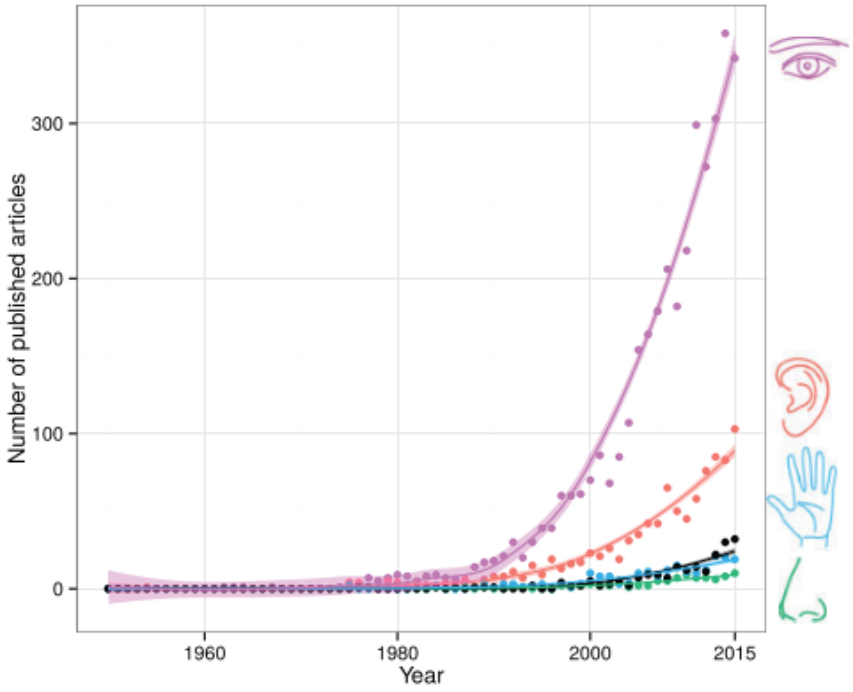

Figure 1.10: Comparison of visual to non-visual studies of consciousness.

Research investigating "awareness" or "consciousness" has exploded since the 1990's with a heavy focus on the visual modality (from Faivre et al., (2017)).

1.4.1. Multi-sensory integration: an opportunity to strengthen the effects of attention during binocular rivalry

Investigations to dissociate the tightly linked processes of attention and visual consciousness have yielded inconsistent results. Key features of concern include whether or not attention and consciousness are dissociable, and at which stage of

processing their influences may diverge. As detailed above, early work in this space has investigated how attention interacts with visual representations that are outside of conscious awareness during inter-ocular rivalry. However, the effects of attention during periods of inter-ocular competition are weak (Dieter, Brascamp, et al., 2016; Meng & Tong, 2004 **Chapter 1.4**), meaning it may be difficult to dissociate any potential effects from those of visual awareness.

By contrast, attention has a strong impact on the consequences of multi-sensory integration (Alais, Newell, et al., 2010; Alsius, Navarra, Campbell, & Soto-Faraco, 2005; Degerman et al., 2007; Mozolic, Hugenschmidt, Peiffer, & Laurienti, 2008; Mozolic, Joyner, et al., 2008; Shinn-Cunningham, 2008; ten Oever et al., 2016). Attending to crossmodal stimulation can impact on the integration of multisensory stimuli into a coherent percept when paired with ambiguous visual stimuli (reviewed in Klink, van Wezel, & van Ee, 2012). For example, while being aware of a suppressed visual stimulus alternative is insufficient to trigger perceptual switches during binocular rivalry (Paffen & Alais, 2011), attending to crossmodal and multisensory information has been shown to greatly increase the dominance durations of a congruent visual stimulus (Deroy et al., 2014; Klink et al., 2012). Dominance durations during rivalry have increased when visual stimuli flicker at a temporal frequency that is matched to a synchronous auditory tone (Guzman-Martinez, Ortega, Grabowecky, Mossbridge, & Suzuki, 2012; Kang & Blake, 2005; Lunghi, Morrone, & Alais, 2014), and when visual and auditory cues share direction-of-motion information (Conrad, Bartels, Kleiner, & Noppeney, 2010). Semantically paired olfactory stimulation can also alter rivalry dynamics (Zhou, Jiang, He, & Chen, 2010), as can ecological audio-visual pairings (Chen, Yeh, & Spence, 2011; van Ee, van Boxtel, Parker, & Alais, 2009). Congruent tactile sensory input also increases dominance durations when matched in either temporal (Lunghi et al., 2014), or spatial-frequency (Guzman-Martinez et al., 2012; Lunghi, Binda, & Morrone, 2010) with a visual stimulus (Figure 1.11).

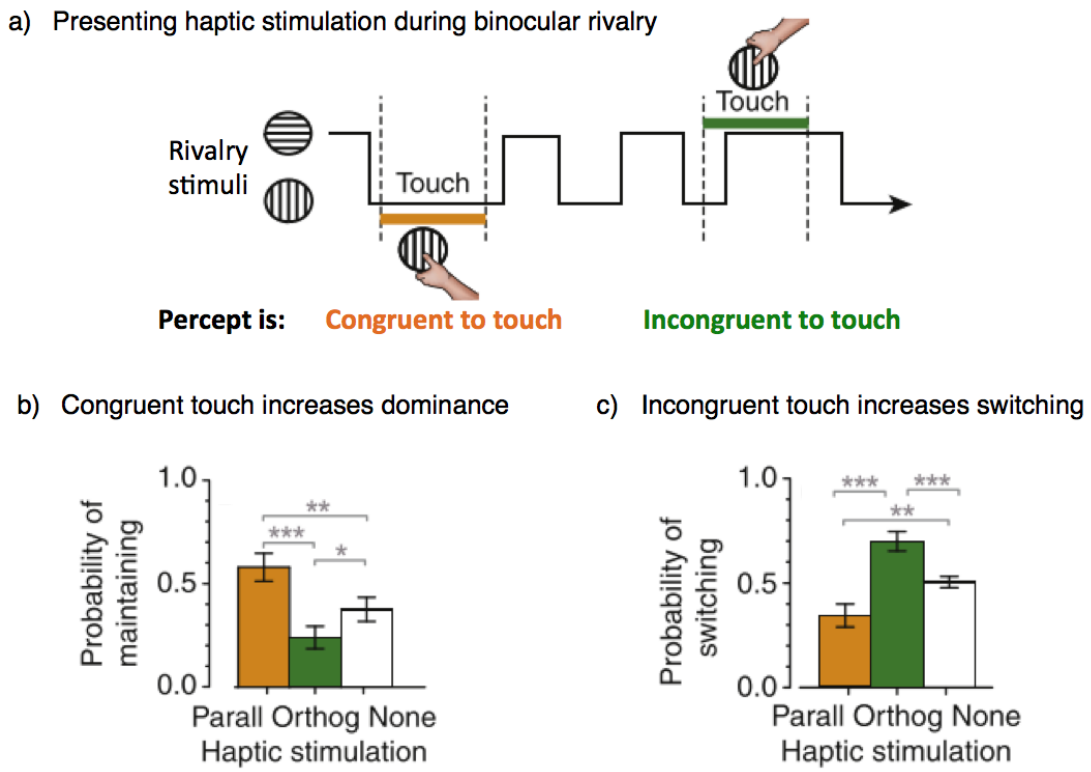

Figure 1.11: Attending to crossmodal information can determine the content of consciousness during binocular rivalry.

a) During binocular rivalry, spontaneous perceptual reversals can be recorded while intermittently presenting crossmodal stimulation that is either congruent or incongruent with the dominant visual stimulus (shown by the black arrow). In this example, participants explored a tactile stimulus with groove orientations that matched one of the visual stimuli. b) Tactile stimulation matched to the dominant visual stimulus increased maintenance. c) Incongruent haptic stimulation increased the probability of switching to the alternate visual stimulus. Adapted from Lunghi et al., (2010).

Common to all these examples is a necessary match between visual information and the crossmodal input, with the result being that matched visual and crossmodal information exert a powerful influence on the normally stochastic dynamics of binocular rivalry. Critically however, only three have independently manipulated selective attention during crossmodal rivalry (Alais, van Boxtel, et al., 2010; Chen et al., 2011; van Ee et al., 2009)[3], and all were behavioural investigations

[3] Lunghi and Alais, (2013) also compared the effects of visuo-tactile pattern matching during active and passive conditions. However, active conditions involved exploring a tactile stimulus with the forefinger, and passive conditions left the hand at rest. As such, selective attention was not independently manipulated per se, as the act of exploration was an additional confound.

without recording neural activity. When attention has been explicitly manipulated, crossmodal effects on the contents of consciousness have been more pronounced. For example, van Ee et al. (2009) presented one looming visual pattern that expanded radially, and another with constant rotation to each eye in their binocular rivalry experiment. At the same time, matched sounds were played which competed for attention, either a stationary tone (matched to the constant rotation), or a looming tone which pulsed at the same rate as the visual display. When attempting to hold the looming visual display perceptually dominant, participants were successful only when attending to the synchronous and matched looming tones. When passively listening to the tones, or attending to the incongruent stimulus features, no change in visual perception emerged (van Ee et al., 2009). The authors interpreted this result as evidence for attention boosting the perceptual gain of the matched stimuli, arguing that the mechanisms of multisensory integration and feature-based attention are tightly linked (van Ee et al., 2009; see also Degerman et al., 2007). In other words, with an appropriate feature-based match between multisensory stimuli, attention enhances the matching visual information, which comes to dominate the contents of consciousness.

Chen and colleagues (2011) also independently manipulated selective attention in their experiment, investigating the effects of semantically matched auditory cues on binocular rivalry. Their results demonstrated that even when passively viewing the audio-visual combination, crossmodal effects still emerged. The authors concluded that the effects of attention and auditory context had additive, although distinct effects, which could be dissociated in their analysis (reported in DeRoy et al., 2014). Selective attention was found to modestly prolong the dominance durations of a congruent stimulus, consistent with previous binocular rivalry investigations (Chong et al., 2005; Dieter & Tadin, 2011; Ooi & He, 1999). The presentation of auditory context, by contrast, decreased the duration of an incongruent percept, suggesting a distinct mechanism to that of selective attention (Chen et al., 2011; Deroy et a., 2014).

This additive effect of attention and crossmodal context is further supported by a series of studies by Lunghi and colleagues (2010; 2013; 2015), investigating the dependency between tactile stimulation and binocular rivalry. Lunghi and Alais (2015) measured the difference in contrast detection thresholds for suppressed visual stimuli, as a measure of suppression depth during crossmodal (visuo-haptic) rivalry. It

was shown that congruent crossmodal stimuli reduced the suppression depth of a visual match, again demonstrating a mechanism for crossmodal congruency that boosts the suppressed visual signal, as opposed to lengthening the dominance durations uniquely.

The upshot of this crossmodal rivalry effect - which is dependent on the congruency between multi-sensory stimuli and the allocation of attention, is that we here have an example of how attentional effects conditionally impact upon the contents of consciousness. Furthermore, unlike traditional unisensory attempts to dissociate attention and conscious perception, we can here extend investigations of the NCCs to include multisensory changes in perception.

But do the correlates of multi-sensory integration meet the requirements of what we expect from NCCs? To complement investigations into the NCCs, I'll next review the empirical evidence for distributed multi-sensory integration in the cortex. The reviewed evidence supports the early integration of multi-sensory information, consistent with contemporary theories of consciousness that assert the importance of widespread neural integration for perception.

1.4.2. Multisensory integration occurs over distributed regions of cortex

As discussed above, contemporary theories of consciousness accept that a common neural mechanism enables consciousness - the integration of sensory information (Faivre et al., 2017; Mudrik et al., 2014). Whether between fronto-parietal networks and a global workspace (Dehaene & Changeux, 2011), enabled along recurrent feedback (Lamme & Roelfsema, 2000), or thalamo-cortical loops (Tononi, Boly, Massimini, & Koch, 2016), true candidate NCCs must also apply to the process of integration outside of visual-only information (Faivre et al., 2017; Koch et al., 2016). While the neural correlates of crossmodal rivalry are unknown, there is reason to believe that multisensory integration is accomplished via distributed cortical processes (Calvert, 2001; Calvert & Thesen, 2004; Schroeder & Foxe, 2005).

The neuroscience of multisensory processing has shifted from a strictly hierarchical view of isolated sensory signals which converge only in multi-modal areas (Felleman & Van, 1991; Meredith & Stein, 1983; Stein & Stanford, 2008), to a growing consensus that cross-modal integration can also occur at early stages of sensory processing (Calvert, 2001; Ghazanfar & Schroeder, 2006; Meredith & Stein,

1983; Schroeder & Foxe, 2005). These early examples of integration cast multi-sensory integration as a similar form of early, and unconscious inference, and thus potential complement to NCC investigations.

Over the last decade, evidence for the convergence of multisensory signals in early sensory areas has come from two main lines of research, emerging first in animal models. Physiological studies have shown that neural activity in traditionally unisensory regions can be altered by the presence of other-sensory stimulation (e.g. (Kayser, Petkov, & Logothetis, 2008; Morrill & Hasenstaub, 2018). For example, single-neuron responses to acoustic stimuli are enhanced within auditory cortex when presented alongside visual or somatosensory input (Bizley & King, 2008). Anatomical projections between primary sensory cortices, that bypass traditional multimodal regions have also been identified (Banks, Uhlrich, Smith, Krause, & Manning, 2011; Falchier et al., 2009; Henschke, Noesselt, Scheich, & Budinger, 2015). In humans, non-invasive EEG (Fort, Delpuech, Pernier, & Giard, 2002; Foxe et al., 2000; Giard & Peronnet, 1999; Mishra, Martinez, Sejnowski, & Hillyard, 2007; Molholm et al., 2002; Murray et al., 2005; Naue et al., 2011; Raij et al., 2010; Thorne, De Vos, Viola, & Debener, 2011) and intracranial recordings (e.g. Mercier et al., 2013) have revealed that multisensory integration occurs rapidly after stimulus presentation, within the time-frame of what is traditionally thought to be early, unimodal processing.

Correlates obtained via fMRI also supporting the emergence of multisensory processing in early sensory cortices (Calvert & Thesen, 2004; Foxe et al., 2002; Kayser & Logothetis, 2007). Owing to their superior spatial resolution, hemodynamic signals obtained via fMRI have been of key interest when localizing the correlates of multi-sensory integration. In one of the first fMRI procedures to present concurrent stimulation in multiple senses, Macaluso, Frith, and Driver (2000) were able to demonstrate increased activity in traditionally unimodal regions of the occipital cortex during synchronous vibrations. Their participants attended a screen upon which flash-events could appear in the left or right hemifield, either accompanied (or not), by synchronous vibrotactile input at the congruent location. When tactile stimulation was congruent with the flash (left hemifield and left hand), touch was found to increase the contralateral occipital response, showing a spatially-specific crossmodal integration in occipital regions. These results have since been replicated (Macaluso, Frith, & Driver, 2002), and extended, as contralateral responses in primarily tactile

regions have also been enhanced with the presence of a task-irrelevant visual flash. These results thus show that crossmodal influences emerge even in unimodal sensory regions, in support of distributed cortical integration at the proposed heart of consciousness.

Crossmodal interactions have also been demonstrated with much higher temporal resolution using ERP measures, also investigating the effects of multisensory stimulation on visually evoked responses. Task irrelevant tactile cues when presented at synchronous locations increase visual N1 ERP components, emerging ~140 ms after visual stimulus onset (Kennett, Eimer, Spence, & Driver, 2001). The rapid nature of these effects was interpreted as an early crossmodal impact on traditional unisensory processing. Similarly, whether auditory and visual stimuli are cued on ipsilateral or contralateral sides of the environment has been shown to enhance early negative visual ERP responses between 120-170 ms after cue onset (McDonald, Teder-Sälejärvi, Russo, & Hillyard, 2003). Similar effects of multisensory interactions have been demonstrated in even earlier time-windows (40-46 ms; Giard & Perronet, 1999), consistent with multisensory interactions occurring at the earliest stages of sensory processing (reviewed in De Meo et al., 2015).

Together, neuroimaging and ERP analyses present converging evidence that multisensory integration can occur early, in a distributed network of traditionally unisensory regions. These early examples of neural integration offer enticing opportunities to investigate the NCCs in the context of multisensory stimulation, which may provide evidence in favour of, or against prominent theories of consciousness. For example, might the emergence of a multi-sensory percept coincide with the ignition of fronto-parietal cortex, as per GWT; or correlate with early sensory and hierarchical stages of cortical information flow, as per recurrent processing? We return to these motivations in **Chapters 3 and 4**, which investigate the effects of selective attention during crossmodal rivalry in the EEG. Next, we turn to the motivations for the second empirical axis of this thesis, which develops paradigms to distinguish the contents of awareness from the focus of attention, *outside* the lens of binocular rivalry.

KEY POINTS

- Consciousness is rarely unimodal, and a true NCC must also apply to cases of multi-sensory integration.
- Binocular rivalry is affected by multi-sensory information, which increases the effects of attention.
- Multi-sensory processes have been implicated in distributed regions of the cortex, consistent with contemporary theories that widespread integration is a necessary prerequisite of consciousness.

1.5. Conceptual advantages of perceptual-filling in

Though binocular rivalry can be relied upon to instigate perceptual suppression, the inter-ocular competition between percepts provides additional constraints (Blake et al., 2014; Sterzer et al., 2009). By contrast, alternate multistable stimuli can be viewed with both eyes simultaneously, and still result in temporal fluctuations of a coherent percept or stimulus feature. In particular, while the effects of attention during unimodal binocular rivalry may be modest, and the rivalry transitions difficult to emulate, a class of multistable phenomena involving the disappearance of peripheral visual stimuli are here argued to confer additional benefits.

1.5.1. Beyond binocular rivalry with filling-in phenomena

By definition, to suppress a visual image from awareness during binocular rivalry requires pronounced and sustained inter-ocular competition. As a consequence, the neural correlates and mechanisms of rivalry concern whether perceptual alternations result from inter-ocular competition, competition between stimulus representations, or some hybrid of the two (**Chapter 1.1.1**). Necessary concerns have also been raised as to whether the peculiar case of binocular rivalry can be generalized to the neural correlates of consciousness more generally (Blake et al., 2014), or even to other forms of multistable images (Dieter, Brascamp, et al., 2016).

As an alternative, filling-in phenomena are multistable events that are induced via the presentation of a salient visual stimulus to a peripheral location (Devyatko, Appelbaum, & Mitroff, 2016; Komatsu, 2006; Ramachandran, 1992, 1993; Ramachandran, Gregory, & Aiken, 1993; Troxler, 1804). If gaze can be maintained and withheld at a central fixation point, peripheral targets disappear in and out of awareness in a manner similar to binocular rivalry (Carter & Pettigrew, 2003; De Weerd, 2006; New & Scholl, 2008; Ramachandran et al., 1993). For example, Troxler fading occurs when a salient peripheral target disappears to match a uniform background (Troxler, 1804; Figure 1.12). These disappearances can be facilitated by superimposing a mask of moving distractors to conflict with the target, known as Motion-Induced Blindness (MIB; Bonneh, Cooperman, & Sagi, 2001; Bonneh et al., 2013). Similar to MIB, peripheral disappearances also increase with the inclusion of a dynamic background that updates continuously around each target (De Weerd, 2006; L Pessoa, Thompson, & Noë, 1998; Ramachandran et al., 1993). This latter process is known as perceptual-filling in (or an artificial scotoma; Anstis, 2010; Anstis & Greenlee, 2014; Pessoa et al., 1998; Ramachandran et al., 1993; Weil, Kilner, Haynes, & Rees, 2007) and is the basis for the empirical investigations in **Chapters 5 and 6**. Figure 1.12 displays these paradigms for comparison.

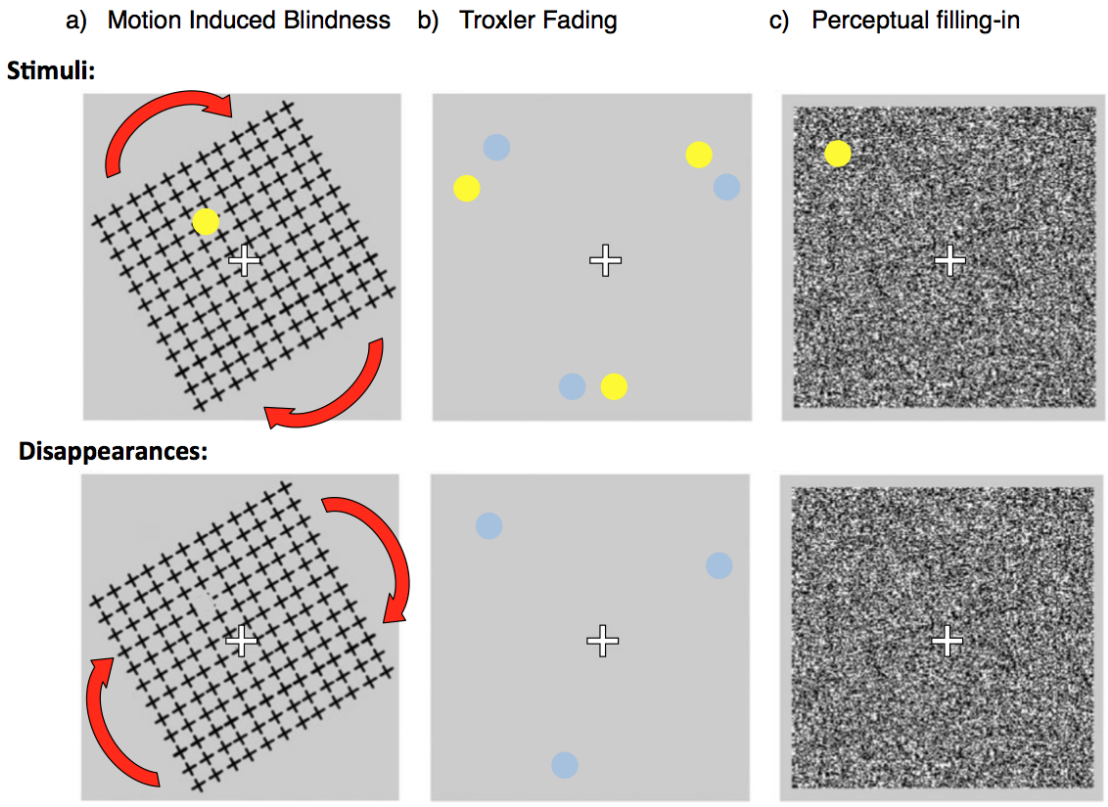

Figure 1.12: Peripheral fading paradigms.

Top row: In all cases, participants are asked to fixate on the central white cross, while covertly monitoring their awareness of salient peripheral targets (yellow discs). a) In Motion-induced blindness a moving mask increases disappearances. b) During Troxler fading, attending to a colour or feature increases disappearances. c) Perceptual filling-in occurs when the target is replaced by surrounding image background. Note in all cases, the yellow discs are always physically present, but disappear from awareness. Adapted from Devyatko, Appelbaum, and Mitroff (2016).

Recently, the notion that disappearances during Troxler fading, MIB, and PFI result purely from low-level adaptation mechanisms has come into question (for that account, see Bonneh, Donner, Cooperman, Heeger, & Sagi, 2013; Ramachandran et al., 1993; Rogers-Ramachandran & Ramachandran, 1998). What makes these paradigms particularly attractive prospects in the context of consciousness research is that unlike binocular rivalry, which is only modestly impacted by selective attention, attention can profoundly impact on the visibility of these multistable phenomena. What's more, attention appears to have a counter-intuitive effect. As attention to peripheral targets increases, visibility decreases, offering a potentially powerful

complement to existing paradigms searching to distinguish between the mechanisms of attention and visual awareness.

1.5.2. Attending to peripheral stimuli decreases their visibility

In the case of Troxler fading, the inhibitory effects of attention remained under-acknowledged for almost 40 years (Babington-Smith, 1961). As identified by Lou (1999), directing attention to a select feature of peripherally presented targets increases their rate of perceptual disappearance. In her original experiment, Lou presented 16 undergraduate participants with spatially discrete coloured discs on a uniform grey background. By alternating the colour of discs between orange and green, subjects attempted to perceive the like-coloured discs to form an equilateral triangle, and report the disappearance of one-colour over the other via button press until it returned. A clear inhibitory effect of attention was observed, with discs fading more frequently when attended, and indeed for a subset of participants, only fading when selectively attended (n=5). Lou interpreted these results as a demonstration that attention does not always enhance perception, here acting to modulate sensory processes in presumably early visual pathways, biasing the contents of consciousness.

Compared to targets which disappear from a uniform background during Troxler fading, peripheral targets that conflict with a moving mask disappear at a higher-rate, and for longer duration, during MIB (Bonneh et al., 2001, 2013). Like Troxler Fading, increased attention has also been shown to increase the disappearance of peripheral targets during MIB (Geng, Song, Li, Xu, & Zhu, 2007; Schölvinck & Rees, 2009). In the first example, Geng et al., (2007) presented two targets over a rotating display that were collocated in either the lower or upper visual field, and separated left and right of fixation. The authors compared disappearance rates when either dividing attention to both targets or selectively attending to the target at left or right of fixation. Again, the results were clear, selectively attending to a target either left or right of fixation significantly increased all measures of disappearance. Schölvinck and Rees (2009) confirmed these results by manipulating attention in two separate ways. First, they presented two targets in the upper visual field as per Geng et al., (2007), and prompted attention to either the left or right target prior to stimulus onset. Schölvinck and Rees (2009) ended trials upon target disappearance, and

recorded that attended targets were significantly more likely to first disappear. In a second experiment, participants performed an attentionally demanding task at fixation while simultaneously reporting on the disappearance of peripheral targets during MIB. Under high attentional load at fixation, the rate of perceptual disappearances was at their lowest, indicating that the allocation of attention away from a relevant target will also inhibit disappearance. Taken together, these results show that the allocation of selective attention can both increase the duration and probability of target disappearance (Geng et al., 2007; Schölvinck & Rees, 2009), as well as reduce the rate of perceptual alternations when attention is diverted (Schölvinck & Rees, 2009).

The use of dual-targets in these designs has also indicated a bias favouring MIB in the left hemisphere, as well as to the lower visual field in the case of Geng and colleagues (2007). In other words, MIB appears to operate non-uniformly across visual quadrants. In **Appendix 1**, this hypothesis was tested explicitly by examining disappearance rates during MIB across all four visual quadrants, under mask parameters approximating self-motion.

Selective attention also increases perceptual-filling in (PFI), whereby a textured background interpolates regions of the visual environment that were previously occupied by salient targets. Extending the design of Lou (1999) by presenting six figures over a textured background, De Weerd, Smith, and Greenberg (2006) varied the colour or shape of spatially discrete targets to measure the influence of selective feature-based attention. These authors specifically tested whether attention decreased the time to first disappearance, or overall probability of filling-in to occur. Consistent with the earlier results of Lou (1999) in Troxler fading, attending to the shared colour, shape, or location among targets increased the probability of becoming interpolated by the surrounding background texture. Notably however, the time until first filling-in was unaffected by selective attention. As selective attention neither increased nor decreased response times, it seemed attention had uniquely increased the interpolation processes that remove an object from visual awareness (de Weerd et al., 2006).

In an effort to disambiguate the effects of attention and consciousness, these classes of peripheral fading and invisibility appear as worthy candidates. In addition to the contradictory effects of attention on visibility, the disappearance of peripheral stimuli can also be mimicked with relative ease, in stark contrast to the physical

replay conditions often used in binocular rivalry. For example, to compare the neural activity evoked by genuine disappearances during PFI and physical replay, Donner, Sagi, Bonneh, and Heeger (2008) recorded the timing of MIB at the subject level. After genuine MIB trials, the timing of disappearances was replayed, with clear differences in the neural activity evoked by perceptual compared to physical target removal. Unlike replay conditions during binocular rivalry, which have evoked similar neural activity that is difficult to disambiguate from perceptual reversals (**Chapter 1.3.2**), activity in the ventral visual processing stream (V4), initially increased during replayed-target disappearances. By contrast, V4 activity during genuine target disappearances decreased relative to baseline, indicating that distinct neural mechanisms for subjective and physical target removal could be identified, even with the relatively course temporal resolution of fMRI (Donner et al., 2008)

In summary, compared to binocular rivalry, selective attention during instances of peripheral target fading appears to be dissociable from visual awareness, increasing target invisibility with increased attention. Moreover, as perceptual changes involve only a small region of the visual field, rather than the complicated and at times piecemeal transitions during rivalry, an effective replay of target disappearances can be used. Early attempts suggest the neural correlates of a genuine, endogenously produced perceptual change can be distinguished from physical replay, in contrast to earlier efforts localizing NCCs with binocular rivalry and fMRI. **Chapters 5 and 6** in the second axis of this thesis concern the development of novel PFI paradigms amenable to investigation in the EEG. Before continuing, it is worth reviewing the methodological advantages of EEG relative to fMRI, with particular focus on frequency-domain analysis and a technique consistent throughout this thesis, known as frequency-tagging.

KEY POINTS

- Troxler fading, motion-induced blindness and perceptual filling-in show attention can have adverse effects on the contents of consciousness, decreasing target visibility.
- Owing to the nature of these multistable paradigms, genuine perceptual reversals can be easily mimicked by physically removing small peripheral targets from a visual display, in contrast to the complicated transitions experienced during binocular rivalry.

1.6. Overview of general methods

1.6.1. Introduction to the Fourier transform

To analyse EEG activity in the frequency domain is to measure the contributions made to observed neural activity from separate frequency specific components. To quantify the frequency-specific activity, a series of analyses can be made based on the assumptions of the French mathematician, Joseph Fourier (Cohen, 2014; Gross, 2014) These assumptions are that any time-varying signal can be decomposed and represented as the sum of sine waves at different frequencies, a process which has come to be known as the Fourier transform (Cohen, 2011, 2014; Hanslmayr et al., 2011).

Fourier transform is used to backwards-infer from an observed EEG signal the relative strength of separate frequencies at any point in time (Figure 1.13). The mathematical details of the analyses can be adapted to experimental advantage, and are consequently presented in the Methods section in the later empirical chapters. From the Fourier analysis, two types of information can be extracted that describe time-frequency dynamics: the power with which an oscillation of a particular frequency is present, and the phase or position along the sine wave of the oscillatory cycle. These power and phase dynamics can be observed in resting-state activity, or observed for changes after the presentation of a transient stimulus.

As well as transient responses, frequency-specific neural activity can also be tracked during continuous stimulus presentation. The continuous, periodic modulation of a stimulus is used throughout the empirical chapters of this thesis to increase frequency-specific neural activity, a method which is known as the flicker effect, or frequency-tagging (Lansing, 1964; Tononi, Srinivasan, et al., 1998a).

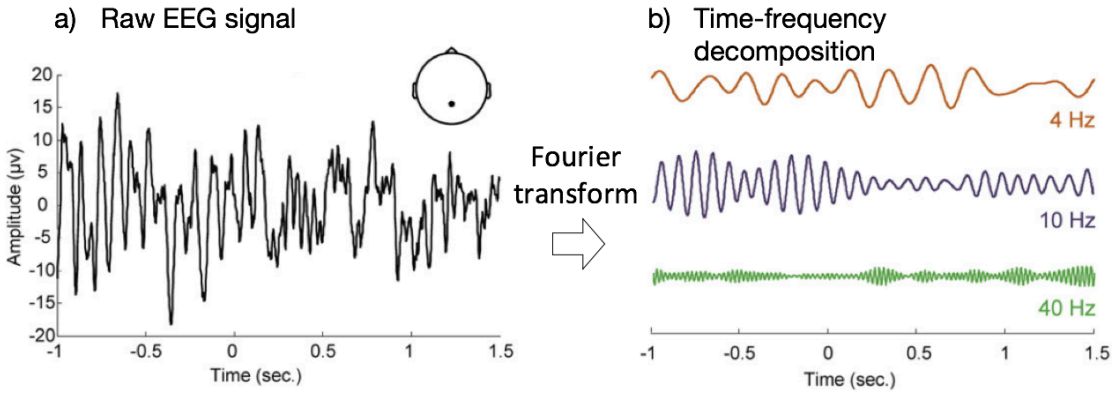

Figure 1.13: Time-frequency decomposition.

a) A raw EEG signal in the time-domain. This time-series contains information about neural activity at various oscillatory frequencies. b) Applying the Fourier transform decomposes the time-series signal, allowing the relative strength in different frequency bands to be quantified. Adapted from Hanslmayr et al. (2011).

1.6.2. The practical advantages of frequency-tagging

Argued to be an underutilized tool in cognitive electrophysiology (Cohen, 2014), frequency-tagging refers to the production of brain activity that corresponds to the rhythmic quality of an externally presented stimulus (Norcia, Appelbaum, Ales, Cottereau, & Rossion, 2015; Regan, 1977; Vialatte, Maurice, Dauwels, & Cichocki, 2010). The produced brain activity is known as a steady-state response, as it is oscillatory in nature and continuous, with a frequency 'tag' observable in the spectral domain after Fourier analysis (Figure 1.14). This frequency-tag is commonly referred to as the steady state visually evoked potential (SSVEP) when in response to a periodically flickering visual stimulus. What makes frequency-tagging so powerful is that the entrainment of neural populations responding to a periodic stimulus also produces narrow-band, stimulus-specific physiological responses (Norcia et al., 2015; Vialatte et al., 2010). This allows for recorded EEG data to be less susceptible to broadband blink and eye-movement artefacts, and for stimulus-specific processing to be isolated from background neural activity (Luck & Kappenman, 2011; Norcia et al., 2015; Perlstein et al., 2003; Vialatte et al., 2010).

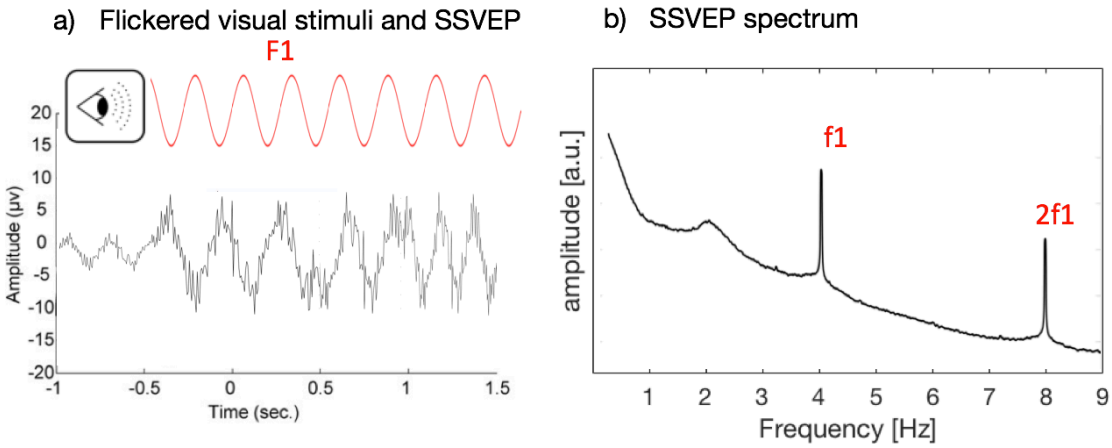

Figure 1.14: Example of the flicker effect producing an SSVEP.

a) A flickering visual stimulus will entrain EEG responses at the same stimulus-locked frequency (F1; the SSVEP). b) After Fourier transform, the amplitude of neural activity at the stimulus-flicker frequency (f1) has a high signal-to-noise ratio (amplitude compared to surrounding frequencies). A linear multiple of the driving flicker frequency is also present (known as harmonic; 2f1). Due to the frequency-specificity nature of the SSVEP, stimulus specific processing can be isolated from background neural activity with ease.

Due to the stability of SSVEPs over time, frequency-tags also have a high signal to noise ratio, making them ideal paradigms to follow temporal fluctuations in stimulus processing (Andersen, Hillyard, & Müller, 2008; Müller, Picton, et al., 1998; Müller, Teder-Salejarvi, & Hillyard, 1998), or for multistable stimuli that are presented for several seconds at a time (Parkkonen, Andersson, Hämäläinen, & Hari, 2008; Srinivasan & Petrovic, 2006; Tononi, Srinivasan, et al., 1998a). Frequency-tagging can also increase the spatial resolution of EEG as a consequence of the high signal-to-noise ratio, and narrow-band response, as neurons preferential to a flickering stimulus increase their activity at the a priori specified frequency (Silberstein, Ciorciari, & Pipingas, 1995; Silberstein et al., 1990a; Wu & Yao, 2007). Current evidence suggests that stimulus flicker can entrain neural activity beyond early sensory cortices (Cohen, 2014; Srinivasan, Fornari, Knyazeva, Meuli, & Maeder, 2007), including higher-order parietal and frontal sites (Ding, Sperling, & Srinivasan, 2006; Srinivasan, Bibi, & Nunez, 2006), though the exact sources of such activity are not well characterised (Herrmann, Strüber, Helfrich, & Engel, 2015; Kamphuisen, Bauer, & van Ee, 2008).

Recently, the use of frequency-tags to investigate cognitive processes has been supplemented by a focus on the harmonics and intermodulation components of the

response (e.g. Katyal, Engel, He, & He, 2016; Kim, Grabowecky, Paller, Muthu, & Suzuki, 2007; Kim, Grabowecky, Paller, & Suzuki, 2011; Zhang, Hong, Gao, Gao, & Röder, 2011; Zhang, Jamison, Engel, He, & He, 2011). Harmonics are linear multiples of the fundamental driving frequency, also observable in the EEG response spectra (e.g. 2 x flicker frequency; 2f1, 3f1, etc.). While the harmonic profile of steady-state responses can vary as a function of flicker parameters (Norcia et al., 2015; Vialatte et al., 2010), a difference between harmonics has also been demonstrated based on cognitive factors, despite being elicited by the same driving stimulus (e.g. Gulbinaite, Roozendaal, & Vanrullen, 2019; Kim et al., 2007, 2011). Specifically, the second harmonic, or frequency double of visually evoked response has been shown to increase in amplitude with selective attention, as well as spread to the contralateral hemisphere when attending to either left or right visual hemifield targets (Kim et al., 2007; 2011). One possible interpretation for this phenomenon is that neural populations in low-level visual areas primarily respond to basic stimulus attributes, such as a change in contrast, whereas the second harmonic represents higher-levels of visual processing accomplished after the aggregation of early sensory signals (Kim et al.,2007; 2011). This interpretation is supported by separate non-overlapping anatomical sources for the first and second harmonic (Pastor et al., 2002), and observations that neural responses in higher visual areas respond most strongly to attentional modulation (Corbetta & Shulman, 2002; Luiz Pessoa, Kastner, & Ungerleider, 2003). At present however, the exact underlying neural mechanism for these differences between first and second harmonics are unknown (Kim et al., 2007; 2011).

When more than one driving frequency is present simultaneously, intermodulation components (IMs) also emerge in response spectra (e.g. increased power at frequency 1 + frequency 2; f1+f2). These IM components are argued to represent *non-linear* interactions in neural processing resources, and argued to be a proxy for the integration strength of separate stimulus features (Gordon, Koenig-Robert, Tsuchiya, Van Boxtel, & Hohwy, 2017; Katyal et al., 2016; Zemon & Ratliff, 1984). For example, presenting spatially non-contiguous pacman-like figures can produce the illusory perception of a shape, which is bound by each pacman figure. By flickering each pacman figure at a distinct frequency (e.g. F1 = 2.94 Hz, F2 = 3.57 Hz), IM components (e.g. F1 +F2 = 6.51 Hz) were shown to increase when an illusory rectangle was perceived – which only could occur if the information from the

non-contiguous flickering objects were combined (Alp, Kogo, Van Belle, Wagemans, & Rossion, 2016; Gundlach & Müller, 2013). By contrast, rotating the pacman like figures to destroy the illusion of shape, and thus the integration between these features, also removes the IM response. Similar examples have arisen for the holistic perception of faces which require the integration of two flickering face-halves (Boremanse, Norcia, & Rossion, 2013), and interaction between semantic and lower visual features (Gordon et al., 2017).

KEY POINTS

- Time-frequency decomposition of EEG data via the Fourier transform allows the power and phase of frequency-specific activity to be investigated.
- Fourier transform can be applied with frequency-tags, which result from periodic modulation of a stimulus.
- Frequency-tags index stimulus-specific processing, can increase the spatial resolution of EEG, and may offer measures of integration and attention via the IM and harmonic components, respectively.

1.7. Steady state responses in consciousness research

1.7.1. Frequency-tagging and attention

Frequency-tags in the visual modality (SSVEPs) have been a powerful tool to study the mechanisms of spatial attention (reviewed in Norcia et al., 2015). Since spatial attention to a flickering stimulus was first shown to enhance SSVEP responses (Morgan, Hansen, & Hillyard, 1996), both covertly and overtly attending to flickering visual features has been shown to increase SSVEP amplitude (Andersen et al., 2008; Müller et al., 2006; Müller, Picton, et al., 1998; Müller, Teder-Salejarvi, et al., 1998; Walter, Quigley, Andersen, & Mueller, 2012), and phase coherence (Kim et al., 2007; Zhang et al., 2011). An increase in SSVEP responses has also demonstrated that attention can be divided in parallel over spatially separated locations of a screen (Müller, Malinowski, Gruber, & Hillyard, 2003). While attending to select features in a spatially overlapping display, which flicker at distinct frequencies, has shown that attention can enhance feature-specific responses even when spatially overlapping (Andersen et al., 2008; Pei, Pettet, & Norcia, 2018).

A smaller group of studies have also examined the effects of attention in non-visual frequency-tagging paradigms. Steady-state auditory evoked potentials (SSAEPs) have primarily been utilized in clinical audiology literature (Plourde, 2006; Tlumak, Durrant, Delgado, & Boston, 2012; Tlumak, Rubinstein, & Durrant, 2007). Both the SSAEP and its somatosensory counterpart (SSSEP), have been shown to increase with selective attention (Giabbiconi, Dancer, Zopf, Gruber, & Müller, 2004; Giabbiconi, Trujillo-Barreto, Gruber, & Müller, 2007; Ross, Picton, Herdman, Hillyard, & Pantev, 2004; Tiitinen et al., 1993). SSVEPs and SSAEPs recorded during the simultaneous presentation of visual and auditory stimuli also increase in amplitude (Budd & Timora, 2013; Covic, Keitel, Porcu, Schröger, & Müller, 2017; de Jong, Toffanin, & Harbers, 2010; Porcu, Keitel, & Müller, 2013, 2014), an additive gain which has been argued to index multisensory integration (Covic et al., 2017; Keil & Senkowski, 2018; Keitel & Müller, 2015; Nozaradan, Peretz, & Mouraux, 2012; Zhang et al., 2011).

Frequency-tagging is thus a useful mechanism to capture stimulus specific neural activity, and increase the signal to noise ratio for the effects of attention. In the context of research into the NCCs, frequency-tagging non-visual information offers the additional opportunity to explore the mechanisms of multi-sensory integration.

1.7.2. Frequency-tagging the contents of consciousness

SSVEPs have been combined with multistable figures to investigate the spatial and temporal correlates of multistable perception. The earliest indication that frequency-tagged activity could be used to indicate the contents of consciousness was provided by Lansing (1964). Using two electrodes over occipital scalp, Lansing observed that increased oscillatory amplitude followed increases to the contrast of a flickering monocular image. Subsequently, the suppression and enhancement of percept-modulated SSVEPs was observed over occipital sites (Cobb, Morton, & Ettlinger, 1967; Lawwill & Biersdorf, 1968), before Brown and Norcia (1997) recorded real-time fluctuations in binocular rivalry, by tracking the SSVEP amplitudes of each separately flickering, and thus uniquely frequency-tagged image (Figure 1.15).

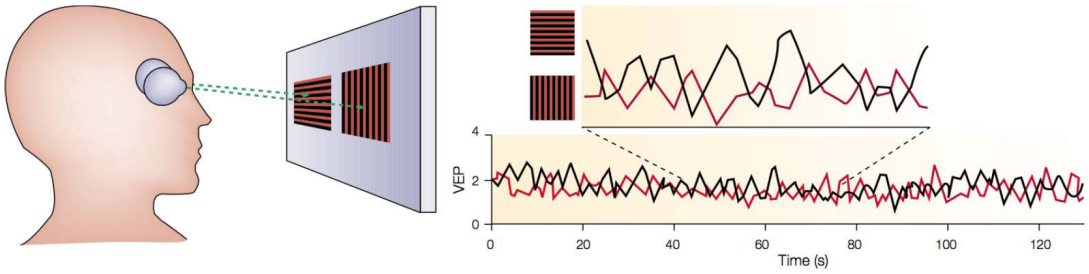

Figure 1.15: Frequency-tagging the contents of consciousness.

During Binocular rivalry, flickering each monocular image at a distinct frequency enables the contents of consciousness to be tracked from neural activity. Critically, the amplitude of these activity fluctuations tracks subjective dominance reports. Adapted from Norcia et al. (2015).

Following this early work, the use of whole-head M/EEG recordings has led to conflicting accounts regarding the frequency-tagged neural correlates of binocular rivalry (e.g. Kamphuisen et al., 2008; Tononi, Srinivasan, et al., 1998). For example, an increase in SSVEP power and phase coherence over the entire scalp has been cited in support of widespread integration, at the heart of contemporary theories of consciousness (Cosmelli et al., 2004; Srinivasan & Petrovic, 2006; Srinivasan, Russell, Edelman, & Tononi, 1999; Tononi, Srinivasan, et al., 1998a). Although these results were taken as evidence that perception correlates with increases in coordinated activity, subsequent research casts doubt over the validity of such claims (Kamphuisen et al., 2008). While Kamphuisen et al (2008) were able to replicate the emergence of stimulus-entrained activity across the entire scalp, they cautioned that the previously reported fronto-parietal NCCs may in fact simply be generated by the transmission of activity from occipital sources. Strengthening their claim, the authors were able to show that the phase-coherence of SSVEP activity in frontal regions was shifted by exactly 180 degrees relative to the simultaneous response in occipital regions. The widespread distribution of SSVEP activity over occipital and frontal regions was then parsimoniously explained to result from the same occipital dipole, rather than a heterogenous network of cortical sources.

More recently, the intermodulation components (IMs) evoked via the interaction of two simultaneous flickers have also become the focus of investigations during binocular rivalry. Consistent with models of binocular rivalry positing early

inter-ocular conflict, intermodulation signals were observed to be strongest over occipital electrode regions (Katyal et al., 2016). The strength of these IM responses was also shown to be highest during periods when a mixture of each eye's image was perceived, further implicating neural populations that detect inter-ocular conflict prior to the resolution of perceptual ambiguity. However, in a contrary result, Zhang et al. (2011), reported an increase in IM power during periods of binocular fusion, in the absence of rivalry. In their study Zhang et al manipulated the allocation of attention toward rivalling images by requiring observers to perform an attentionally demanding task at fixation. Visual regions around this central task could engage in rivalry, with each eye's parafoveal image flickering at a unique frequency. When attending to the central task, IM power was higher than when viewers could actively attend to ongoing rivalry patterns. In a follow-up experiment, Zhang et al emulated a mixture of each eye's image in these parafoveal regions, which could only have occurred in the absence of binocular rivalry. IM power was again high, which the authors took as evidence that the IMs recorded in their non-attend condition were evidence for an absence of rivalry, in the absence of attention (see also Brascamp & Blake, 2012). Although promising, these conflicting accounts for the relevance of IMs during rivalry - either reflecting inter-ocular conflict (Katyal et al., 2006) or lack thereof (Zhang et al., 2011), require clarification.

Subsequent research has extended the use of SSVEPs beyond binocular rivalry to study other correlates of perception. In the absence of inter-ocular suppression, SSVEPs have demonstrated sensitivity to emergent semantic content in phase scrambled images (Boremanse et al., 2013; Gordon et al., 2017; Kaspar, Hassler, Martens, Trujillo-Barreto, & Gruber, 2010; Koenig-Robert & VanRullen, 2013), as well as the perceptual dominance of multistable figure/ground displays (Brooks & Palmer, 2010) and Rubin's face/vase illusion (Parkkonen et al., 2008). Notably in all these cases, an enhanced SSVEP was present in early visual areas, providing evidence the neural correlates of object representations and perceptual rivalry can be tracked to specific cortical loci.

KEY POINTS

- Steady-state responses are enhanced by attention, and can be entrained by both visual and non-visual sensory stimulation.
- This enhancement of steady-state responses is sensitive to the spatial and feature-based allocation of attention.
- SSVEPs have been combined with binocular rivalry and other multistable phenomena, and shown to positively correlate with the contents of consciousness.

Chapter 2: Introduction to empirical contributions

2.1 Outstanding questions

A long history of enquiry has framed the contents of consciousness as the end result of constructive and inferential processes. When viewing multistable stimuli we are dramatically reminded of these processes, as the contents of consciousness change despite an unchanged physical object. To investigate which brain regions may mediate these changes in consciousness, visually multistable images have been combined with fMRI and EEG recordings, to spatially and temporally localize the NCCs.

Through the careful review of these paradigms, new concerns have emerged that motivate the empirical contributions of this thesis. These regard whether the most commonly used multistable illusions are best suited to distinguish the contents of consciousness from focus of attention, and whether these processes may be functionally dissociated. More specifically, the present work was motivated to explore how multisensory and non-visual information may contribute to our understanding of consciousness during binocular rivalry, as well as develop novel paradigms that may differentiate the overlapping phenomena of attention and conscious perception. The empirical chapters that follow progress through the following specific questions:

2.1.1. How are attention and consciousness related?

The link between attention and consciousness has begun to be explored from a novel vantage point, with evidence to suggest that though they are tightly linked, attention and consciousness may be two distinct brain processes (Koch & Tsuchiya, 2007; van Boxtel et al., 2010a, 2010b). Given this possibility, it is necessary to examine how attention is related to consciousness, and in particular the ways in which attention may (or may not) mediate the contents of conscious experience. To examine the effects of attention in new ways, several new approaches are presented in this thesis which can be grouped along two main axes of research. To begin, manipulating

attention is combined with simultaneous multisensory stimulation during a binocular rivalry paradigm. In the second, the frequency-tagging of visual information is used to develop and refine paradigms aimed at disentangling the neural correlates of attention and consciousness. These distinct approaches also generate a subset of questions which can be explored within the framework of examining the link between attention and consciousness:

2.1.2. How does attention interact with multisensory integration to mediate the neural correlates of conscious experience?

The first empirical axis of this thesis builds upon the psychophysical literature combining auditory and tactile stimulation with binocular rivalry. In particular, while the effects of attention on rivalry are modest, attending to crossmodal stimulation can induce strong effects on rivalry dynamics. I explore this intersection between crossmodal stimulation and attention, and investigate the neural correlates within electroencephalographic recordings. In particular, it was hypothesized that the effects of attention on binocular rivalry dynamics could be enhanced in the context of multi-sensory integration, enabling a novel avenue to explore the NCCs outside of the previous unimodal visual focus.

Chapters 3 and 4 explore the possibility of enhancing the effect of attention during binocular rivalry via multi-sensory stimulation, in an effort to disambiguate the neural correlates of attention from the contents of consciousness. Chapter **3** builds on the application of frequency-tagging during binocular rivalry in a combined visual, auditory, and tactile stimulation protocol. By manipulating attention during these periods of crossmodal rivalry, this approach tested the spatiotemporal correlates of increased neural activity co-occurring with changes in perception. Using the same crossmodal binocular rivalry paradigm, **Chapter 4** consists of investigating the changes in consciousness that were induced by attended crossmodal stimuli in more detail. Recently, the notion that attention is a sustained and unitary construct, like an unbroken stream, has been called into question (Landau, 2018; Landau & Fries, 2012; VanRullen, 2016b, 2016a). Converging evidence suggests that periods of high and low attentional focus leave distinct neural signatures which can be captured in both behaviour and physiological recordings (Fiebelkorn, Pinsk, & Kastner, 2018; Fiebelkorn, Saalmann, & Kastner, 2013; VanRullen, 2016b, 2018). Capitalizing on

this possibility, I go on to show that these unique attentional features are present when attending to changes in consciousness.

2.1.3. Can distinct neural correlates of attention and consciousness be captured?

As reviewed in **Chapter 1,** distinguishing the neural correlates of attention from conscious awareness may also be possible using perceptual filling-in. Compared to binocular rivalry, this class of multistable visual phenomena confer distinct advantages for the investigation of attention and consciousness. During perceptual filling-in, distinct target regions in the visual periphery become interpolated by the surrounding background, spontaneously disappearing from conscious awareness. Most notably, increased attention toward these regions has been shown to increase their disappearance. This feature is an attractive resource to investigate the interaction between attention and consciousness, as normally selective attention enhances sensory processing and sensitivity. During PFI however, attention results in a decrease in stimulus visibility, suggesting that attention and conscious awareness may be distinct neural processes.

Building upon the tradition of frequency-tagging changes to the contents of conscious perception during binocular rivalry, **Chapters 5 and 6** apply frequency-tagging to investigations of perceptual filling-in. **Chapter 5** first shows how frequency-tagging visual background information provides distinct advantages for attempts to capture the NCCs. Particular emphasis is placed on the methodological advantages of combining PFI with frequency-tagging compared to binocular rivalry. The paradigm is then further refined in **Chapter 6,** to examine the neural correlates of attended visual information as it disappears from conscious awareness.

Chapter 3: Attending to frequency-specific crossmodal cues alters binocular rivalry dynamics

Authors and affiliations: Matthew James Davidson[1, 2] *, David Alais[3], Naotsugu Tsuchiya[1, 2, *, +], Jeroen J.A. van Boxtel[1,2,4, *, +]

[1)] School of Psychological Sciences, Faculty of Medicine, Nursing, and Health Sciences, Monash University.
[2)] Monash Institute of Cognitive and Clinical Neurosciences, Monash University.
[3)] School of Psychology, The University of Sydney.
[4)] School of Psychology, The University of Canberra.
* corresponding author, lead contacts.
+ equal contribution

3.1. Article Introduction

In their experiment, Lunghi et al (2014) demonstrated that cross-modal auditory and tactile stimulation could influence the perceptual outcome of binocular rivalry. Subjects were significantly more likely to maintain a flickering visual percept if it was congruent in temporal frequency with a cross modal stimulus, and significantly more likely to switch if stimulation was incongruent with the current visual percept.

The functional combination of cross-modal signals was argued to be indicative of a temporal binding mechanism, capable of combining information across sensory modalities to disambiguate visual perception (Lunghi et al., 2014). Behaviourally, Lunghi et al (2014) were able to show that binocular rivalry could provide psychophysical evidence for the entrained cross-modal synchronization of neural activity, and that perceptual alternations are sensitive to synchronized stimulation.

As the psychophysical evidence provided by Lunghi et al revealed that a visual stimulus could enter awareness in the context of multisensory stimulation, their paradigm was chosen to explore whether attention may also mediate the strength of this effect. In particular, we were motivated to explore the neural correlates of these changes by combining frequency-tagging with an instruction to either attend or ignore crossmodal cues, and in doing so localize NCCs outside of the traditional, strictly unimodal visual-only lens. By including a manipulation of attention, we demonstrate an interaction between attention and multisensory stimulation, in support of the facilitative role that both bottom-up and top-down processes play in mediating the contents of conscious experience. In doing so we show that the strength of attention on binocular rivalry dynamics can be increased when paired with frequency-matched crossmodal cues. After demonstrating this increase in the strength of attention, we then investigate whether neural markers can be obtained to support that attention and consciousness are two distinct brain processes in **Chapter 4,** using the same data.

Both **Chapters 3 and 4** contribute to the first empirical question outlined in **Chapter 2**, namely: How does attention interact with multisensory integration to mediate the neural correlates of conscious experience? Here we show that attention affects the ITPC and log(SNR) of frequency-tagged responses during crossmodal rivalry, before examining how changes in perception are mediated by attention in more detail in **Chapter 4**.`

3.2. Abstract

When attending to an ambiguous visual scene, simultaneous crossmodal information can enhance perceptual processing when congruent with the ambiguous visual stimulus. Recent investigations have shown that this crossmodal facilitation can also determine the content of consciousness during binocular rivalry, when conflicting images normally alternate in-and-out of awareness. Here we investigated how attending to congruent auditory and tactile stimulation impacts on behavioural and neural dynamics during binocular rivalry through the use of frequency-tagging in the EEG. Across two-days of testing, participants (*N*=34) continuously reported their dominant visual percept during 24 x 3-minute binocular rivalry periods. Throughout

rivalry, each image continuously flickered at either 4.5 or 20 Hz (F1 and F2), and on separate days participants either attended or ignored simultaneous crossmodal cues (2-4 seconds duration, amplitude modulated at F1, F2). We found striking frequency-specific differences in the efficacy of frequency-tagged rivalry and crossmodal cueing effects. During visual-only rivalry, we found F1 frequency-tags strongly correlated with conscious percepts, while F2 frequency-tags either positively or negatively correlated with percepts on an individual basis. These results show that phenomenal capture via frequency-tagging may depend on frequency-specific mechanisms. Moreover, only attended F1 crossmodal cues altered binocular rivalry dynamics, breaking the normally stochastic perceptual alternations to bring a congruent visual image into awareness. This dependency on both attention and flicker-frequency exemplifies the role of both top-down and bottom-up stimulus driven factors in facilitating the binding of multisensory information into a unified conscious percept.

3.3. Introduction

Perception is an inherently multisensory process. To successfully navigate in the world, our brains make use of finite processing resources to sift through and combine incoming sensory stimulation. This process has long been recognized to be a type of unconscious inference (Brascamp et al., 2018; Von Helmholtz, 1867), with certain information distilled into meaningful conscious experience. The dramatic nature of this inferential process is evident when viewing perceptually ambiguous or multistable images, during which a constant physical stimulus is perceived to change over time (Kim & Blake, 2005). For example, when incompatible images are presented to each eye, binocular rivalry ensues, a process whereby the content of consciousness temporarily changes between each image (Alais, 2012).

Binocular rivalry has been extensively investigated to understand how a conscious percept emerges from the influence of visual features, as well as other cognitive factors (Alais, 2012; Blake, 2001) The lower-level properties of visual stimuli during rivalry have been shown to alter rivalry dynamics, as the relative dominance durations of one percept increasing with, for example, a change in image contrast (Hollins, 1980) or flicker rate (Blake & Fox, 1974; O'Shea & Blake, 1986). By contrast, the influence of cognitive factors such as attention have been less

pronounced (Chong & Blake, 2006; Dieter et al., 2015; Paffen & Alais, 2011; Paffen et al., 2006; yet see Dieter, Melnick, & Tadin, 2016, for success with extensive training). Selective attention may prolong the dominance of one percept, but generally is unable to bring a suppressed image into awareness (Blake, 2001; Dieter et al., 2015; Paffen & Alais, 2011). On balance, it appears that the ambiguous visual scene experienced during binocular rivalry is strongly affected by visual properties, and only modestly affected by attention.

Counter to this traditional account, recent studies have shown how the allocation of attention toward non-visual information can determine what is consciously perceived during binocular rivalry (Deroy et al., 2014; Klink et al., 2012). For example, matching the spatial/temporal frequency of visual images during rivalry with the frequency of simultaneously attended tactile/auditory tones can bring a previously suppressed image into awareness (Guzman-Martinez et al., 2012; Lunghi & Alais, 2015; Lunghi et al., 2010, 2014; Morrone & Lunghi, 2013). These results indicate that simultaneous multisensory information is capable of resolving visuo-perceptual competition during rivalry, bringing a previously suppressed visual stimulus into conscious awareness.

But where might this transition from a non-conscious to conscious percept be taking place? Normally, the neural correlates of perceptual multi-stability can be traced to a right-lateralized region of fronto-parietal cortex (see Brascamp et al., 2018 for review). It is currently unknown however, whether changes in awareness that are catalysed by multi-sensory integration share these correlates, which would support their role in perceptual awareness. Recent evidence has also displayed that multi-sensory integration does not occur uniquely in higher-cortical areas, but can take place at the earliest stages of sensory processing (Calvert, 2001; Ghazanfar & Schroeder, 2006; Meredith & Stein, 1983; Schroeder & Foxe, 2005). Leading theories of consciousness also differ in regard to the predicted locus of perceptual awareness, ranging from uniquely frontal, to parietal, and more distributed regions of sensory cortex (Boly et al., 2017; Odegaard et al., 2017; Tsuchiya, Wilke, Frässle, et al., 2015). To explore whether the previously reported neural correlates of consciousness are consistent with crossmodal rivalry when a previously suppressed image enters into awareness, we investigated the neural correlates of these changes via the use of steady-state evoked potentials, or frequency-tagging in the EEG.

Frequency-tagging is a powerful technique to increase the identification of stimulus-specific neural responses (reviewed in Norcia et al., 2015). By modulating stimuli at a particular frequency, neural populations are entrained to respond at the same unique frequency, with consequent increases in the power and/or phase coherence of frequency-specific activity compared to frequency neighbours. As these steady-state evoked potentials persist during stimulus presentation, and are stimulus-specific, they are ideal for capturing continuous responses to more than one stimulus. For example, by flickering each image at a unique frequency during binocular rivalry, the power of each image's frequency-tagged activity has been shown to positively correlate with the contents of consciousness over time (Jamison, Roy, He, Engel, & He, 2015; Lansing, 1964; Zhang et al., 2011). By combining visual, auditory, and tactile frequency-tagging protocols, we explored whether a change from non-conscious to conscious percept during rivalry was impacted by the allocation of attention, and whether the right-lateralized neural correlates previously implicated in visual-only multistability could be extended to crossmodal rivalry. We hypothesized that

3.4. Methods

3.4.1. Participants

Thirty-four healthy individuals (21 female, 1 left handed, average age 23 ±4.7) were recruited via convenience sampling at Monash University, Melbourne, Australia. All had normal or corrected-to-normal vision and gave written informed consent prior to participation. Monash University Human Research and Ethics Committee approved this study, and subjects were paid 15 AUD per hour of their time, over an approximate total of 5 hours.

3.4.2. Apparatus and Stimuli

Stimuli were generated using Psychtoolbox (Brainard, 1997) and custom MATLAB scripts. Each visual stimulus was viewed through a mirror stereoscope

placed at an approximate viewing distance of 55 cm from computer screen (29 x 51 cm, 1080 x 1920 pixels, 60 Hz refresh rate) with the subject's head stabilized via chin rest. Rivalry stimuli were red and green gratings displayed on a black background, with a white frame to aid binocular fusion, embedded within the wider grey background of the remaining portions of the screen. Beside each white framed image, coloured arrows indicated the direction for button press (e.g., right for red, left for green). Gratings were sinusoidal with spatial frequency of 0.62 cycles per degree, oriented ± 45° from vertical, and subtended 6.5° visual angle (240 x 240 pixels on the display). Visual stimuli were sinusoidally contrast-modulated at either 4.5 or 20 Hz using a temporal sinusoidal envelope. The phase of each grating was static throughout each 3-minute binocular rivalry block, yet shifted after each block to reduce the effects of visual adaptation. The stimulus size was chosen after piloting the largest images that could support minimal incidences of piecemeal rivalry. The very low spatial frequency of 0.6 cycles per degree and the rapid temporal modulations both favour neurons with large receptive fields and thus reduce the incidence of piecemeal rivalry. In addition, by rivalling red and green stimuli, each image had a consistent colour which helps group rivalry alternations and maintain perceptual coherence rather than piecemeal switching. We explained to participants that piecemeal percepts may occur and, in such cases, they should indicate the stimulus that was most dominant (Supplementary Figure 3.8).

For crossmodal stimuli 50 Hz carrier tones were amplitude modulated by 4.5 or 20 Hz sine waves to create digital waveforms, which were either 2, 3.1 or 4 seconds in duration. For tactile stimulation, subjects clasped a wooden ball with their left hand attached to a Clark Synthesis Tactile Sound Transducer (TST429 platinum) housed in a custom sound insulated box (Lunghi et al., 2014). Auditory stimulation was delivered binaurally through Etymotic HD5 noise reduction headphones, with ACCU-Fit foam ear tips to reduce ambient noise. Throughout this paper, we refer to stimulus flicker or amplitude modulation as either F1/F2, and denote recorded responses to stimulus flicker in lower-case (f1/f2 and harmonics).

3.4.3. Stimulus timing

Accurate stimulus timing of synchronous visual and crossmodal stimuli was ensured with a WDM-compatible, hardware-independent, low-latency ASIO driver (www.asio4all.com), which was necessary to minimize audio buffer duration to sub-millisecond intervals and reduce latency compensation. The time-course of stimulus presentation was also physically recorded in the EEG for offline analysis. Photodiodes were used to record the flicker-envelope of visual stimuli and stored as separate channels in the ongoing EEG. The waveforms for crossmodal stimulation were simultaneously sent to both the presentation hardware and external electrode channels using a digital splitter (Redback A2630 4 Channel Headphone Distribution Amplifier). Stimulus presentation lag was assessed by computing the difference between the recorded frames of trigger-codes and actual physical trace within the EEG as part of data pre-processing. We adjusted the relative timing of behavioural and EEG data accordingly as part of this analysis. In most cases, no adjustment was necessary, requiring a maximum change of 3 frames in duration on <1% of blocks across all subjects.

3.4.4. Calibration of visual stimuli

A maximum of 10 one-minute binocular rivalry blocks were performed prior to experimentation on the first day for all subjects. These blocks served to familiarize subjects with reporting their visual percepts during binocular rivalry, and to calibrate approximately equal dominance durations for the flickering stimuli in each eye. Contrast values for left/right eye, green/red colour, and low/high frequency stimulus combinations (in total, 8 combinations) were adjusted on a logarithmic scale until approximately equivalent total dominance durations were reached (between 1:1 and 1:1.5), with the additional requirement that the average perceptual duration for each stimulus was longer than 1 second. As there were 24 unique 3-minute binocular rivalry blocks on each day of experimentation, each of the 8 combinations of visual parameters was balanced across all three crossmodal conditions. Supplementary Figure 3.9. displays the results of this calibration procedure.

3.4.5. Calibration of auditory stimuli

Prior to experimentation, subjects were also tasked with equating the perceptual intensity of tactile and auditory stimulation for each low- and high-frequency condition, to achieve approximately equal phenomenological intensity across subjects and stimulus conditions. For all subjects, the amplitude of tactile vibrations was set to the same comfortable, supra-threshold level (approximately equivalent to 65 dB SPL). In the absence of visual stimulation, simultaneous auditory and tactile stimuli were then presented in a staircase procedure, with subjects adjusting the amplitude of auditory tones to match the perceived intensity of simultaneous tactile vibrations. They performed the matching task separately within low-frequency auditory tones and tactile vibrations and within high-frequency auditory tones and tactile vibrations. This calibration procedure was performed on each day of testing, to account for differences in the insertion depth of inner-ear headphones across separate days.

3.4.6. Experimental Procedure and Behavioural Analysis

Twenty-four three-minute binocular rivalry blocks were presented on each of the two separate days of testing. In each block, subjects reported their dominant visual percept during rivalry while receiving occasional crossmodal cues, which were either auditory, tactile, or simultaneous auditory and tactile. In a given three-minute block, we presented only one of the three types of crossmodal cues. The order of these blocks was randomized for each subject and each day of experimentation. In each block, 12 trials of crossmodal cues were presented. Each cue was either low (4.5 Hz) or high (20 Hz) frequency auditory and/or tactile stimulation. Six cues were presented for each frequency, with durations composed of three x 2 s, two x 3.1 s, and one x 4 s cues. To increase uncertainty of the timing of the cues, we defined three null cue periods (which we call visual-only periods, Figure 3.1) without any crossmodal stimulation for a duration of 2.6 s (the average of crossmodal cue durations). These periods were not distinguished from the continuous rivalry presentation to the participant. We also used these visual-only periods as baseline for behavioural

analysis (Figure 3.5). We randomized the order of all cues, which were separated with uniform jittering by 7-10 s ISI within each block.

Across all sessions, subjects were told to focus on accurately reporting their dominant visual percept at all times via button press. As the state of the button-press was sampled at 60 Hz, the same rate as the video refresh rate, we were able to estimate the probability and time-course of binocular rivalry dynamics over 16.7 ms intervals.

Over two sessions on separate days, subjects distributed attention between visual rivalry and crossmodal cues based on separate task instructions. On Day 1 for n= 18 or Day 2 for n=16, subjects were instructed to ignore the crossmodal cues and to focus on reporting only visual rivalry. For the other session, subjects were instructed to distribute attention across both visual rivalry and crossmodal cues. To ensure their attention was on task, these alternate days included task instructions for subjects to silently tally the number of times the temporal frequency of an attended crossmodal cue matched that of their dominant visual percept at the time of crossmodal cue's offset. Due to the varied duration of crossmodal cues, this task ensured that attention was allocated consistently throughout the presentation of crossmodal cues. To familiarize subjects with these task demands, an additional two practice blocks (three minutes each) were included during the calibration procedure on the relevant day of experimentation. Although 34 subjects were retained for final analysis, five others were recruited and began the experiment, yet failed to complete their second day of experimentation. One other subject was removed due to their failure in following task instructions and excessive movement during EEG recording.

3.4.7. Evaluation of attention-on-task

To evaluate the attentional allocation to both visual and crossmodal stimuli, at the end of each 3-minute block we asked subjects to verbally report their subjective estimate of the number of crossmodal stimuli which were matched in temporal frequency to the flicker of their dominant visual percept at the point of attended-crossmodal cue offset. Then, we defined an index, 'attention to cues' (Figure 3.5b, x-axis) as the correlation coefficient between 24 subjective estimates (one per attended block) and the actual recorded occurrences of congruent stimuli. Supplementary

Figure 3.10 displays the correlation between subjective and actual congruent stimuli for a representative subject.

3.4.8. Behavioural data analysis

We pre-processed the button press data to accurately estimate the timing of changes in visual consciousness during binocular rivalry. First, we categorized each time-point according to the flicker frequency of the dominant visual stimulus reported. To analyse the time-course of the probability of a button press state, we categorized button-presses (which could correspond to either low- or high-frequency) as either congruent or incongruent with the ongoing crossmodal stimulus frequency. Then, we obtained the probability of a congruent button press state as a function of time per subject, by averaging responses at each time point across all 144 trials per attention x frequency cue subtype.

For visual-only periods, the left button (corresponding to left-eye dominance) was arbitrarily set to congruent prior to the averaging of probability traces within subjects. As visual parameters were balanced across all blocks, this selection necessarily balanced across visual frequency and colour parameters, and we note that the identical analysis performed using right-eye congruence produced equivalent results. In Figure 3.5a, we compared among six conditions with one-way repeated-measures ANOVAs: 1,visual-only on attend days; visual only on non-attend days; 3, attended low-frequency; 4, attended high-frequency; 5, unattended low-frequency; and 6, unattended high-frequency. We defined significant differences among conditions at those time points that survived corrections for multiple comparisons with planned comparisons between cue types and the visual-only baseline, using FDR at q = .05 (Benjamini & Yekutieli, 2001).

3.4.9. Perceptual Switch Index (PSI)

To quantify crossmodal effects during binocular rivalry, we defined the perceptual switch index (PSI). PSI is the difference in the probability of a change in percept when comparing attended low-frequency to four other crossmodal cues. For the y-axis in Figure 3.5b, we calculated the PSI as the difference in the probability of viewing a congruent visual flicker over the period 1-4 s after cue onset.

3.4.10. EEG recording and analysis

EEG was recorded at a sampling rate of 1000 Hz using three BrainAmp amplifiers and 64-channel ActiCap (BrainProducts), with the impedance of each electrode kept below 10 kΩ. Ground and reference electrodes were AFz and FCz at recording, respectively. After re-referencing the data to the average of all channels, we performed linear detrending and bandpass filtering (0.1- 60 Hz with a Hamming-windowed finite impulse response filter) and down-sampled the data to 250 Hz before time-frequency analysis. We performed all time-frequency analyses using the Chronux toolbox (http://chronux.org; Bokil, Andrews, Kulkarni, Mehta, & Mitra, 2010), and custom MATLAB scripts.

3.4.11. SSVEP and signal-to-noise ratio (SNR) calculation

In our SSVEP paradigm, we first calculated the natural log of the power spectrum via single taper fast Fourier transform, during the period 0 to 4 seconds after the onset of a visual-only cue period. We computed the SNR at each frequency by subtracting from the natural log power at each frequency, the natural log power across neighbourhood frequencies. For this 4 second window (half-bandwidth = 0.25 Hz), we compared log power at f (Hz) to a neighbourhood defined as [f-1.5, f-0.5] Hz, and [f+ 0.5, f+1.5] Hz. When calculating the log(SNR) time-series for binocular rivalry analyses, we first calculated the log power using a sliding window of 1.5 s duration (half-bandwidth = 0.66 Hz), wish a step size of 0.15 s. Given this shorter window, and large half-bandwidth, the SNR time-course was computed using the neighbourhood [f - 2.66, f - 1.22] Hz, and [f+ 1.22, f + 2.66] Hz.

3.4.12. Binocular Rivalry SSVEP analysis via rhythmic entrainment source separation (RESS)

We applied RESS to optimally extract SSVEP components with different topographical profiles, without relying upon post-hoc electrode selection (Cohen & Gulbinaite, 2017). RESS takes advantage of spatial and temporal selectivity of SSVEP responses and creates a map of spatial weights across all electrodes, tailored

to maximally differentiate the covariance between a signal flicker frequency (F) and neighbourhood or reference frequencies (R). The largest eigenvalue obtained when comparing F and R covariance matrices denotes the eigenvector which can be used as channel weights to reduce the dimensionality of multi-channel data into a single component time-series, thereby increasing the signal to reference ratio at the flicker frequency F. As a result of multiplying the multi-channel data by a single eigenvector, the reduction of multi-channel data into a single component time-series reduces multiple comparisons across channels in statistical testing (Cohen & Gulbinaite, 2017).

We constructed RESS spatial filters from 64-channel EEG data, separately per frequency of interest [f1,2f1, 3f1, IM, f2, 2f2], and participant. Per participant, we used all visual-only periods (10s epoch; -3.5 to 6.5s after onset) as input data. Signal data covariance matrices were extracted after a narrow-band filter via frequency-domain Gaussian centred at F (full width at half maximum = 1 Hz), and we used broadband neural activity to construct reference matrices. The comparison of signal to broadband activity has previously been shown to enable the reconstruction of SSVEP signals using RESS (Cohen & Gulbinaite, 2017), and in our paradigm enables a common reference to be used for all signals, despite the irregular spacing between stimulus frequencies, their harmonics, and intermodulation components. Critically, RESS filters were constructed without first distinguishing whether a low- or high-flicker image was being perceived during rivalry, so as to avoid the possibility of overfitting our filters based on one possible percept during rivalry. After application of RESS spatial filters to reduce multi-channel data into a single component time-series, we analysed the time-course of log(SNR) responses using the procedure described above.

3.4.13. Participant-by-participant image analysis of binocular-rivalry SNR

To complement the mean SNR time-series during binocular rivalry, we performed a participant-by-participant image analysis to visualize the consistency of SNR responses across subjects. For this analysis, we first sorted switches by direction (e.g. perceiving high-flicker to perceiving low-flicker; pf2 to pf1), and separated SNR

responses for our two main frequency-tags of interest (f1 and f2). To compare across subjects, we first normalized the SNR-time course per participant, dividing by standard deviation, after subtracting the mean SNR value per participant. After normalizing SNR responses in this way, we then sorted participants in descending order based on the magnitude of their SNR response after button press (averaged from 0 to 3 s). This participant order was calculated for the period 0 to 3s after a switch from pf2 to pf1 for both f1- and f2-SNR, allowing patterns in the across subject consistency of responses to be visualized separately for each flicker frequency. The same participant order (at f1 and f2) was then used to display SNR responses during switches in the opposite direction (from pf1 to pf2), to visualize the within-subject consistency of SNR responses based on switch direction. Similar to previous image-based analyses (Fujiwara et al., 2017), we also smoothed along the y-dimension (5 point moving average) to visualize the temporal consistency of SNR changes across subjects. To quantitatively analyse the differences in participant-responses to low and high-flicker during rivalry, we performed a median-split of sorted-participants at each frequency-tag of interest.

3.4.14. ITPC analysis

To assess crossmodal stimulus-locked neural synchronization, we analysed the inter-trial phase coherence (ITPC) within electrodes, over multiple time-frequency points (Bastos & Schoffelen, 2016). ITPC is an amplitude-normalized measure of the degree to which EEG responses are phase-locked to the onset of an exogenous cue, ranging between 0 (random phase over trials) and 1 (perfect phase consistency over trials). To compute ITPC, the consistency of phase angles is computed as the length of the average of unit phase vectors in the complex plane over trials. For a given time, t, and frequency, f,

$$\mathrm{ITPC(t,f)} = \left| \frac{1}{N} * \sum_{n=1}^{N} e^{i(\theta(t,f,n))} \right|$$

where N is the number of trials, and θ is the phase angle at time t, frequency f, and trial n.

3.4.15. Statistical analysis - EEG.

In the EEG spectra, we tested the significance of SSVEP peaks after correcting for multiple comparisons with a False Discovery Rate (FDR) of .05 (Benjamini, Krieger, & Yekutieli, 2006; Benjamini & Yekutieli, 2001). To assess the statistical significance when comparing between SNR time-courses, we used temporal cluster-based corrections for multiple comparisons (Maris & Oostenveld, 2007). For this analysis, we detected any temporally contiguous cluster by defining a significant time point as $p < .05$ uncorrected, based on repeated-measures *t*-tests at each time-point of interest. We then obtained a summed cluster-level test statistic for the cluster, which was the sum of observed test-statistics (e.g., *t* scores) for comparison with a permutation-based null distribution. We then repeated this procedure after shuffling the subject specific averages within each participant 2000 times, retaining the maximum cluster-level test-statistic from each shuffled data. If the original summed cluster-level statistics exceeded the top 97.5% of the null distribution of this null-distribution, we regarded the original observed effect to be significant (as $p_{cluster} < .025$). When the observed effect exceeds the entire null distribution, we report this value as $p_{\text{cluster}} < .001$).

To assess the statistical significance of topographic SNR and ITPC responses, we used a similar two-stage cluster-based approach. We retained electrodes only after identification of a significant spatial-cluster, where inter-electrode distances did not exceed 3.5 cm. We also altered the criterion for significance based on the measure of interest in order to avoid overfitting (van Driel, Knapen, van Es, & Cohen, 2014). We set the uncorrected threshold for ITPC at $p < .05$ uncorrected, and threshold for log(SNR) responses to $p < .01$ uncorrected. This more stringent criterion accounts for the relatively increased magnitude of SNR responses, as a cluster threshold of $p < .05$ (uncorrected) would result in a single contiguous spatial cluster being submitted for the second stage of statistical analysis.

At the second stage, we again retained the sum of observed *t*-scores within the identified cluster, which we retained as our observed test-statistic. To create the null distribution, condition labels (e.g. attend vs. non-attend cue conditions) were randomly shuffled for each electrode within each subject, to create two surrogate datasets the same size as our original condition comparison. Then the *t*-scores were computed for each electrode based on our surrogate datasets, and the electrode with

the maximum *t*-score and the maximum *t*-score of its neighbours was retained. Similar to the procedure described above, this procedure repeated 2000 times to obtain a null distribution. Against this distribution, the sum of observed *t*-scores for the candidate cluster was then compared. When the observed sum of *t*-scores was within the top 5% (or cluster corrected to $p < .05$) then we concluded that there was a significant difference between conditions.

3.5. Results

3.5.1. Successful frequency-tagging at stimulus frequencies and harmonics during visual only periods

We investigated the strength of frequency-tagged activity while manipulating the allocation of attention and conscious visibility of images during a crossmodal binocular rivalry experiment. Subjects (N=34) participated in two days of 24 x 3-minute binocular rivalry blocks, during which they constantly reported on the content of their consciousness via button press. To frequency-tag the rivalling visual stimuli, each image underwent 4.5/20 Hz (F1/F2) sinusoidal contrast modulation while we simultaneously recorded 64-channel EEG (see Methods). On both days of experimentation, 15 cue periods were presented per 3-minute block (mean duration 2.6 s), which were either sinusoidally amplitude-modulated crossmodal cues (12 cues per block; auditory, tactile, or combined auditory and tactile), or empty visual-only periods (3 cues per block). All cue periods were separated by 7-10 s, with visual-only cue periods serving to increase the uncertainty of stimulus timing, as well as form a baseline for subsequent analyses.

To build upon previous psychophysical research investigating the impact of crossmodal cues on rivalry, we also explicitly manipulated the allocation of attention either toward or away from crossmodal stimulation. To manipulate the allocation of attention to crossmodal cues, on one day of experimentation (day 1 or n=16, day 2 for n=18) subjects performed an additional task on top of their continuous rivalry report. We instructed subjects to note the flicker-frequency of each crossmodal cue (as either low/high corresponding to F1/F2) and to silently count whenever the flicker-frequency of the crossmodal cue matched the flicker-frequency of their dominant

visual percept at the point of crossmodal cue offset. This ensured continued attention throughout crossmodal cues (2- 4 s in duration), with the number of congruent visual-crossmodal cue periods reported verbally after each 3-minute block. Figure 3.1 displays a schematic of this procedure.

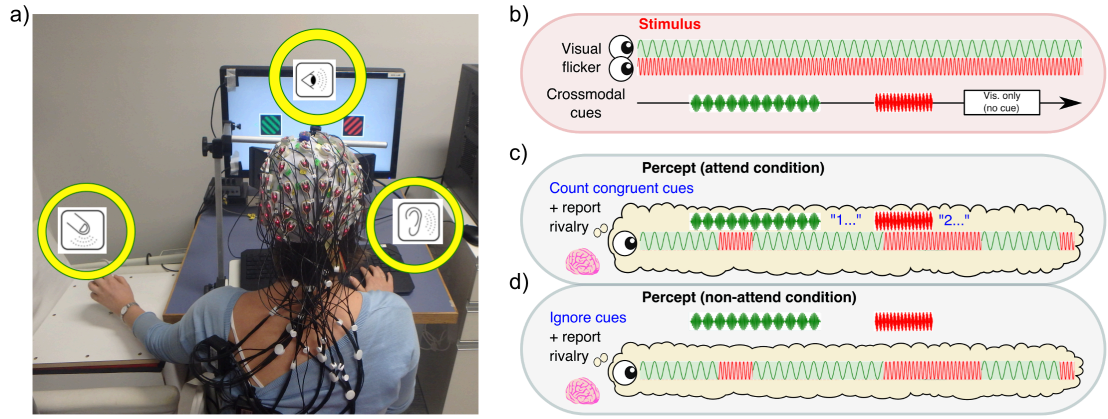

Figure 3.1. Experiment overview.

a) Experimental apparatus. b) Schematic time-course of stimulus presentation. Each eye was presented with a 4.5 or 20 Hz sinusoidal flicker (low [F1] and high [F2] flicker, respectively) throughout 3-minute blocks. Crossmodal cues (also F1 or F2; 2, 3.1 or 4 s in duration) or visual-only periods (2.6 s in duration) were separated by inter-stimulus intervals of 7–10 s. On two days of testing, subjects reported their dominant visual percept during rivalry via continuous button press. c) On attend days (day 1 for n=16), subjects were also tasked with silently counting the number of crossmodal cues that were congruent with their dominant visual percept at the time of cue offset ("2" in this example). d) On non-attend days (day 1 for n=18), subjects were instructed to ignore all crossmodal cues and report rivalry only.

Across both days of experimentation, we found strong and occipitally localized steady-state visually evoked potentials (SSVEPs) during visual-only periods. We measured the relative strength of responses at each flicker frequency by computing the natural log of the power spectrum via fast Fourier transform, operationalized as a signal-to-noise ratio (SNR). In our SSVEP paradigm, the SNR represents the relative power at each tagged frequency in log scale, subtracted by the mean log power across neighbourhood frequencies (see Methods). Over the period 0 to 4 seconds after the onset of a visual-only cue period, significant increases in log(SNR) were observed for both flicker frequencies (f1 and f2), and their harmonics (2f1; 3f1; 2f2). Significant increases in the intermodulation component (IM; f2-f1), previously advocated as markers of inter-ocular conflict (Katyal et al., 2016; Zhang et

al., 2011) and/or neural integration (Alp et al., 2016; Boremanse et al., 2013; Gordon et al., 2017) were also observed over occipital regions. Figure 3.2 displays a summary of these results.

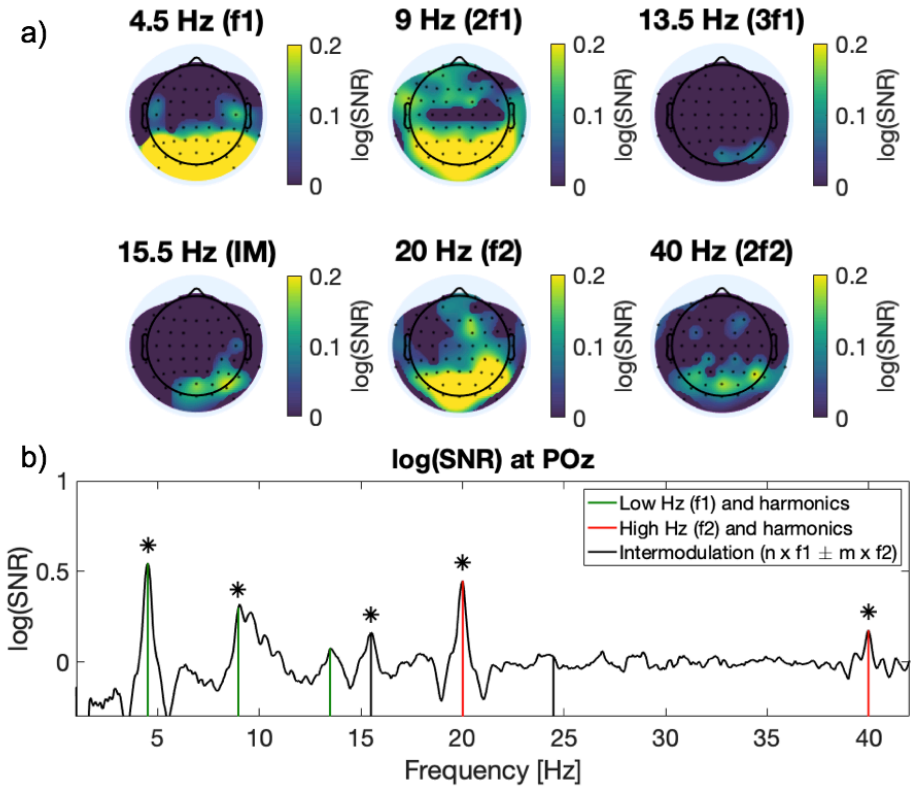

Figure 3.2. Across participant SSVEP-SNR topoplots and spectrum at POz during visual-only stimulation periods.

a) Topoplots for log(SNR) responses averaged across subjects, non-significant electrodes are masked ($p < .05$, FDR corrected). Subjects reported on binocular rivalry between two flickering 4.5 Hz (F1) and 20 Hz (F2) gratings. b) Representative spectra at parietal-occipital channel POz. Green lines show SSVEP responses to F1 and harmonics, red lines show responses to F2 and harmonics. Black lines indicate intermodulation frequencies resulting from the combination of f1 and f2 components. Asterisks mark significant peaks compared to zero at POz (calculated from 0 to 4 seconds after onset; $p < .05$ Bonferroni corrected).

3.5.2. Low- but not high-flicker frequency-tags correlate with the contents of consciousness across subjects

Given the significant increases in log(SNR) across frequency-tags, we next investigated whether the strength of frequency-tagged activity correlated with perceptual reports over time. As previous investigations have indicated that the strength of frequency-tagged activity positively correlates with dominant visual

flicker, we proceeded by epoching EEG around button press responses (± 3 seconds), and comparing the strength of frequency-tagged activity during perceptual switches to the low or high frequency visual flicker, respectively.

To accurately compare the strength of frequency-tagged activity from whole-head responses, we first applied rhythmic entrainment source separation (RESS) to extract the time-course of frequency-tagged activity without a priori electrode selection (Cohen & Gulbinaite, 2017). RESS creates a map of spatial weights across all electrodes, tailored to optimize the SNR at a particular frequency, per participant. By applying RESS, we were able to analyse the time-course of log(SNR) responses despite differences in source topography, and to avoid correcting for multiple comparisons following electrode selection (see Methods). Unless otherwise specified, SNR values reported reflect RESS log(SNR).

After applying RESS per participant at each frequency (f1, 2f1, 3f1, IM, f2, 2f2), we compared the across-participant mean SNR during a reported switch to either the low-flicker (pf2 to pf1) or high-flicker percept (pf1 to pf2). Figure 3.3 displays a summary of these results, collapsed across both days of testing. No differences were found between attention conditions during these visual only periods.

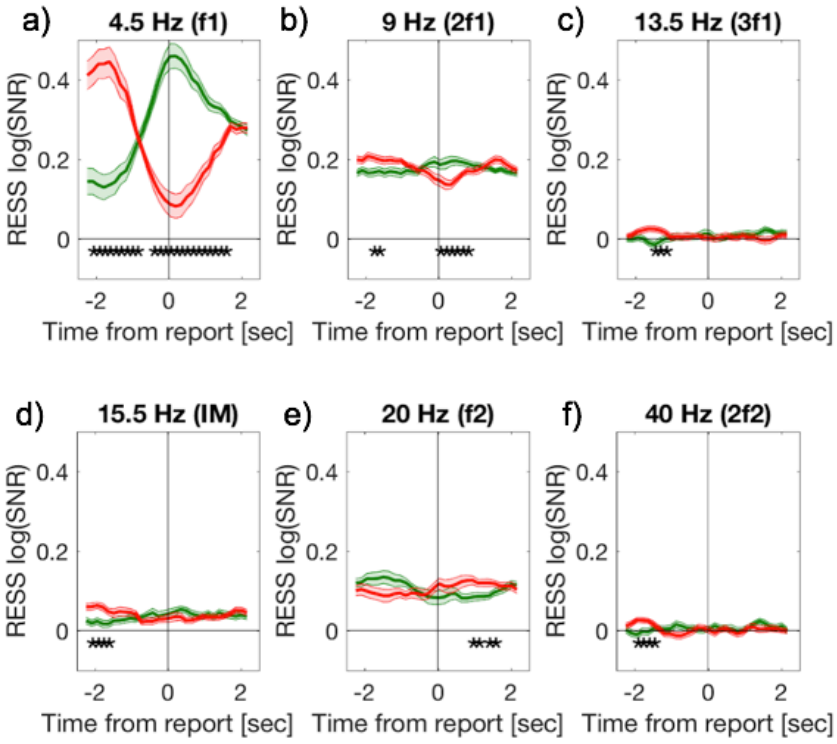

Figure 3.3. SNR time-series of RESS log(SNR) during binocular rivalry (visual only conditions).

Green lines and shading show SNR during a switch from pf2 to pf1 (perceiving high- to perceiving low-flicker), red lines and shading show a switch from pf1 to pf2 (low to high flicker). Shading indicates 1 SEM corrected for within-subject comparisons (Cousineau, 2005). Asterisks mark significant differences in RESS log(SNR) at each time point based on switch direction (paired samples t-tests, $p < .05$ temporal cluster-corrected, Maris & Oostenveld, 2007).

We captured positive correlations between the low-flicker frequency-tag (f1) and first harmonic component (2f1) with the contents of consciousness. SNR was significantly weaker preceding a switch to the low-flicker at f1 (-2.24 to - 1.03 s; $p_{cluster} < .001$), 2f1 (-1.94 to -1.78 s; $p_{cluster} = .001$), and 3f1 (-1.64 to -1.33 s; $p_{cluster} < .001$). By contrast, SNR strength was significantly higher after a switch to the low-flicker, for both f1 (-0.57 to 1.40s; $p_{cluster} < .001$) and 2f1 responses (-0.12 to 0.64 s, $p_{cluster} < .001$). We also observed percept related modulation of the IM response (f2-f1). SNR at IM was significantly lower prior to a switch to the low-flicker percept from -2.44 to -1.79 s ($p_{cluster} < .001$).

In contrast to the f1 and IM components, RESS log(SNR) strength was strongest for f2 responses when perceiving the high-flicker frequency image.

Specifically, SNR strength was significantly increased over the periods 0.80 to 0.95 s ($p_{cluster}$ = .015) and 1.25 to 1.40 s ($p_{cluster}$ < .001) after switching from the low to high-flicker during rivalry. Unlike the F1 harmonics, this pattern did not continue for the first harmonic of F2 (2f2). At 2f2, SNR responses were highest from -2.09 to -1.64 s ($p_{cluster}$ < . 001) prior to a switch to the high-flicker image.

To better understand these differences in the efficacy of F1 and F2 frequency-tags, we next investigated the time-course of frequency-tagged RESS log(SNR) per individual subject. For this analysis, we first normalized the SNR per participant (see Methods). To determine whether the pattern of increasing and decreasing SNR was consistent within individual subjects, we next sorted subject-level SNR time-courses by the total increase in SNR after a switch to the low-frequency visual flicker (pf1). This results in participants with the strongest increases in f1-SNR during a switch from pf2 to pf1 to be located at the highest rows in Figure 3.4a, and strongest responses at f2 after a switch from pf2 to pf1 in the highest rows of Figure 3.4e. To visualize whether the magnitude of changes to SNR were consistent within each participant, the same participant order based on f1 and f2 responses was used to plot the corresponding time-course during a switch in the opposite direction (Figure 3.4b and 3.4f).

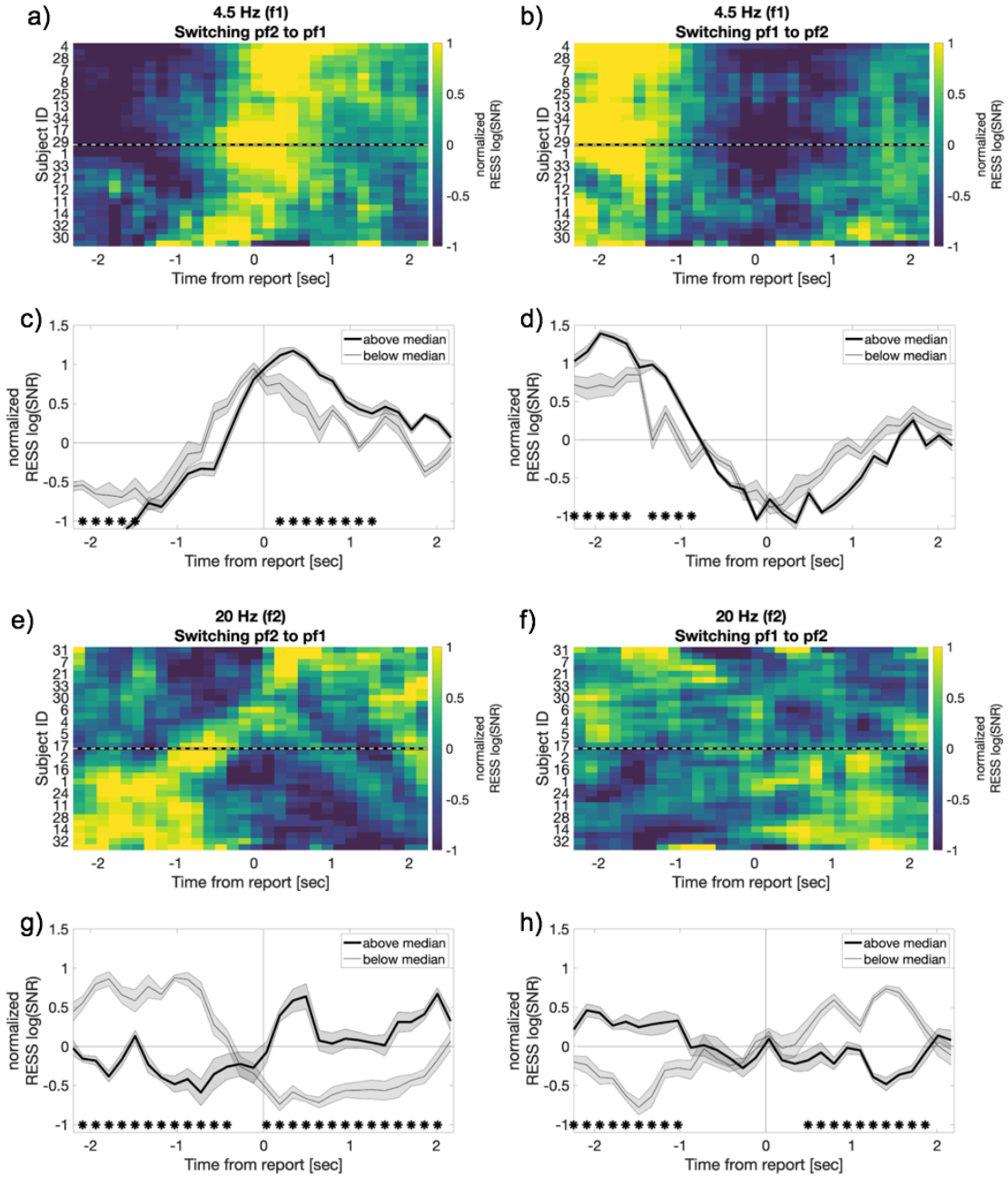

Figure 3.4. F1-frequency-tags are positively correlated with perception across subjects, but F2-frequency-tags are not.

a-b) f1- and e-f) f2-RESS log(SNR) time-series after sorting individual subjects based on the strength of increased SNR after button press, when switching from pf2 to pf1. Y-axis identifies individual subject order determined in a) and e), with corresponding decrease in SNR during a switch from pf1 to pf2 shown in b) and f). Note that the order of subjects is identical in a-b an e-f, though only every 2nd subject ID is displayed. c-d) and g-h) The time-course of RESS log(SNR) after separating subjects based on the strength of responses after median split (cut-off indicated by broken black and white lines in a-b,e-f). Asterisks mark significant differences in SNR time series (cluster corrected, $p < .05$). Though the direction of SNR changes are highly consistent across subjects at f1, a bimodal distribution of SNR time-series is evident for f2.

From this analysis it is clear that at f1, increases and corresponding decreases in SNR upon changes in perceptual content were highly consistent within individual subjects (comparing equivalent rows in Figure 3.4a and 3.4b). The direction of changes to SNR strength were also highly consistent across participants (comparing down the y-axis). In other words, during a switch to pf1, f1-SNR reliably increased across subjects (Figure 3.4a). By contrast an opposite trend was observed for f1-SNR strength when switching from pf1 to pf2 (Figure 3.4b). To quantify this across-subject consistency we performed a median-split of SNR time-series across subjects. Figure 3.4c and 3.4d display the results of this analysis. Although significant differences exist between the median-split SNR time-series (pf2 to pf1; -2.24 to -1.49 s, .019 to 1.25 s, $p_{\text{cluster}} < .001$; pf1 to pf2; -2.24 to -1.64 s, -1.33 to -0.88 s , $p_{\text{cluster}} < .001$), most critically, the shape of SNR is highly consistent and positively correlates with the dominance or suppression of the low-flicker image. This pattern indicates that f1-SNR positively correlates with the contents of consciousness across subjects.

Similar to f1-SNR responses, f2-SNR responses were consistent within subjects (i.e. comparing the same rows in Figure 3.4e and 3.4f). That is to say that increases and decreases in SNR were occurring with a change in switch direction, respectively. Strikingly however, the direction of these changes was not consistent across subjects. Across subjects, a bimodal distribution of SNR responses was observed, such that the direction of SNR changes inverted on an individual basis; either positively or negatively correlating with the contents of conscious perception (Figure 3.4e). To quantify this across-subject inconsistency, we again performed a median-split of SNR time-series across subjects (Figure 3.4g and 3.4h). Compared to the consistency of f1-SNR responses, the shape of f2-responses was shown to either positively or negatively correlate with the contents of consciousness when perceiving the high-flicker. These differences were significant for both directions of switches (pf2 to pf1; -2.09 to -0.42 s, 0.04 to 2.01 s, $p_{cluster} < .001$; pf2 to pf1, -2.24 to -1.03 s, 0.49 to 1.86 s, $p_{\text{cluster}} < .001$). We return to this pattern of results in our discussion.

3.5.3. Attending to low-frequency crossmodal cues alters binocular rivalry dynamics

Previous research has also indicated that during binocular rivalry, the normally stochastic alternations between each eye's image can be altered by the presentation of

crossmodal cues. In our paradigm, visual and crossmodal cues were matched in temporal frequency, at either 4.5/20 Hz (F1/F2). We continued by investigating whether crossmodal cues altered the ongoing probability of perceiving one image over the other, by aligning all button press data at cue onset and averaging over subjects for each cue condition (Figure 3.5).

We found a strong interaction between attend condition and flicker frequency mediating the effect of crossmodal cues on rivalry. The probability of perceiving a congruent visual flicker increased only when attending to low-frequency crossmodal cues. Compared to the probability of perceiving either flicker during visual only periods, attended low-frequency crossmodal cues increased the probability of perceiving the low-frequency flicker over the period 0.68 to 3.97 s after cue onset (repeated measures ANOVA followed by planned comparisons, FDR q = .05, Figure 3.5a). When comparing the strength of this effect between auditory, tactile, or simultaneous auditory and tactile cues, no differences emerged (Figure 3.11).

To confirm that the increased probability of seeing low-flicker during attended crossmodal cues was indeed due to attention, we performed a correlation-based analysis. In our paradigm, after each 3-minute block during attend conditions, participants verbally reported the number of congruent visual and crossmodal cues. We calculated the Pearson's correlation coefficient between this verbal report and the actual number of congruent cues based on button-press data. We labelled this as attention to cues (x-axis in Figure 3.5b), and compared subject attention to cues, to the strength of crossmodal cue effects displayed in Figure 3.5a. Specifically, we defined the strength of the attended low-frequency crossmodal cue effects as the difference in the probability of seeing the congruent visual flicker 1 to 4 seconds after cue onset, compared to the same probability during visual-only periods (y-axis in Figure 3.5b). As this measure reflects the degree of perceptual switch after cue onset, we label it as the perceptual switch index (PSI). The relationship between attended low-frequency PSI and attention to cues exhibited a strong positive correlation ($r(32) = 0.46$, $p = .0006$, two-tailed), confirming that the increased probability to perceive congruent visual flicker was mediated by attention.

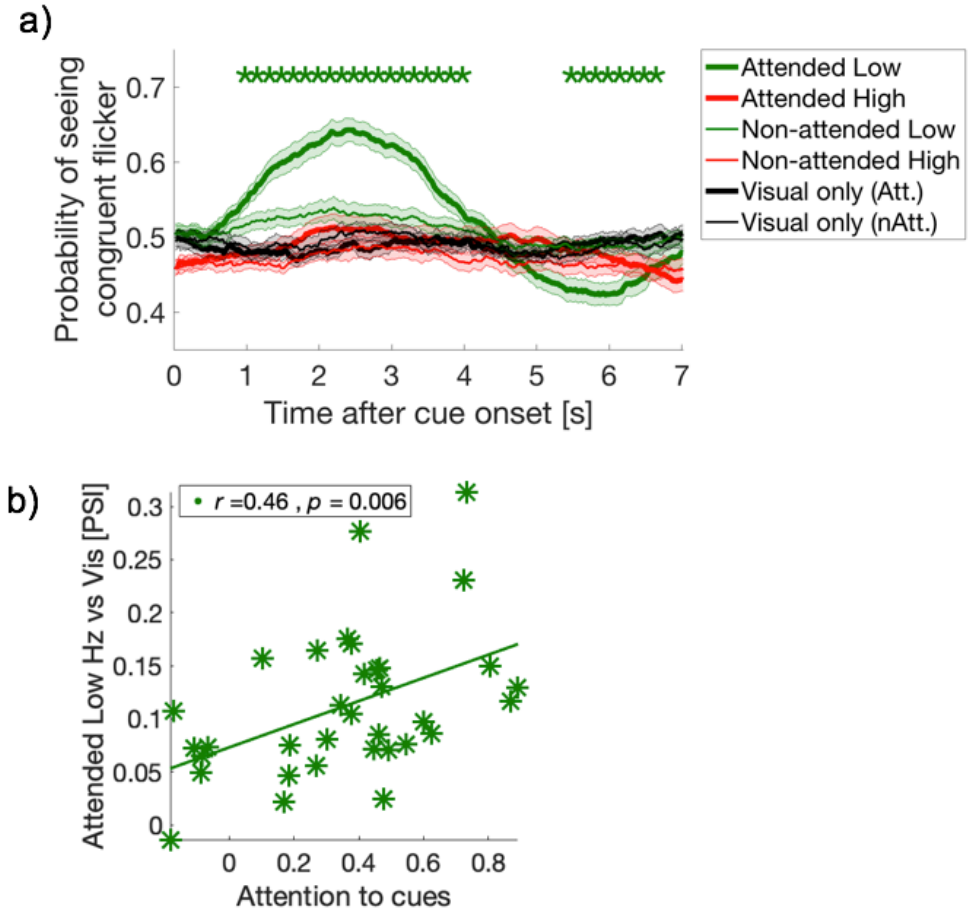

Figure 3.5. Crossmodal cue effects on binocular rivalry dynamics.

a) Button-press data. Y-axis represents the proportion of button-presses reporting congruent crossmodal and visual flicker at each time point, sampled at 60 Hz (or every 16.7 ms). Coloured lines and their shading show mean ±1 standard error across 34 subjects during attended and ignored cues (thick and thin lines) for low and high frequency (green and red colours). Black lines represent the equivalent probability for visual-only periods, serving as baseline (Methods). Asterisks indicate a significant difference between cues at each time point (repeated-measures ANOVA followed by planned comparisons). We use FDR q = .05 for the statistical threshold unless noted otherwise. b) Crossmodal effects are mediated by task-relevant attention. Y-axis shows the mean probability of perceiving the congruent flicker during 1-4 sec after cue onset for attended-low-frequency cues (thick green in panel a). Attention to cues (x-axis) is the correlation coefficient between the reported and actual congruent stimuli when comparing between rivalry percepts and crossmodal cues at offset (See Methods for details).

3.5.4. Frequency-specific topographic responses to crossmodal cues

Our previous analysis showed a significant behavioural effect of attended low-frequency crossmodal cues compared to all other cue types. We next investigated the

spatiotemporal neural correlates of these effects mediating a change in conscious awareness. In our paradigm, crossmodal cues are presented at the same time as continuous and ongoing visual flicker. As such, to analyse cue specific neural effects, we computed an inter-trial phase coherence (ITPC) measure after aligning all epoched cue periods at crossmodal cue onset. If the ITPC measure captures the timing of consciousness related change, then we hypothesized that these neural correlates could be localized with this measure. For comparison, we also computed the log(SNR) across electrodes, again comparing activity during crossmodal periods to visual only periods (0 to 2 s window; see Methods).

We first localized regions increasing in frequency-specific ITPC compared to visual-only periods, focusing on attended crossmodal cue types. We label these attended-cue responses as evoked ITPC and evoked log(SNR). We observed frequency-specific changes to the whole-head ITPC response when attending to either f1 or f2 compared to visual-only cue periods. Specifically, attended f1-crossmodal cues evoked significantly greater f1-ITPC over bilateral fronto-central-temporal electrodes (Left; [FT7, FC5, FC3, C5], $p_{cluster} < .001$; and right side; [F4, F6, FC2, FC4, FC6, C4, C6], $p_{cluster} < .001$ Figure 3.6a). In contrast, f2-crossmodal cues evoked significantly greater f2-ITPC over an extended fronto-central and parietal-occipital region, and only over the right hemisphere ([Fz, F2, F4, F6, FC2, FC4, Cz, C2, C4, C6, CP4, CP6, TP8, Pz, P4, P2, P6, POz, PO4, Oz], $p_{cluster} < .001$).

These differences in evoked ITPC topography were largely mirrored in evoked log(SNR), but with one notable exception. As well as an increase in f1-SNR over bilateral fronto temporal sites (both $p_{cluster} < .001$), increases in predominantly parieto-occipital f1-SNR were also observed during f1-cues compared to visual only-periods ([TP9, TP7, CP5, CP3, CP4, TP8, TP10, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, POz, PO4, PO8, PO9, O1, Oz, O2, PO10], $p_{cluster} < .001$). No increase in occipital f2-SNR was observed following f2-cues. We return to these results in the discussion.

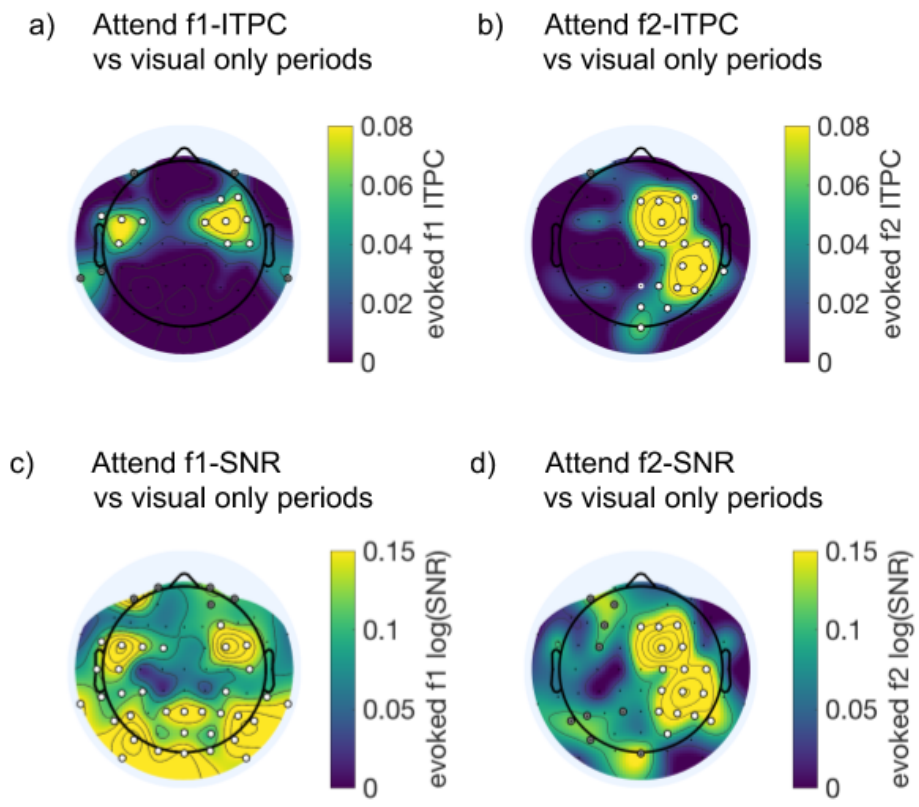

Figure 3.6. Evoked ITPC and evoked log(SNR) responses for attended crossmodal cues at f1 and f2.

a-b) ITPC and c-d) log(SNR) during crossmodal cues compared to visual only periods, calculated 0 to 2 s after cue onset. In a-b) grey disks mark p <.05 uncorrected, in c-d) grey disks mark p < .01 uncorrected due to differences in SNR affecting the cluster threshold (see methods). Significant increases in ITPC or log(SNR) when comparing crossmodal cues to visual only periods are marked with white disks (p < .05, spatial cluster corrected, Maris & Oostenveld, 2007).

3.5.5. Crossmodal SNR and ITPC do not significantly correlate with altered binocular rivalry dynamics

As attended low-frequency cues uniquely contributed to a change in consciousness during binocular rivalry, we focused our remaining analysis on this condition. We hypothesized that attending to crossmodal cues would result in an increase in the SNR or ITPC at f1 compared to non-attended f1-cues. In addition, we also expected the differences in these metrics could correlate with the PSI across individuals.

We did not find that an increase in log(SNR) or ITPC during attended cues could account for the increase in probability of perceiving the low-flicker across subjects. Figure 3.7, displays the results of this analysis. When comparing attended to non-attended cues, we again observed an increase in f1-ITPC strength over bilateral electrodes, although only a right frontal cluster survived spatial cluster corrections for multiple comparisons ([AF8, F6], $p_{cluster}$ = .009; Figure 3.7a). We next tested whether differences in the strength of f1-ITPC between attended and non-attended crossmodal cues would correlate with PSI strength across subjects. Neither the right frontal cluster identified, nor ITPC in any regions significantly correlated with altered binocular rivalry dynamics (Figure 3.7b).

As a complement to these ITPC analyses, we also compared the strength of f1-log(SNR) between attended and non-attended crossmodal cues (Figure 3.7c). Attending crossmodal cues increased the f1-log(SNR) over right fronto-central regions ([FC4, FC6], $p_{cluster}$ < .001), yet this difference in SNR strength did not mediate the PSI across subjects (Figure 3.7d). We will briefly review these differences in frequency-specific effects, before returning to analyse the attended low-frequency condition in more detail in **Chapter 4**.

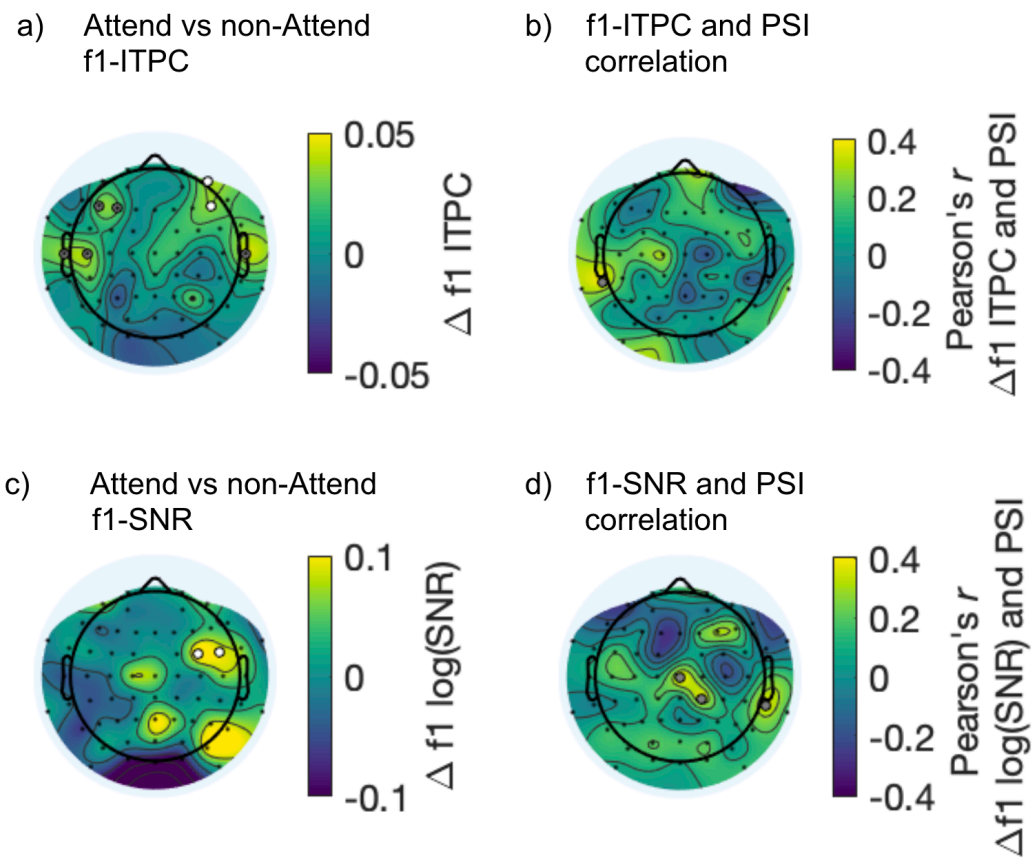

Figure 3.7. Attention effects ITPC and log(SNR) strength, but these changes do not correlate with PSI across subjects

a) f1-ITPC strength during attended compared to non-attended cues (0 to 2 s after onset, see methods). Grey disks mark $p < .05$ uncorrected, white disks mark $p < .05$ cluster corrected for multiple comparisons. b) Despite significant differences in f1-ITPC during attended crossmodal cues over right frontal regions, no regions significantly correlated ITPC strength with PSI across subjects (see methods). c) f1-log(SNR) strength comparing attended to non-attended crossmodal cues, white disks denote $p<.05$ cluster corrected. d) No significant correlations between SNR strength and PSI survived spatial cluster-corrections for multiple comparisons (grey disks denote $p < .01$ uncorrected).

3.6. Discussion

By simultaneously frequency-tagging visual, auditory and tactile stimuli, we have extended previous psychophysical investigations which demonstrated that crossmodal cues can alter the contents of consciousness during binocular rivalry (Lunghi & Alais, 2015; Lunghi et al., 2014). Our findings provide evidence that frequency-specific mechanisms mediate the integration of multi-sensory stimuli into a

coherent percept, and show that these effects can depend upon the allocation of attention. With respect to the driving empirical questions of **Chapter 2,** we found that attention interacts with multisensory integration in a frequency-specific manner. By flickering visual stimuli at widely spaced frequencies (4.5 and 20 Hz), we have also unveiled a novel feature of frequency-tagged responses during binocular rivalry, in that neural responses at the high-flicker frequency-tag can either positively or negatively correlate with the contents of consciousness on an individual basis. We here discuss this novel frequency-dependent behaviour of SSVEP amplitude during rivalry, before continuing in **Chapter 4** to further investigate the unique effect that attended-low frequency cues held over rivalry dynamics.

3.6.1. Frequency-dependent tracking of consciousness during binocular rivalry

Frequency-tagging visual stimuli during binocular rivalry has been heralded as a powerful technique to unveil the neural correlates of conscious perception (Norcia et al., 2015; Tononi, Srinivasan, et al., 1998a). Since it was first shown that the frequency-content of occipital EEG positively correlates with an increase in contrast for a flickering image (Lansing, 1964), the sequential enhancement and suppression of cortical responses have been interpreted to reflect the alternating visual-dominance and suppression phases of flickering images during rivalry (Brown & Norcia, 1997; Cosmelli et al., 2004; Kamphuisen et al., 2008; Katyal et al., 2016; Srinivasan & Petrovic, 2006; Srinivasan et al., 1999; Tononi, Srinivasan, et al., 1998a; P. Zhang et al., 2011).

What then might explain the negative correlation between f2-SNR in our paradigm, and the current contents of consciousness? The individual differences we observe in tag-efficacy for high-frequency flicker (20 Hz) could be due to differences in the effects of attention on SSVEP amplitude. For example, the effects of attention vary across SSVEP frequency-bands (Ding et al., 2006; Gulbinaite et al., 2019), a feature which may have been overlooked given typical rivalry investigations have used closely spaced flicker frequencies, predominantly below ~10 Hz (Alpers, Ruhleder, Walz, Mühlberger, & Pauli, 2005; Brown & Norcia, 1997; Cosmelli et al., 2004; Jamison et al., 2015a; Katyal et al., 2016; Roy, Jamison, He, Engel, & He,

2017; Srinivasan & Petrovic, 2006; Srinivasan et al., 1999; P. Zhang et al., 2011) or either side of the canonical alpha band (Kamphuisen et al., 2008; Tononi, Srinivasan, et al., 1998; yet see Sutoyo & Srinivasan, 2009). Contributing to this asymmetry, SSVEPs are generally strongest at frequencies below 10 Hz (Ding et al., 2006; Norcia et al., 2015; Vialatte et al., 2010), and importantly, although attention can increase stimulus-specific SSVEP amplitudes (Morgan et al., 1996; Müller, Picton, et al., 1998), responses to unattended stimuli may also increase in amplitude in the presence of a simultaneous and strong distracting flicker (Chen, Seth, Gally, & Edelman, 2003; Ding et al., 2006; Gulbinaite et al., 2019; Wang, Clementz, & Keil, 2007).

In our paradigm, f2-responses were recorded in the presence of a simultaneous, and relatively stronger low-frequency flicker. The superposition of flickering gratings has previously been shown to create asymmetric effects of attention - such that SSVEP responses to an unattended stimulus increase in amplitude (Chen et al., 2003; Wang et al., 2007). For example, after superimposing vertical and horizontal gratings, Wang et al directed their subjects to allocate attention only to foreground features of the horizontal grating. Although distinct in feature (red/green) and flicker frequency (7.13/8.33 Hz), increased SSVEP amplitude was also recorded for the unattended grating, an effect which increased with more focused attention (Wang et al., 2007). We propose that in our paradigm, attending to f1-flicker during rivalry also increased f2-SNR for a subset of individuals, to the extent that f2-SNR became anticorrelated with perceptual contents in a number of our subjects. The source which characterizes this subset is currently undefined, as our exploratory analyses could find no effect of condition order, or correlation between f2-SNR and effect size of f2 crossmodal cue effects. Indeed we note that while the typical increases and decreases observed in frequency-tagged responses have been taken to reflect the contents of consciousness per se, an alternative possibility is that they reflect increased SSVEP amplitude with attention toward images as they enter and exit conscious awareness (Dieter & Tadin, 2011; Li et al., 2017; Zhang et al., 2011). As such, other indices of individual differences in attention, such as occipital alpha peak frequency and/or lateralization, may offer inroads for investigation. This speculative proposition will need to be investigated in future work to understand the source of individual susceptibility to this effect.

3.6.2. Frequency-dependent effects of crossmodal cues

We also observed striking frequency-specific effects for the efficacy of crossmodal cues on binocular rivalry dynamics. Only attended low-frequency crossmodal cues were capable of altering the normally stochastic probability of perceiving one visual flicker over the other. Our result does not support the earlier work of Lunghi et al., (2014), who demonstrated that passively hearing/feeling amplitude modulated auditory/tactile tones could alter rivalry dynamics when modulated at both 3.75 and 15 Hz. When investigating these effects in the EEG, we found that attending to crossmodal cues enhanced both f1-ITPC and f1-SNR over bilateral fronto-temporal electrodes (Figure 3.6.). Unexpectedly, these regions-of-interest were not shared by attended f2-ITPC and SNR, which increased in magnitude uniquely over the right hemisphere.

Compared to SSVEPs, the topographic profiles of cross-modal steady-state responses have been less well characterized (Colon, Legrain, & Mouraux, 2012; Norcia et al., 2015; Tanaka, Kuriki, Nemoto, & Uchikawa, 2013). In the only study we are aware of which has tested auditory steady-state responses (ASSRs) at a similar range (4/ 20 Hz), EEG source clusters for 4 Hz modulation were most frequent across subjects within left and right frontal lobes, with medial and right hemispheric sources demonstrated for 20 Hz responses (Farahani, Goossens, Wouters, & van Wieringen, 2017). In the tactile domain, steady-state somatosensory potentials are strongest when evoked by approximately 20-30 Hz modulation (Snyder, 1992), which we note may have contributed to the asymmetry observed in our topoplots of f1 and f2 crossmodal responses. Specifically, we suggest that strong right-lateralized responses to our left-handed tactile-stimuli disproportionately contributed to the crossmodal-topoplots, resulting in a right-lateralized response for f2-cues (Figure 3.12). When attending to low-frequency cues an increase in occipital f1-SNR was also observed (Figure 3.6). The most parsimonious explanation for this result regards the increased probability to perceive the low-frequency visual flicker during these cues, which given the strength of f1-SNR frequency-tags we observed (Figure 3.3), would result in an increase in occipital f1-SNR 0 to 2 seconds after cue onset.

Given the inability in our data to determine whether the aforementioned increase in occipital f1-SNR was a cause of, or secondary to the change in

consciousness, we also compared attended and non-attended crossmodal cues to isolate attention-related changes to topographic activity. For both f1-SNR and f1-ITPC, we observed right fronto-temporal clusters increased in magnitude in response to attended crossmodal cues. Across participants however, these increases in f1-response magnitude did not correlate with the likelihood of a change in conscious percept over time. As a result, we continue with **Chapter 4** to investigate the attended low-frequency cue effects in more detail.

3.7. Supplementary figures

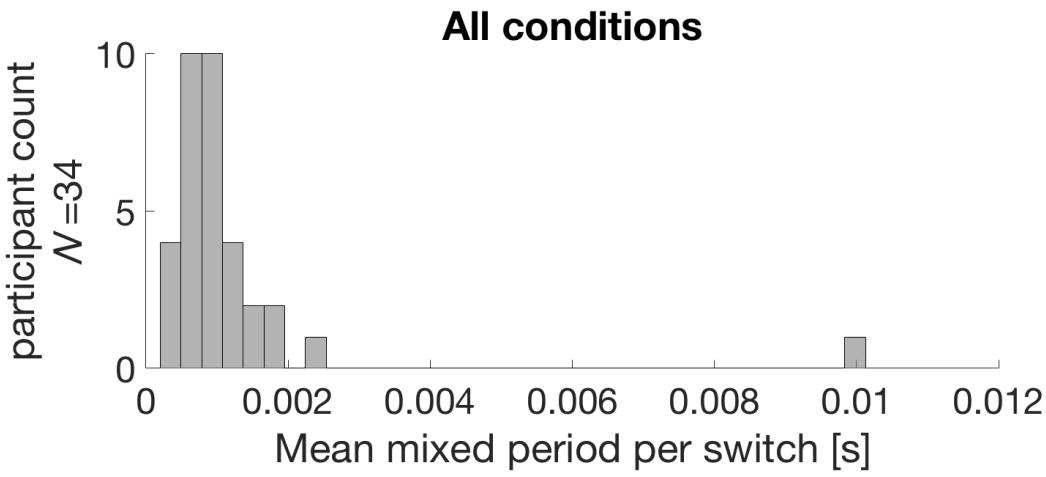

Figure 3.8. Average mixed periods during binocular rivalry.

Across all experimental periods, the average duration of mixed periods per switch per subject was less than 16.7 ms (our binning width), thus showing that mixed percepts are unlikely to have contributed to an increase in the variance of perceptual report timing. Switches happened instantly, with zero or one mixed frame (16.7 ms) on average.

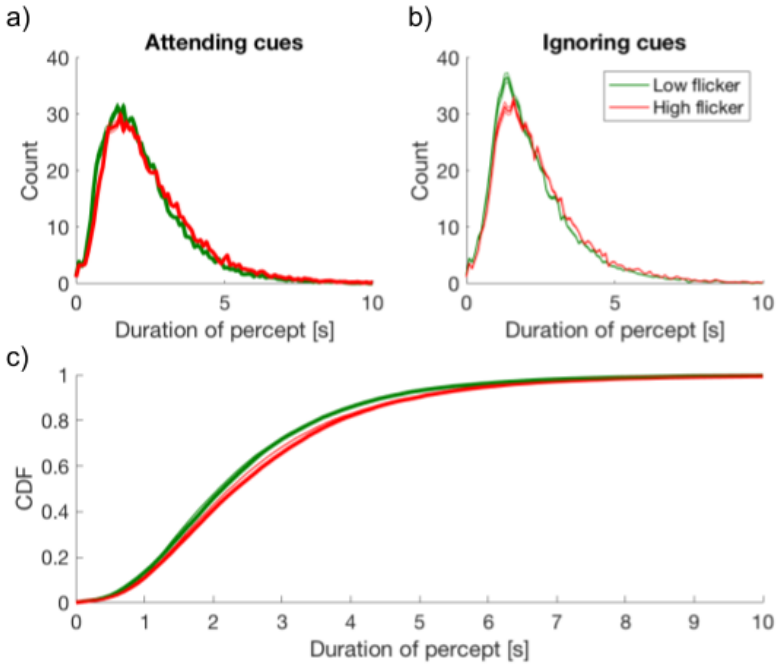

Figure 3.9. Results of binocular rivalry calibration procedure when attending and ignoring crossmodal cues.

a-b) Across the entire experimental session, when either attending or ignoring crossmodal cues binocular rivalry percept durations did not differ depending on flicker type (f1 vs f2). Coloured lines and their shading show mean and SE. Percept durations were taken across all experimental periods (Green (F1) 4.5 Hz flicker, Red (F2) 20 Hz flicker). c) cumulative-density function of the data in a-b. Thick lines show attend conditions, thin lines show non-attend (ignore cues) conditions. The thick and thin lines for both Low and High flicker largely overlap.

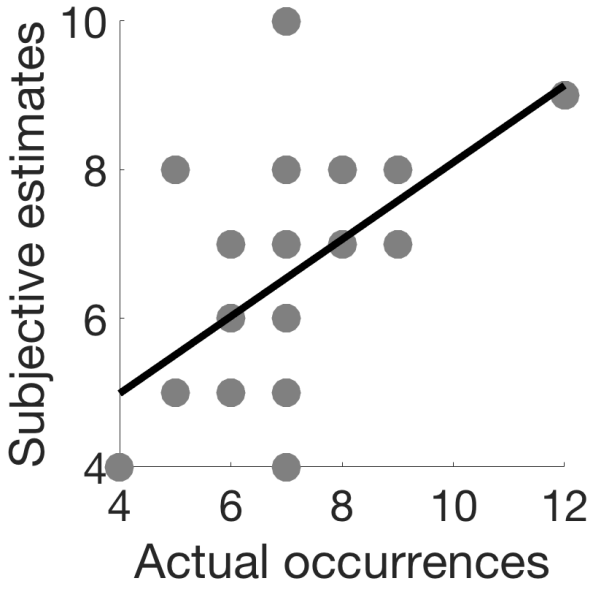

Figure 3.10. Definition of "attention to cues" in Figure 3.5b.

Y- and x-axes of this figure are the subjective and actual congruent crossmodal and visual stimuli in the attended sessions from one representative subject. Each data point is the subjective report after a single block, note that some datapoints overlap. The actual number of occurrences of crossmodal cues is not fixed, as it is defined by the ongoing moment-to-moment percepts of subjects during binocular rivalry. We defined the correlation coefficient between the actual occurrences and subjective estimates (here, r = .55) as the 'attention to cue' index used as x-axis in Figure 3.5b.

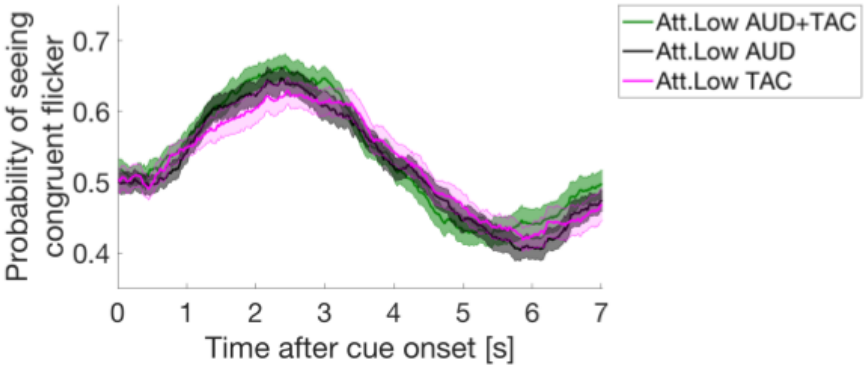

Figure 3.11. Attended low-frequency cue effects by crossmodal stimulus modality.

Green bars and shading show simultaneous auditory and tactile cues, black lines and shading show Auditory cues, magenta lines and shading show Tactile cues. There was no difference between cue modalities regarding the change in binocular rivalry dynamics (repeated measures ANOVA at each time point, corrected for multiple comparisons).

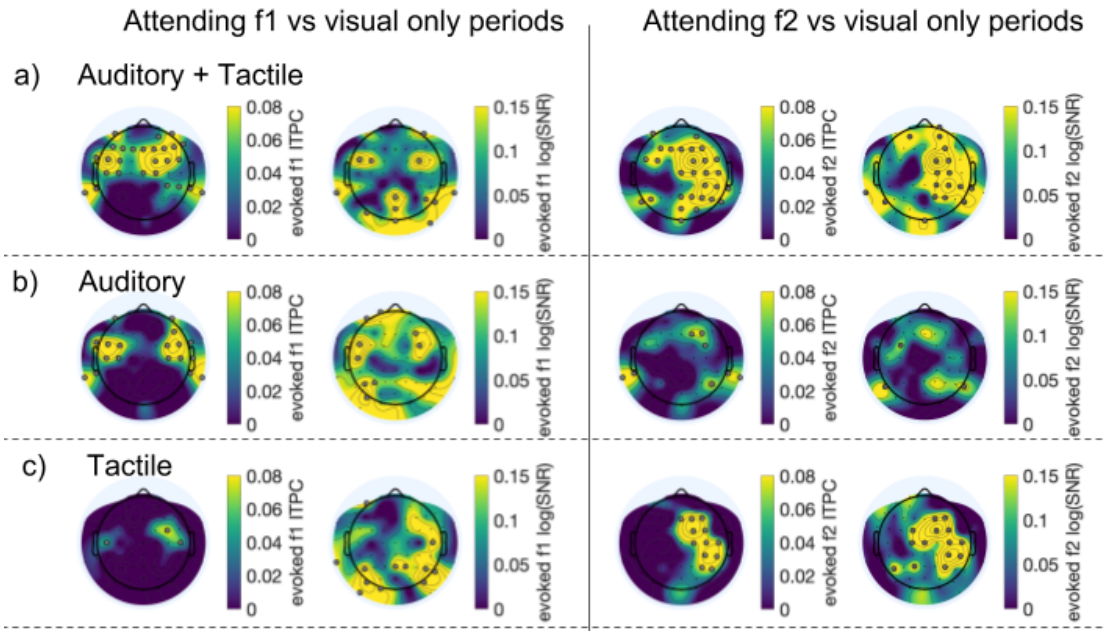

Figure 3.12. Evoked ITPC and SNR responses by crossmodal cue modality.

All panels show responses to attended low-frequency (f1; Left side) and attended high-frequency (f2; right side) crossmodal cues, compared to visual only periods (0 to 2 s after cue onset). Row a) shows all responses to simultaneous auditory and tactile cues, row b) to auditory cues, row c) to Tactile cues. Grey disks mark p <.05 uncorrected for ITPC values (interior left columns), and mark p < .01 uncorrected for log(SNR) values (interior right columns) due to differences in SNR affecting the cluster threshold (see methods).

Chapter 4: Attention periodically samples competing stimuli during binocular rivalry

Authors and affiliations: Matthew James Davidson[1, 2] *, David Alais[3], Jeroen J.A. van Boxtel[1,2,4*,+] , Naotsugu Tsuchiya[1, 2,*,+]

[1)] School of Psychological Sciences, Faculty of Medicine, Nursing, and Health Sciences, Monash University.

[2)] Monash Institute of Cognitive and Clinical Neurosciences, Monash University.

[3)] School of Psychology, The University of Sydney.

[4)] School of Psychology, The University of Canberra.

* corresponding author, lead contacts.

+ equal contribution

4.1. Article Introduction

In the previous study, attending to low-frequency crossmodal cues was shown to mediate the contents of consciousness during rivalry. This effect of attention on the content of binocular rivalry is noteworthy, as though attention has been argued to be necessary for binocular rivalry to occur (Brascamp & Blake, 2012; Zhang et al., 2011), attentional control over moment-to-moment percepts during binocular rivalry has been largely unsuccessful (Dieter, Brascamp, et al., 2016; Dieter et al., 2015; Meng & Tong, 2004; Paffen & Alais, 2011; yet see Dieter, Melnick, et al., 2016; Hugrass & Crewther, 2012).

In the previous chapter, neither increased phase coherence or SNR emerged as correlates of changes in perception. Here we explored a novel approach. The interaction between attention and crossmodal cues, combining to bring an invisible visual image into conscious awareness raises the exciting possibility that the locus of attention may be distinct from the current contents of consciousness. The next investigation searched for a recently established property of attention, in order to determine whether the allocation of attention could indeed be distributed away from a

conscious visual image. **Chapter 4** uses the same experimental data from **Chapter 3.** As such, the methods section has not been reproduced entirely, with only new analyses described in detail.

4.2. Abstract

The attentional sampling hypothesis suggests that attention rhythmically enhances sensory processing when attending to a single (~8 Hz), or multiple (~4 Hz) objects. Here we investigated whether attention samples sensory representations that are not part of the conscious percept during binocular rivalry. When crossmodally cued toward a conscious image, subsequent changes in consciousness occurred at ~8 Hz, consistent with rates of undivided attentional sampling. However, when attention was cued toward the suppressed image, changes in consciousness slowed to ~3.5 Hz, indicating the division of attention away from the conscious visual image. In the electroencephalogram, we found that at attentional sampling frequencies, the strength of inter-trial phase-coherence over fronto-temporal and parieto-occipital regions correlated with changes in perception. When cues were not task-relevant, these effects disappeared, confirming that perceptual changes were dependent upon the allocation of attention, and that attention can flexibly sample away from a conscious image in a task-dependent manner.

4.3. Introduction

Recent behavioural and electrophysiological evidence suggests that despite our seamless visual experience, incoming visual information is periodically enhanced for analysis in the visual system (VanRullen, 2016a, 2016b; Zoefel & VanRullen, 2017). This periodic sampling mechanism is proposed to result from the allocation of visual attention (Busch & VanRullen, 2010; Dugué, Roberts, & Carrasco, 2016; Dugué & VanRullen, 2017; VanRullen, Carlson, & Cavanagh, 2007; Zoefel & VanRullen, 2017), wherein alternating windows of high and low attentional resources operate to parcel incoming visual information, similar to the sequential frames that capture film within a video camera (Chakravarthi & VanRullen, 2012; VanRullen & Dubois, 2011). Whether stimuli are presented at the appropriate phase (Busch,

Dubois, & VanRullen, 2009; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009; VanRullen et al., 2007) or location (Dugué, McLelland, Lajous, & VanRullen, 2015; Dugué et al., 2016; Dugué & VanRullen, 2014, 2017; Dugué, Xue, & Carrasco, 2017; Huang, Chen, & Luo, 2015; Landau & Fries, 2012; Song, Meng, Chen, Zhou, & Luo, 2014) of this sampling mechanism has been shown to modulate the accurate detection of a visual stimulus, in stark contrast to our experience of an uninterrupted visual environment.

To date, primary neural evidence for the rhythmic gating of visual processing stems from the dependence of target detection on the pre-target phase of neural oscillations at approximately 7-8 Hz (Busch et al., 2009; Busch & VanRullen, 2010). These spontaneous fluctuations in detection may result from the allocation of visual attention toward a single location (Busch & VanRullen, 2010; Dugué et al., 2015; Spaak, de Lange, & Jensen, 2014; VanRullen, 2016a; Zoefel & VanRullen, 2017) and support the assumption that neural excitability cycles gate and filter incoming information for further processing (Schroeder & Lakatos, 2009; VanRullen, 2013; Zoefel & VanRullen, 2017).

This periodic gating of visual perception is also prominent behaviourally in the time-course of detection accuracy. Spectral analyses applied to high temporal resolution behavioural measures reveal 7-8 Hz modulations in performance following cues to reorient attention (Fiebelkorn et al., 2013), which slow proportionately when attention is divided between two or more locations (e.g. Chen, Wang, Wang, Tang, & Zhang, 2017; Holcombe & Chen, 2013; Huang et al., 2015; Landau & Fries, 2012; Landau, Schreyer, Van Pelt, & Fries, 2015; VanRullen, 2013). For example, Landau and Fries (2012) observed that following a cue to reorient attention to either the left or right visual hemifield, target detection oscillated at a 4 Hz counterphase rhythm depending on whether cues were congruent or incongruent with the target location. Critically, this counterphase sampling of visual information persisted at ~ 4 Hz when attention was directed to two locations on a single object (Fiebelkorn et al., 2013), and when cues to reorient attention were incongruent with target location – requiring a subsequent shift in the allocation of attention to a second location (Huang et al., 2015). These successive fluctuations in target detection and counterphase sampling between locations have led to the suggestion that an intrinsic ~7-8 Hz attentional rhythm can be allocated over space and time in a sequential manner (Dugué et al.,

2016; Dugué & VanRullen, 2017; Fiebelkorn et al., 2013; Holcombe & Chen, 2013; Landau & Fries, 2012; VanRullen, 2013; Zoefel & VanRullen, 2017).

Here, we tested if rhythmic attentional sampling is at play during binocular rivalry, when visual stimuli overlap regarding their retinal location, but differ based on their level of awareness. We hypothesized that evidence for distributed attentional sampling could be obtained during binocular rivalry, in support of a putative dissociation between the phenomena of attention and consciousness (Koch & Tsuchiya, 2007; van Boxtel et al., 2010b, 2010a). During binocular rivalry, incompatible images are presented to each eye which results in stochastic perceptual alternations, with one image visible at a time while the other is suppressed (Alais, 2012; Alais & Blake, 2005; Maier et al., 2012). In an experiment designed to induce or delay these transitions using auditory and tactile cues, we found that changes in consciousness were occurring rhythmically after the reorientation of attention. These fluctuations occurred depending on whether the crossmodal cue directed attention toward either the dominant or suppressed visual image, resulting in ~8 Hz and ~3.5 Hz oscillations, respectively. Critically, these rhythms were observed in both behaviour and the electroencephalogram (EEG), and were absent when cues were not task-relevant. This approximate halving of frequency suggests that when non-visual input is inconsistent with the ongoing visual percept, attentional sampling can flexibly orient away from a consciously perceived image, seemingly 'searching for' alternative sensory information to resolve the conflict.

4.4. Results

4.4.1. Attending to low-frequency crossmodal stimulation promotes the perceptual dominance of low-frequency flicker during binocular rivalry

We manipulated the conscious visibility of images across two sessions of 24 x 3-minute binocular rivalry blocks. Subjects (N=34; the same as **Chapter 3**) continuously reported the content of their visual consciousness via button press to indicate which image they currently perceived, while neural activity was simultaneously recorded via 64-channel EEG (see Methods). Rivalry stimuli were orthogonal sinusoidal gratings, which underwent sinusoidal contrast modulation, one

at 4.5 Hz and the other at 20 Hz (Figure 4.1). In each 3-minute block, we intermittently presented 12 crossmodal cues (mean duration 2.6 s), which were sinusoidally amplitude-modulated signals presented in the auditory and/or tactile modality (auditory, tactile, or combined auditory and tactile) at a frequency congruent with one of the visual stimuli (4.5 or 20 Hz). Three null cues (visual-only periods) without any crossmodal stimulation were also presented to increase the uncertainty of stimulus timing. The visual-only periods also served as a baseline to compare the behavioural effects of crossmodal cues (see below). We separated all cue periods by jittering the ISI between 7-10 s. As a result, the timing of crossmodal cues was completely independent to perceptual reports, and cues were presented at any point relative to the onset of the currently dominant percept (i.e., no closed-loop control).

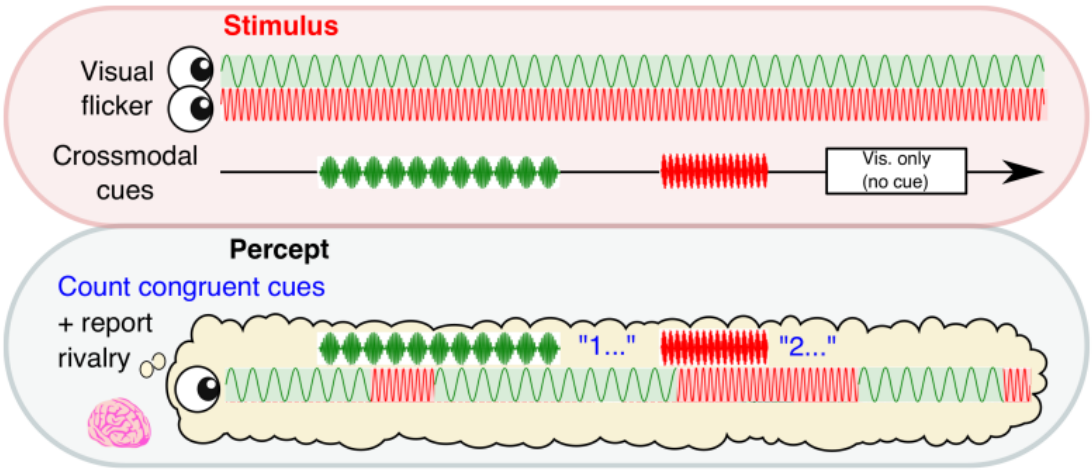

Figure 4.1. Experimental paradigm.

A schematic time course showing stimulus presentation and reported visual percept. Each eye was presented with a 4.5 or 20 Hz sinusoidal flicker throughout 3-min blocks. Subjects reported their perceptual state through button press. Crossmodal cues (also 4.5 or 20 Hz; 2, 3.1 or 4s in duration) or visual-only periods (2.6s in duration) were separated by inter-stimulus intervals of 7-10 s.

In order to investigate whether the allocation of attention to crossmodal cues alters the contents of visual consciousness during binocular rivalry, we varied attentional instructions over two sessions of the experiment. For one of their two sessions (day 1 for n=16, day 2 for n=18), we asked subjects to count the number of times that the temporal frequency of crossmodal cues coincided with their conscious visual percept at crossmodal cue offset (see Methods). For their other session,

subjects were instructed to focus on reporting their visual percept alone – ignoring any crossmodal cues.

Following the onset of a crossmodal cue, the probability of perceiving a congruent visual image increased only during attended low-frequency cues compared to all other cue types, during the period 0.68 to 3.97 s after cue onset (repeated measures ANOVAs followed by planned comparisons, FDR q = .05, Figure 4.2a). There was no difference in this effect when comparing the three types of crossmodal cues (Chapter **3; Figure 3.11**). To confirm that this effect was due to attention, we performed a correlation-based behavioural analysis. First, we computed the correlation coefficient (x-axis in Figure 4.2b), between each subject's verbally reported number of congruent cues (i.e., their attentional task during attend conditions), to the actual number of cues that were congruent with their visual percepts based on button-press data. Second, we defined the strength of the crossmodal cueing effect for attended low-frequency cues compared to other cue types (y-axis in Figure 4.2b), as the difference in the probability of seeing the congruent visual flicker during 1 to 4 s after cue onset. We call this the perceptual switch index (PSI), as it reflects the degree of perceptual switch after cue onset. The magnitude of these two variables displayed a strong positive correlation ($r(32) = .46$, $p = .006$, two-tailed), suggesting that the cross-modal cueing effect was indeed mediated by attention.

Due to the ongoing dynamics of binocular rivalry, this cueing effect can be calculated when visual and crossmodal information mismatched or matched at cue onset. When crossmodal cues mismatched with the visual percept at cue onset, the likelihood of switching to the previously suppressed, yet matched visual stimuli significantly increased for attended low-frequency cues compared to all other cue types over a time period from 0.62 to 4.12 s (FDR q = .05, Figure 4.2c). By contrast, when visual and crossmodal cues matched at cue onset, the effect of attending to low-frequency crossmodal cues delayed changes to the previously suppressed visual percept compared to all other cue types, over the period from 1.05 to 3.58 s (FDR q = .05, Figure 4.2d). Comparison against the visual-only cue period yielded the same conclusion, confirming that the attended low-frequency cues significantly influenced rivalry dynamics, while other cue types did not. As the overall crossmodal effects were unique to the attended low-frequency condition, we focused our subsequent attentional sampling and EEG analysis on this condition.

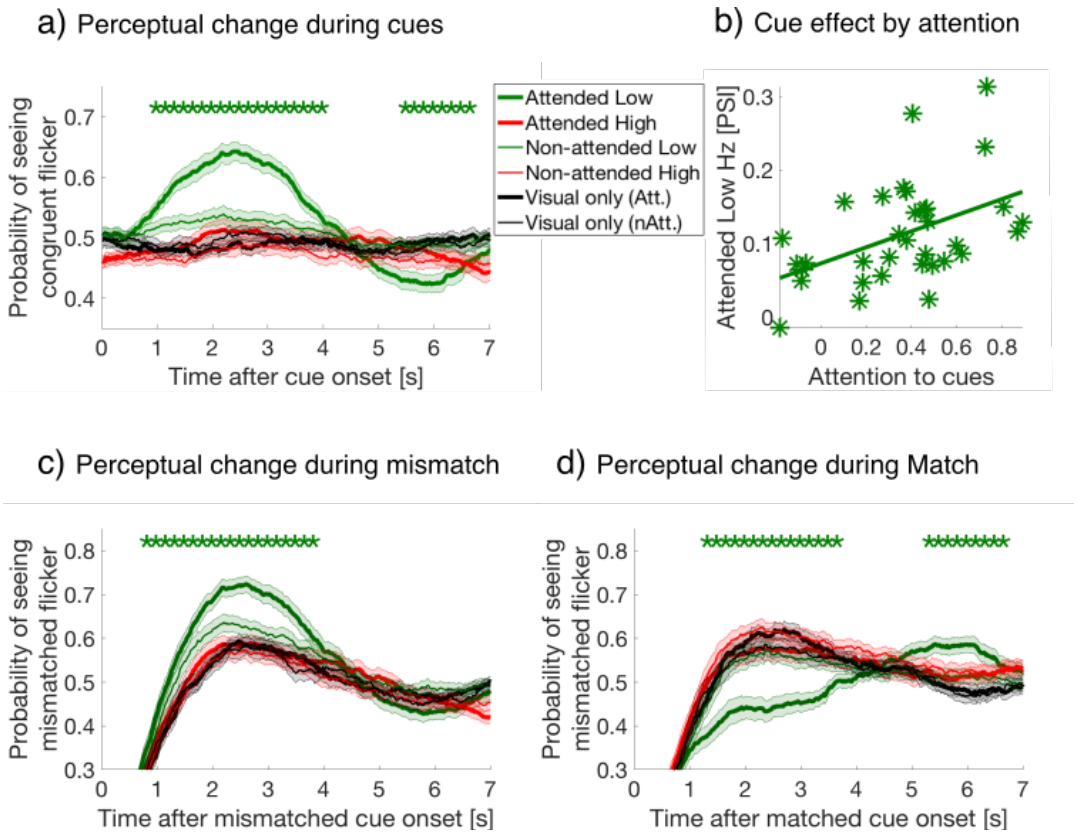

Figure 4.2. Behavioural results.

a) Button-press data, aligned at cue onset, were averaged over all crossmodal cue and visual-only periods per subject, then averaged over subjects for each cue condition. Y-axis represents the proportion of button-presses reporting congruent crossmodal and visual flicker at each time point, sampled at 60 Hz (or every 16.7 ms). Coloured lines and their shading show mean ±1 standard error across 34 subjects during attended and ignored cues (thick and thin lines) for low and high frequency (green and red colours). Black lines represent the equivalent probability for visual-only periods, serving as baseline (Methods). Asterisks indicate a significant difference between cues at each time point (repeated-measures ANOVA followed by planned comparisons). We use FDR q = .05 for the statistical threshold unless noted otherwise. b) Crossmodal effects are mediated by task-relevant attention. Our measure of crossmodal effects, the perceptual switch index (PSI, y-axis), is defined as the mean difference for the probability of seeing congruent flicker during 1-4 sec after the cue onset for attended-low-frequency cues (thick green in panel a) compared to other cue types. Attention-task performance (x-axis) is the correlation coefficient between the reported and actual congruent stimuli when comparing between rivalry percepts and crossmodal cues at offset (See Methods for details). The across-subject correlation between the two variables was strong (r(32) = .46, p = .006, two-tailed), demonstrating the crossmodal effects were strongly dependent on performance during the attention task. c) and d) Button-press data aligned at cue onset, with lines and shading as in panel a). Y-axis showing the proportion of button-presses reporting the mismatched flicker at each time point, after c) visual-crossmodal mismatch, or d) visual-crossmodal match at cue onset. Only the data of the attended-low-frequency condition differed significantly from visual only periods.

4.4.2. Binocular rivalry dynamics during attended low-frequency crossmodal cues

Our previous analysis showed that during attended low-frequency crossmodal cues, mismatched crossmodal cues lead to more perceptual switches, as the visually perceived image changed from high-frequency to low-frequency to become congruent with the crossmodal input. In the context of the attentional sampling hypothesis, we directly tested if these changes were occurring rhythmically after the reorientation of attention, and specifically investigated the timing of the *first* switch after cue onset, defined as the first change in button-state after cue onset.

To determine if cues affected the timing of first switches, we calculated the cumulative density function of each subject's first switches after cue onset (Figure 4.3a). Compared to visual-only cue periods, first-switches after cue onset occurred earlier for mismatched cues, indicating an earlier change to the congruent, previously suppressed, visual flicker. By contrast, following matched cues first-switches during rivalry were delayed, indicating an extended maintenance of the congruent visual percept when matched with attended low-frequency crossmodal cues. The facilitation of switches by mismatched cues was observed from 0.63 to 2.45 s and 3.78 to 6.87 s relative to cue onset, with matched cues delaying switches from 1.27 to 3.77 s after onset (paired samples *t*-tests, FDR q = .05, in Figure 4.3b).

After cue onset, the time-course for the probability of first switches displayed oscillatory patterns for mismatched and matched conditions (Figure 4.3c and d), but not the visual only condition (Figure 4.3e). Each data point represents the proportion of first switches which occurred at each time bin (16.7 ms intervals), calculated first per individual, and then averaged across subjects.

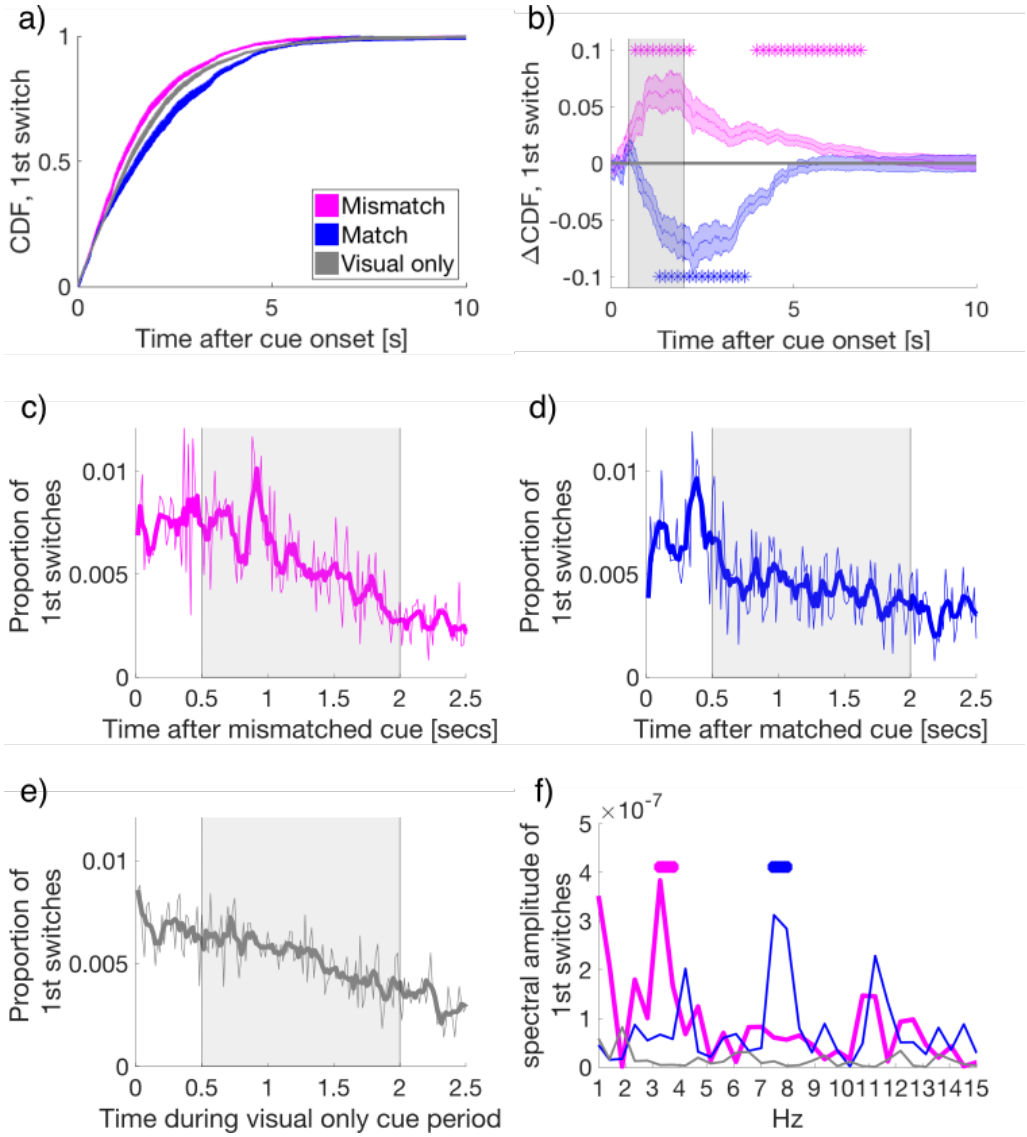

Figure 4.3. Binocular Rivalry dynamics during mismatched and matched cues.

a) The cumulative density function (CDF) of the time to first-switch. Mismatched, matched, and visual-only conditions are coloured in magenta, blue, and grey in all panels a-f. Lines and shading show mean and standard error across subjects (N=34) for a and b. b) The difference in CDFs between conditions. Asterisks mark statistical significance (paired-samples t-tests) comparing mismatched or matched cues to visual-only periods. FDR q = .05. c-e) The time course of the proportion of first switches made after cue onset in c) mismatched, d) matched, and e) visual-only conditions. Thin lines show the mean proportion of first-switches, binned in 16.7 ms increments and averaged across subjects. Thick lines show the smoothed data for visualization (83.5 ms sliding window). Grey shaded regions show the time window used for spectral analysis in f). f) The amplitude spectra for the time course of switches in conditions in c-e). Clustered asterisks indicate significant clusters (at least two neighbouring frequency bins) after permutation and cluster-based corrections for multiple comparisons (see Methods). The permuted null distribution and critical value for the identified clusters in f) are shown in Figure 4.8.

To quantify these patterns, we applied the Fourier transform to the period 0.5 to 2 s after cue onset (skipping the first 0.5 s to avoid an onset transient, see Figure 4.3, supplement 1) as performed by previous investigations of attentional sampling (Dugué et al., 2015, 2016; Fiebelkorn et al., 2013; Landau & Fries, 2012). For these analyses, we corrected for multiple comparisons by using non-parametric cluster-based permutations (Maris & Oostenveld, 2007), with thresholds set to $p < .005$ (Benjamin et al., 2017) for identification within a cluster, and a final critical value for significance set to $p = .05$, cluster corrected (see Methods).

Strikingly, when the temporal frequency of the cue matched the conscious visual flicker at cue onset, the first perceptual switches followed a 7.5-8 Hz rhythm ($p_{cluster} < .001$, Figure 4.3f blue). This rate is consistent with suggestions that attention samples sensory stimulation at a rate of approximately 7-8 Hz (Dugué & VanRullen, 2017; Fiebelkorn et al., 2013; VanRullen, 2013)., possibly due to fluctuations in neural excitability cycles (Schroeder & Lakatos, 2009; VanRullen, 2013; Zoefel & VanRullen, 2017). However, when crossmodal cues were mismatched with the dominant visual image at cue onset, the amplitude spectrum of perceptual switches peaked between at approximately half this rage, at 3.3-3.75 Hz ($p_{cluster} < .001$, Figure 4.3f magenta). This slower rhythm of perceptual changes is consistent with findings that show attention samples two locations at a rate of approximately 3.5-4 Hz (Fiebelkorn et al., 2013; Landau & Fries, 2012; Landau et al., 2015). No significant peaks were detected for the visual only condition (Figure 4.3f, grey), suggesting this effect was contingent upon the reallocation of attention in response to cues, rather than ongoing throughout rivalry. As to the remaining three cue combinations (attended high-, ignored low- and ignored high-frequency cues), all failed to exhibit any significant crossmodal effects on perceptual switches compared to visual only periods (shown Figure 4.2a, c, d, and Figure 4.9). Thus, we did not pursue further spectral or neural analyses of these conditions. We note that this analysis was performed on the averaged time-course, consistent with previous behavioural investigations of attentional sampling (e.g. Fiebelkorn et al., 2013; Landau & Fries, 2012). The pattern in individual participants is similar, though is not present for each individual. This is because the number of switches per condition when separating by attention/mismatch/frequency type was low, and the strength of attentional effects themselves varied across participants (Figure 4.2b).

4.4.3. The neural correlates of divided and focused attentional sampling

We hypothesized that at our behaviourally observed attentional sampling frequencies (3.5 and 8 Hz), we should be able to identify the neural correlates of attentional sampling in the EEG signal using an inter-trial phase coherence (ITPC) measure. Previously, the phase of ongoing cortical oscillations have been shown to be reset by external crossmodal events (Frey, Ruhnau, & Weisz, 2015; Lakatos et al., 2009; Romei, Gross, & Thut, 2012; Van Atteveldt, Murray, Thut, & Schroeder, 2014) and to modulate the probability of target detection (Busch et al., 2009; Landau et al., 2015; Mathewson et al., 2009; Thorne & Debener, 2014; VanRullen et al., 2007). To isolate the specific neural correlates of attentional sampling we compared the evoked ITPC, the increase in ITPC during 0 to 2 s after onset compared to -2 to 0 s before onset, in mismatched and matched cue conditions at the attentional sampling frequencies (3.5 and 8 Hz). Importantly, in these conditions, the physical sensory input was identical (i.e., attending low-frequency cues during binocular rivalry), with the only difference between conditions being the subject's percept at cue onset. Thus, any differences between conditions reflect differences due to the subjective visual percept matching or not with crossmodal cues.

For this analysis, we retained electrodes only after identification of a significant effect in evoked ITPC ($p < .05$, uncorrected) which also satisfied a spatial cluster-based criterion for selection and used non-parametric permutation distributions to control for multiple comparisons (Maris & Oostenveld, 2007; Figure 4.4, supplement 1). We tested Mismatch vs Match ITPC from 1 to 15 Hz (in 0.5 Hz increments), to match the behaviourally relevant frequencies of interest from our previous analysis. We found that at our attentional sampling frequencies, the mismatched cues induced stronger ITPC than the matched cues, at 3.5 Hz over right fronto-central-temporal electrodes [FT8, C6] (Figure 4.4a) and at 8 Hz over right parietal-occipital electrodes [P6, PO8] (Figure 4.5a). We note that a stronger ITPC for Mismatched cues also emerged at 9.5 Hz over right parietal-occipital electrodes [P4, P6,P8], as well as at 4.5 Hz over central midline electrodes [Cz, CPz]. Critically however, our time-window was long enough to distinguish the 3.5 from 4.5 Hz response (and 8 from 9.5 Hz; with half bandwidth = 0.5 Hz) to resolve the attentional sampling frequencies of interest. As mid-central theta (~5 Hz) is a well characterized

hallmark of task switching (reviewed in Gratton, Cooper, Fabiani, Carter, & Karayanidis, 2017) we continued by focusing our analysis on the novel attentional sampling ROIs located at right-fronto-central-temporal, and parieto-occipital electrodes. Figures 4.4b and 4.5b compare the evoked ITPC spectra in these regions based on mismatched and matched subjective percepts at cue onset, and confirm that our time window was long enough to resolve the frequencies of interest.

4.4.4. Attentional-sampling ITPC strength predicts perceptual outcome

Next, we investigated whether the evoked ITPC at the attentional-sampling frequencies in the above-identified regions (Figures 4.4a and 4.5a) predicted the magnitude of behavioural effects across subjects, shown in Figure 4.2c-d. To minimize any circularity problem (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009), we changed the time-window for our behavioural analysis to be temporally non-overlapping from the window used to define our ROIs. We also altered the critical comparison to compare ITPC responses between either mismatched or matched cue types (including visual only periods) instead of between mismatched and matched cues which defined our ROIs. More specifically, we again computed the difference in behavioural effects when comparing attended low-frequency to all other cue types (PSI; 2:4 s after cue onset), as a measure for the degree of perceptual change following mismatched and matched cues. Note that when considering a wider time-window (0:4s for behavioural effects, data not shown) a similar pattern of results was obtained, though weaker due to the lack of differences between cue types in early cue periods (i.e. 0:1s, cf. Figure 2c-d). We used the evoked ITPC from 0 to 2s after cue onset to restrict our analysis to within attended crossmodal cueing periods (which were 2, 3.1 and 4s in duration), and to capture the period where the majority of first switches were made after cue onset (Figure 2c and d). Similar to the PSI, we also subtracted the evoked ITPC across all other conditions from those in the attended low-frequency condition, and abbreviate this as the normalized ITPC (nITPC) below.

In the right fronto-central-temporal electrodes ([FT8, C6]) which significantly differed in 3.5Hz ITPC based on mismatched or matched percepts (Figure 4.4a), we found that 3.5 Hz nITPC and PSI were positively correlated for both mismatched ($r(32) = .38$, $p = .027$, two-tailed, Figure 4.4c), and matched cue types ($r(32) = .34$, $p = .049$, two-tailed, Figure 4.4d). Indicating that for both mismatched and matched

cues, increases in 3.5 Hz nITPC facilitated a change in visual consciousness across subjects (Figure 4.4c-d).

In the parieto-occipital electrodes ([P6, PO8]), we found that 8 Hz ITPC was not correlated with the PSI for mismatched cues (Figure 4.5a). However, 8 Hz ITPC was negatively correlated with the PSI during matched cues ($r(32) = -.39$, $p = .023$, two-tailed, Figure 4.5c), demonstrating that increased 8Hz nITPC resulted in fewer perceptual switches across subjects (Figure 4.5d).

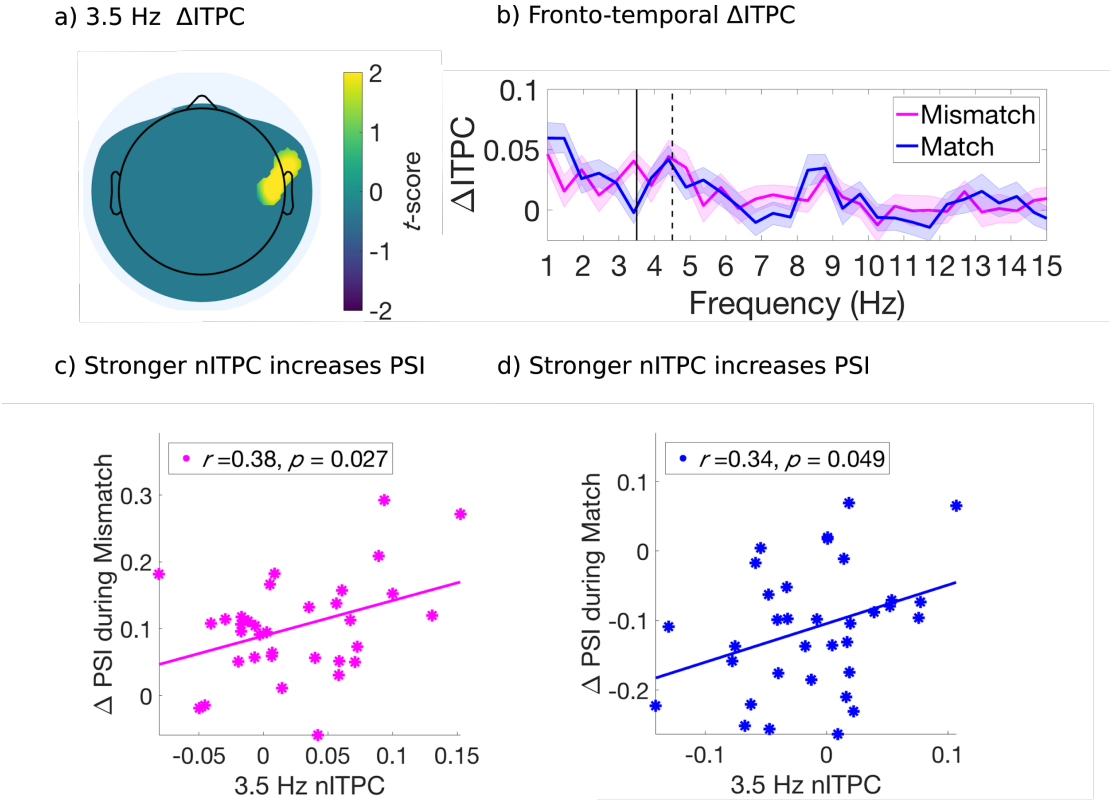

Figure 4.4. Evoked ITPC at 3.5 Hz mediates the probability of switches during mismatched and matched cues.

a) Significant differences in evoked ITPC between mismatched and matched cue conditions (multiple comparisons corrected using a cluster-based criterion; Methods). Non-significant electrodes after spatial-cluster based corrections are masked. b) Evoked ITPC spectra at significant regions in (a). The magenta and blue lines and their shading show mean ±1 standard error of the mean across 34 subjects for mismatched and matched cues, respectively. Solid and dotted vertical black lines mark the behaviourally observed attentional sampling frequency at 3.5 Hz, stimulus frequency at 4.5 Hz respectively. c, d): Stronger 3.5 Hz nITPC correlates with increased PSI during (c) mismatched and (d) matched conditions. The x and y-axes represent the normalized ITPC and perceptual switch index respectively (see text for definitions). Straight lines represent least-squares regression predicting PSI from nITPC.

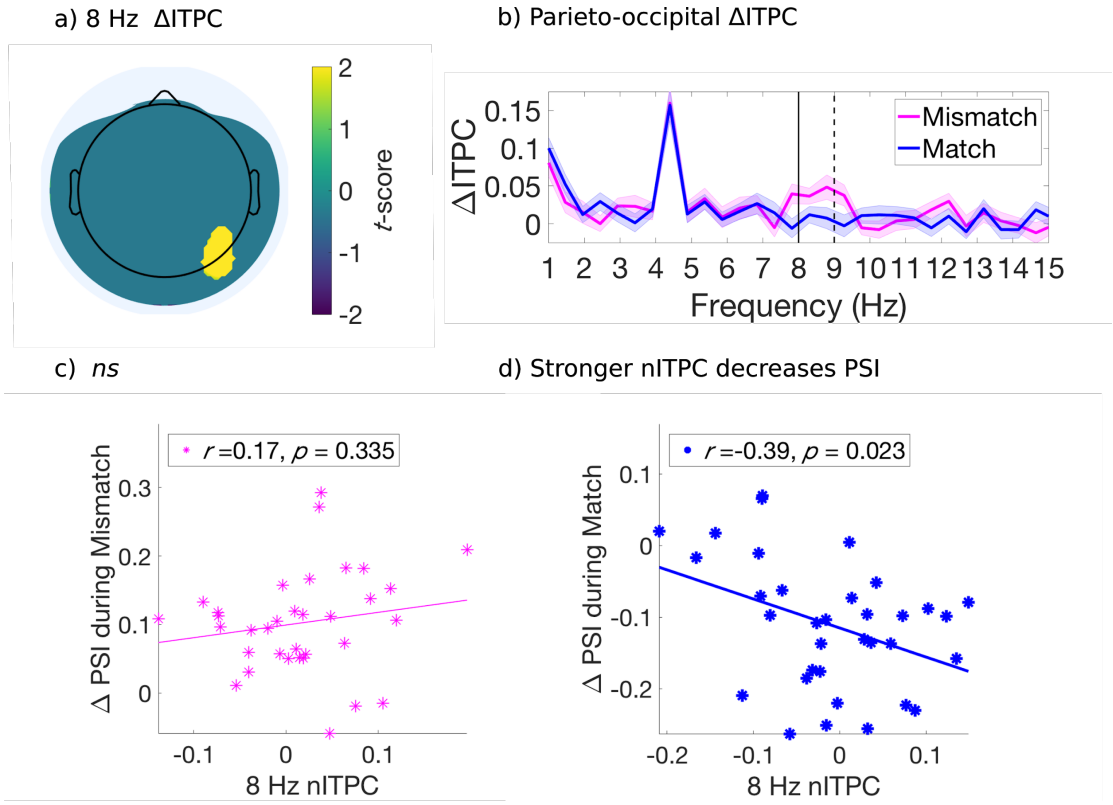

Figure 4.5. Evoked ITPC at 8 Hz mediates the probability of switches during matched cues only.

a) Significant differences in evoked ITPC between mismatched and matched cue conditions (multiple comparisons corrected using a cluster-based criterion; Methods). Non-significant electrodes after spatial-cluster based corrections are masked. b) Evoked ITPC spectra at significant regions in (a). The magenta and blue lines and their shading show mean ±1 standard error of the mean across 34 subjects for mismatched and matched cues, respectively. Solid and dotted vertical black lines mark in (b) the 8 Hz sampling frequency observed behaviourally and stimulus harmonic, respectively. c,d): Stronger 8 Hz nITPC correlates with a decreased PSI for (d) matched, but not the (c) mismatched condition. The x and y-axes represent the normalized ITPC and perceptual switch index respectively (see text for definitions). Straight lines represent least-squares regression predicting PSI from nITPC.

4.5. Discussion

Our findings provide novel evidence that attentional sampling exists during binocular rivalry, demonstrated in both behaviour and the electroencephalogram (EEG). Behaviourally, we replicated previous evidence that stimulus-driven cues can cause a switch to previously suppressed visual stimuli when mismatched with the current percept (to bring about congruence), as well as increase the maintenance of a

dominant visual image if cues matched perception (Figure 4.2; Dieter et al., 2015; Lunghi & Alais, 2015; Lunghi et al., 2014). Critically, we found distinct attentional sampling frequencies evident in the time-course of first switches during these cues (Figure 4.3). When crossmodal cues were incongruent in temporal frequency with the dominant visual stimulus, switches in visual consciousness occurred earlier, and within a distinct ~3.5 Hz rhythm. This 3.5 Hz rhythm is consistent with previous reports of divided attentional sampling between two locations (Fiebelkorn et al., 2013; Landau & Fries, 2012; Landau et al., 2015). However, when crossmodal cues were matched in temporal frequency to the dominant visual stimulus, changes in visual consciousness demonstrated an ~8 Hz rhythm, consistent with periodicities in behavioural measures observed when attending to a single visual location (Fiebelkorn et al., 2013), and suggestions that a cortical 7-8 Hz attentional rhythm may gate visual processing (Busch & VanRullen, 2010; Dugué & VanRullen, 2017; Fries, 2015). In the EEG (Figures 4.4, 4.5), distinct correlates of these divided and focused attentional sampling frequencies emerged over right fronto-temporal and right parieto-occipital sites, respectively, with ITPC strength at these frequencies correlating with the behaviourally reported change in consciousness across subjects.

Traditionally, top-down, voluntary attention has been thought to have limited control over perceptual dynamics during binocular rivalry; attention may alter dominance durations, but cannot halt the process of perceptual reversals entirely (Chong & Blake, 2006; Chong et al., 2005; Chopin & Mamassian, 2010; Dieter, Brascamp, et al., 2016; Dieter et al., 2015; Mitchell et al., 2004; Paffen & Alais, 2011 for bottom-up control, including crossmodal stimulation, see Conrad et al., 2010; Deroy, Spence, & Noppeney, 2016; Guzman-Martinez et al., 2012; Kang & Blake, 2005; Lunghi & Alais, 2013; Lunghi et al., 2010, 2014; van Ee et al., 2009). Our results clearly show additional dependence on the top-down deployment of attention, as without explicit instruction to attend to crossmodal signals, no facilitatory crossmodal effects emerged (see also Jack & Hacker, 2014; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010; van Ee et al., 2009). This interaction between low-level stimulus features (temporal frequency) and the allocation of attention indicates the facilitative role of both crossmodal stimuli (Deroy et al., 2014; Noppeney et al., 2016) and attention for perceptual transitions during binocular rivalry (Dieter, Brascamp, et al., 2016; Dieter et al., 2015; Dieter & Tadin, 2011; Paffen & Alais, 2011; Zhang, Jamison, Engel, He, & He, 2011). Our results are consistent with

previous research which has shown that exogenous feature-based cues can bias rivalry dynamics (Dieter et al., 2015; van Ee et al., 2009) and extend these reports by revealing an oscillatory basis to these changes in visual perception.

Our demonstration of these oscillatory changes in visual consciousness, which have been evoked by attended crossmodal cues, are also relevant to current computational models of binocular rivalry (e.g. Laing & Chow, 2002; Li et al., 2017; Wilson, 2003). No current model has accounted for the interaction we observe between attentional allocation and crossmodal stimuli, nor attention as an oscillatory process. Most recently, Li et al., (2017) have described a model for binocular rivalry incorporating attention and mutual inhibition. In their model, attentional modulation is dealt to the sensory representation that has the stronger sensory responses, by providing feedback to monocular-excitatory drives that otherwise increase monotonically with stimulus contrast. Building on our findings, future models could incorporate an oscillatory increase in excitatory drive as a result of periodic, rather than sustained attentional modulation (Fiebelkorn & Kastner, 2019; Fiebelkorn et al., 2018; Helfrich, 2018; VanRullen, 2018).

Previous behavioural investigations of attentional sampling have relied upon a brief cue to reorient attention, before estimating the time-course of target detection by densely sampling subject responses over closely spaced target-presentation intervals. Our design is unique in that 'target-detection' here is operationalized as the first reported change in visual consciousness for a continuously presented stimulus, resolved at 16.7 ms (or 60 Hz) from 500 ms to 2000 ms following cue-onset.

Past research has demonstrated approximately 7-8 Hz fluctuations in perceptual performance following the allocation of visual attention (Dugué et al., 2015; Fiebelkorn et al., 2013; VanRullen, 2013; VanRullen et al., 2007; Zoefel & VanRullen, 2017), commensurate with suggestions that cortical oscillations at approximately 7-8 Hz gate the content of visual perception (Busch & VanRullen, 2010; Dugué & VanRullen, 2017; Hanslmayr, Volberg, Wimber, Dalal, & Greenlee, 2013). In our binocular rivalry paradigm, we also observed changes in visual consciousness occurring within an 8 Hz rhythm, yet unique to when cues were congruent with the dominant visual percept at cue onset. By contrast, perceptual sampling has previously been observed at ~4 Hz when cues have encouraged dividing attention between two objects or locations (Dugué et al., 2016; Fiebelkorn et al., 2013; Huang et al., 2015; Landau & Fries, 2012; Landau et al., 2015; Song et al.,

2014). As such, the ~3.5 Hz rhythm we observed when crossmodal cues mismatched with the conscious visual percept extends the evidence for divided attentional sampling to binocular rivalry.

From our data alone, we cannot infer whether during conventional binocular rivalry, attention samples at 8 or 4 Hz. We surmise that increased attention to stimulus competition may be required to observe the attentional sampling rhythms we report here. Indeed, outside of attended cue periods we did not observe periodic behavioural responses (Figure 4.7), suggesting that attentional sampling is commensurate with sustained and goal-directed attention, instead of persisting throughout rivalry. We also note that the issue of trial-to-trial variability when reporting on perceptual changes cannot be completely avoided in binocular rivalry research, and is important to consider. Here, one might argue that variable timing in perceptual reports may blur any effects of temporal periodicity. However, our results clearly demonstrate that a change in perceptual state occurred for attended low-frequency cues that were unique in changing visual consciousness compared to visual-only baseline, at or near frequencies of attentional sampling that have been reported in previous literature. By comparison, the time-course of visual switches during visual-only periods did not exhibit periodic sampling, nor did the first-switches during non-attended, or high-frequency cues.

Distinct neural correlates of these attentional sampling rhythms were also found in the EEG. We found significantly greater 3.5 Hz ITPC strength for mismatched compared to matched cue types over right fronto-centro-temporal electrodes [FT8 and C6], suggesting this region may be a candidate neural correlate for divided periodic attentional sampling (Figure 4.4a). Accordingly, following both mismatched and matched cues, increased 3.5 Hz ITPC in this region also positively correlated with the likelihood of switching to the previously suppressed visual image across subjects (Figure 4.4c-d). Using visual-only stimulation, previous research has identified a pre-target ~4 Hz phase-dependency for peri-threshold perception when attention is divided across visual hemifields (Landau et al., 2015). While our right fronto-temporal region is different to those previously implicated in attentional sampling (e.g. Landau et al., 2015), we note that in our paradigm, attention was not divided between visual hemifields, but between competing stimuli during binocular rivalry. Right fronto-temporal regions have previously been implicated in the reorientation of attention to unattended locations (Corbetta & Shulman, 2002;

Downar, Crawley, Mikulis, & Davis, 2000; Proskovec, Heinrichs-Graham, Wiesman, McDermott, & Wilson, 2018), and most recently, Helfrich et al. (2018) have demonstrated that the phase of ~ 4 Hz dependent sampling in frontal and parietal areas determines visual perception. Taken together, our results support previous research indicating that a periodic attentional sampling mechanism modulates visual perception, here extending this finding into binocular rivalry when visual stimuli spatially overlap and compete for perceptual dominance.

We also found behavioural and neural correlates of focused attentional sampling during binocular rivalry when cues were consistent with the prevailing visual percept. Specifically, 8 Hz ITPC over parieto-occipital electrodes was negatively correlated with the likelihood of switching to the incongruent perceptual outcome (Figure 4.5d). Previously, phase-dependent peri-threshold perception has been reported for focused attention tasks in the visual domain (Busch et al., 2009; Busch & VanRullen, 2010; Hanslmayr et al., 2013; Mathewson et al., 2009), and has primarily implicated an approximately 7 Hz component located over fronto-central electrodes (Busch et al., 2009; Busch & VanRullen, 2010). Given the differences between paradigms, it is unsurprising that our identified region for focused attentional sampling does not coincide with those reported in previous research regarding phase-dependent perception. Particularly as our right-lateralized response may be due to the left-lateralized tactile input used to investigate crossmodal attentional sampling (though ITPC was not different among the three crossmodal stimulation types, data not shown). While promising, future experiments that control for this lateralization are needed to characterize the contributions of fronto-centro-temporal and parieto-occipital regions to this effect, particularly as activity over each of these regions has previously been implicated in the reorienting of visuo-spatial attention (Corbetta & Shulman, 2002; Downar et al., 2000; Dugué, Merriam, Heeger, & Carrasco, 2017; Proskovec et al., 2018), and for the integration of multisensory stimuli into a coherent percept (Beauchamp, 2005; Bushara et al., 2003; Calvert & Thesen, 2004; Driver & Noesselt, 2008; Zhang et al., 2011). Increases in right parieto-occipital theta power (4-8 Hz) have also been shown when attending to visual stimuli in the presence of auditory distractors (van Driel et al., 2014), with the phase of right parieto-occipital alpha (8-10 Hz) or theta (6-7 Hz) oscillations determining the perceptual outcome of multistable stimuli (Ronconi, Oosterhof, Bonmassar, & Melcher, 2017). As such, the present modulation for 8 Hz parieto-occipital ITPC is consistent with the idea that

right-parietal networks may preferentially represent temporal information in the visual modality (Battelli, Pascual-Leone, & Cavanagh, 2007; Guggisberg, Dalal, Schnider, & Nagarajan, 2011).

Reporting binocular rivalry switches involves both a change in perception, and a decision to press the button. Accordingly, the attentional sampling in binocular rivalry we report here may reflect the fluctuation of perception or of decision criterion. Recent studies of behavioural oscillations that have employed signal detection theory have reported that sensitivity and response criterion both exhibit oscillations (at distinct frequencies) in the high theta/low alpha band, for both vision (Zhang, Morrone, & Alais, 2019) and audition (Ho, Leung, Burr, Alais, & Morrone, 2017a). Consequently, whether our oscillations reflect perceptual or decision-level effects must be clarified (Ho, Leung, Burr, Alais, & Morrone, 2017b; Iemi & Busch, 2018; Iemi, Chaumon, Crouzet, & Busch, 2017; Limbach & Corballis, 2016). Our paradigm as it is cannot resolve this, although a future investigation using our paradigm combined with signal detection theory could do so.

Our analysis has so far revealed that when crossmodal cues mismatched the dominant binocular rivalry stimulus, that rates of attentional sampling slowed to ~ 3.5 Hz – implicating the division of attention over multiple locations. However, our exogenous cues oriented attention toward the congruency of visual and crossmodal stimuli, prompting the question: between what was attentional sampling divided? One possibility is that attentional sampling during mismatched cues was divided between two sensory modalities, as the brain tried to resolve a conflict between concurrent auditory/tactile and visual information. Figure 4.6a provides a schematic of this multisensory interpretation. If the neural activity in our identified region is representative of divided sampling between modalities, it constitutes the first evidence that an attentional sampling mechanism can flexibly orient between temporally co-modulating crossmodal stimuli. Although the facilitative role of attention in multisensory integration remains controversial (Hartcher-O'Brien et al., 2016; Talsma et al., 2010), we see it as a viable possibility that this mechanism resolved perceptual ambiguity through a visual perceptual switch to the competing image, rendering the multisensory stimuli congruent.

As only attended, low-frequency modulated cues enabled a change in visual consciousness, we must consider whether the lack of a high-frequency effect reflects an upper limit in temporal frequency on crossmodal interactions or attention. Such a

limit on crossmodal interactions may explain why we observed low- but not high-frequency behavioural effects in the present task, and is supported by previous investigations regarding the binding of multisensory stimulus attributes (Fujisaki & Nishida, 2005, 2010; Lunghi et al., 2014; Vroomen & Keetels, 2010), and the limits of crossmodal temporal judgments (Fujisaki & Nishida, 2005, 2010; Holcombe, 2009; Vroomen & Keetels, 2010). For example, Fujisaki and Nishida (2005) have shown that judgments of temporal synchrony between rhythmic sensory streams degrade above ~4 Hz. It is plausible that the ineffective crossmodal cueing that we found is related to the above-mentioned findings, reflecting a limit on crossmodal integration processes, rather than attention.

Having said that, one previous study using a similar design to ours was successful in eliciting a high-frequency crossmodal effect (15-20 Hz; Lunghi et al., 2014), and notably, in the absence of explicit attentional demands. However, these differences are not wholly unexpected, as to optimize the present task for EEG recordings we used larger (6.5° visual angle) luminance-modulated sinusoidal gratings to facilitate subsequent steady-state visually evoked potential analyses. While in comparison, Lunghi et al succeeded in showing a high-frequency effect with rivalry stimuli that were contrast-modulated narrow-band random noise patterns (3.2° visual angle), and did so under conditions analogous to our non-attend conditions. This difference in the composition of visual stimuli is noteworthy, as stimulus size is known to strongly affect rivalry dynamics (Blake, O'Shea, & Mueller, 1992). To our knowledge, whether stimulus size impacts upon crossmodal effects during binocular rivalry is unknown. However, given the strength of our results for attended low-frequency flicker (Figure 4.2a), we note that the low- and high-frequency effects observed by Lunghi et al (2014) are not generalizable to the larger and attended-rivalry stimuli employed here. Similarly, whether the type of stimuli (e.g., gratings vs random noise patterns) also impacts upon crossmodal effects during rivalry represents a fruitful endeavour for research, particularly given the novel possibility of distinguishing between crossmodal and attentional limits on attentional sampling.

An alternate possibility to crossmodal attentional sampling is that the 3.5 Hz rhythm in our paradigm reflects divided attentional sampling between dominant and suppressed visual images during binocular rivalry (Figure 4.6b). The frequency of divided attentional sampling that we observed is consistent with those obtained when

visual attention has been divided between two objects or locations (Fiebelkorn et al., 2013; Landau & Fries, 2012). As our binocular rivalry stimuli necessarily occupied the same spatial location, attention in our paradigm was likely divided between either features or objects, instead of locations. Indeed, feature-based attention has already been shown to modulate neural processes when an attended target is suppressed during continuous flash suppression (Kanai, Tsuchiya, & Verstraten, 2006). During binocular rivalry, perceptual dominance is also influenced by object-based attention (Mitchell et al., 2004), with unconscious selection mechanisms argued to mediate perceptual transitions (Lin & He, 2009). This second alternative is also indirectly supported by the temporal limits of binocular rivalry when conflicting visual stimuli are presented asynchronously, without temporal overlap between the two eyes (O'Shea & Blake, 1986; van Boxtel, Knapen, Erkelens, & van Ee, 2008; van Boxtel, van Ee, & Erkelens, 2007). The maximum stimulus onset asynchrony that can sustain this type of rivalry is approximately 350 ±50 ms, beyond which alternating stimuli introduced to one eye are perceived immediately, without rivalry occurring (van Boxtel, Alais, Erkelens, & van Ee, 2008). This limit is consistent with a 7-8 Hz attentional sampling rhythm distributed between the two conflicting stimuli (each sampled at ~3-4 Hz). When stimuli are presented rapidly enough, they are temporally bound together and can engage in ongoing rivalry; when stimuli are presented slower than at 3-4 Hz, they are temporally individuated by attention, and rivalry ceases. A recent computational model that explicitly modelled time-varying attention could indeed reproduce this finding (Li et al., 2017), suggesting that attention is the process that temporally binds the successive stimulus presentations together. Based on our findings, we propose the persistence of percepts may also be modelled using an oscillatory, and feature-selective attentional mechanism (Li, Carrasco, & Heeger, 2015; Li et al., 2017).

The suggestion that attention can sample between conscious and nonconscious vision is also consistent with a view that the underlying neuronal processes for attention and consciousness are supported by distinct neural mechanisms (Bahrami et al., 2007; Smout & Mattingley, 2018; Watanabe et al., 2011; for review see Koch & Tsuchiya, 2007). We note that while attentional sampling of a suppressed image suggests that attention is not sufficient for consciousness (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Koch & Tsuchiya, 2007; Lamme, 2003; van Boxtel, Tsuchiya, & Koch, 2010a, 2010b), this interpretation remains consistent with

a view that attention may still be necessary for conscious perception (Chica & Bartolomeo, 2012; Cohen & Dennett, 2011; Merikle & Joordens, 1997; O'Regan & Noë, 2001; Posner, 1994, 2012).

Whether attributable to conscious-nonconscious, or visual-crossmodal attentional sampling, the present results also complement the 'active-sensing' hypothesis (Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010), whereby perceptual selection is determined by routine exploratory behaviours. According to the active-sensing hypothesis, attention is critical to 'search for' task-relevant information from the environment (Schroeder et al., 2010), particularly via the rhythmic coordination of multisensory information (Schroeder et al., 2010; Thorne & Debener, 2014). Intriguingly, early contributions from multi-sensory (non-visual) information have been shown to modulate perception (Morillon, Schroeder, & Wyart, 2015; Schroeder et al., 2010; Thorne & Debener, 2014). The rhythmic modulation of visual performance has also been demonstrated to follow the onset of both voluntary (Hogendoorn, 2016), and preparatory motor behaviours (Tomassini, Ambrogioni, Medendorp, & Maris, 2017; Tomassini, Spinelli, Jacono, Sandini, & Morrone, 2015). Here, in further support of the active-sensing hypothesis, we have shown that task-relevant crossmodal information can change the rhythmic modulations of perceptual selection during competition for perceptual dominance.

In summary, here we have provided novel evidence in support of attentional sampling during binocular rivalry through the use of crossmodal cues matched to either a conscious or nonconscious visual stimulus. As the attention sampling hypothesis continues to garner traction from various psychophysical and neuronal paradigms (Fiebelkorn et al., 2018; Helfrich et al., 2018; VanRullen, 2016a, 2016b, 2018), future targeted experimentation can confirm whether attention can indeed sample across modalities (Figure 4.6a), as well as if attention can sample between conscious and nonconscious neural representations during binocular rivalry (Figure 4.6b). The interactions between crossmodal stimuli and conscious perception represent a fruitful avenue for experimentation (Faivre et al., 2017), here uncovering the previously unknown dependence of attention and consciousness on rhythmic neural dynamics of the human brain.

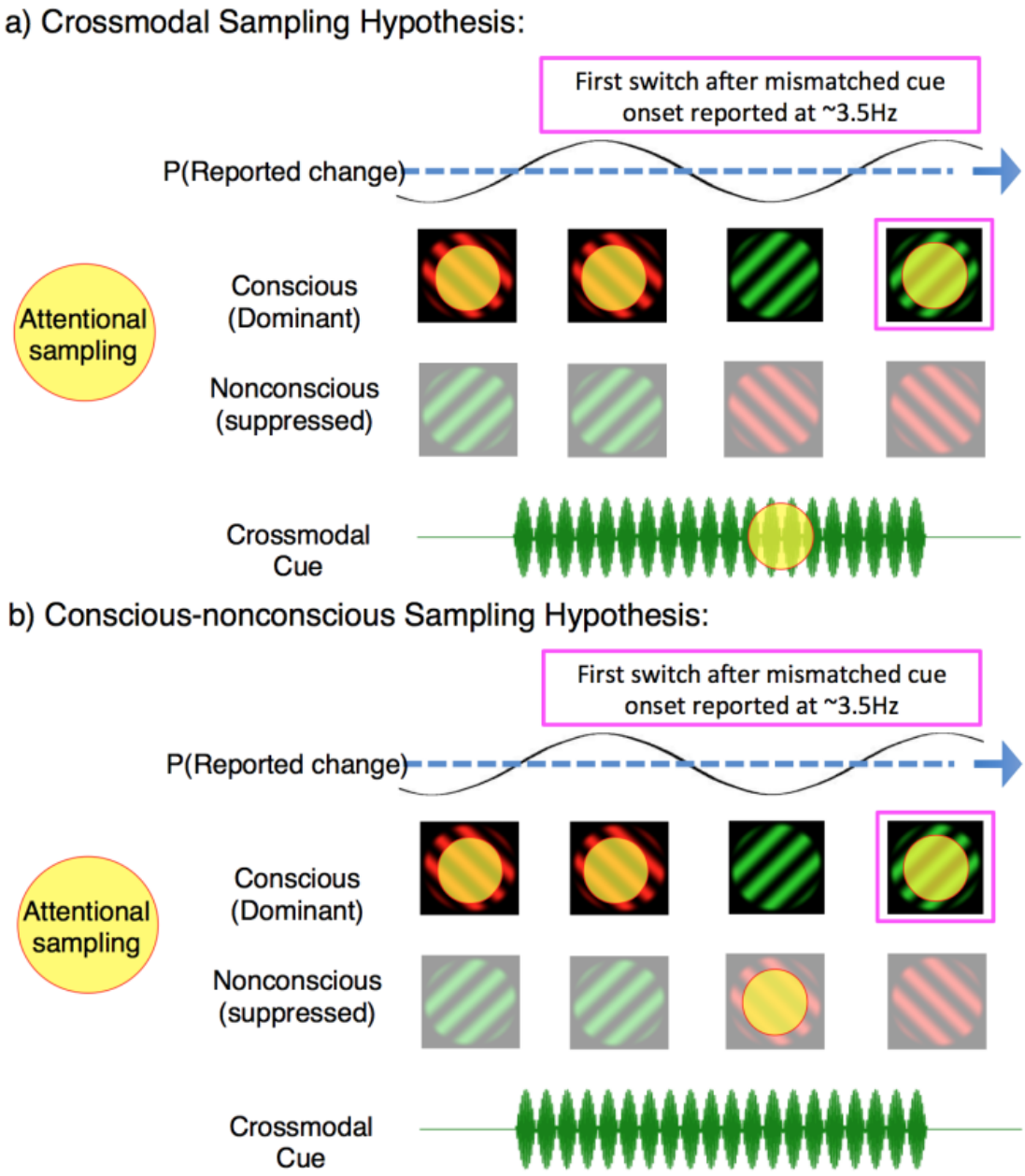

Figure 4.6. Two possible interpretations of attentional sampling during mismatched crossmodal cues

Schematic representation of attentional sampling and perceptual oscillations during binocular rivalry. a) Crossmodal sampling hypothesis: While perceiving the high-frequency visual flicker, an attended low-frequency crossmodal cue mobilises attention to sample between the dominant image and mismatched crossmodal cue at ~3.5 Hz. As a consequence, the likelihood of the first perceptual switch is modulated at ~3.5 Hz. b) Conscious-nonconscious sampling hypothesis: The onset of a mismatched cue prompts attention to sample between separate visual features, which in our paradigm consists of dominant and suppressed visual images. We do not suggest that these are the only mechanisms of attentional sampling during binocular rivalry, and only illustrate the interpretations discussed.

4.6. Methods

The subsections for Participants, Apparatus and Stimuli, Stimulus timing, Calibration of visual stimuli, Calibration of auditory stimuli, Experimental procedure and behavioural analysis, Evaluation of attention-on-task, EEG recording and analysis are as described in **Chapter 3**. The same participants were used in both **Chapters 3** and **4.** Additional relevant methodological information for this published manuscript follows below.

4.6.1. Behavioural data analysis

For crossmodal cue and visual-only periods, the calculation of congruent button-press was identical as previously described. In addition, we also defined Mismatched (Figure 4.2.c) or matched (Figure 4.2.d) condition comparisons by whether the congruent button (left-eye dominant) was pressed at cue onset.

In Figure 4.2.c and d, we set the y-axis for 'Probability of seeing mismatched flicker', to reflect the probability of perceptual states that differ in temporal frequency from the crossmodal cue. In Figure 4.2.a, c, and d, we compared among six conditions with one-way repeated-measures ANOVAs: 1,visual-only on attend days; visual only on non-attend days; 3, attended low-frequency; 4, attended high-frequency; 5, unattended low-frequency; and 6, unattended high-frequency. We defined significant differences among conditions at those time points that survived corrections for multiple comparisons with planned comparisons between cue types and the visual-only baseline, using FDR at q = .05 (Benjamini & Yekutieli, 2001).

4.6.2. Perceptual Switch Index (PSI)

The PSI was calculated as described in **Chapter 3**, as the difference in the probability of viewing a congruent visual flicker over the period 1-4 s after stimulus onset. The same subtraction was used to compare the probability of viewing the previously suppressed visual flicker following mismatched (Figure 4.4.c, 4.5.c) and matched cues (Figure 4.4.d, 4.5.d), for the period 2-4 s after onset. This shorter time window was selected to capture when the crossmodal effects on binocular rivalry

emerged for mismatched and matched cues. A similar pattern to the results displayed in Figures 4.4 and 4.5 was shown when a wider window was used (e.g. 0-4 s, data not shown).

4.6.3. Spectral analysis of first switches

For our spectral analysis (Figure 4.3), we focused on the first perceptual switches, which were the first time-point recording a change in button-press state after cue onset. To account for individual variation in the number of overall switches, the proportion of switches at each time point was first calculated at the subject level, before averaging across all subjects. We sampled button presses at 60 Hz (or every 16.7 ms). For the spectral analysis of perceptual switches (Figure 4.3f), we applied a single-taper fast Fourier transform (FFT) to the period 0.5 - 2 seconds after cue onset (Nyquist = 30 Hz, a half bandwidth = 0.67 Hz). This window was selected to restrict the analysis so that all the analysed trials occurred during an attended cueing period (as the minimum crossmodal cue duration was 2 seconds), and to remove transient button presses occurring early in the cue period, which were unlikely to be caused by crossmodal match or mismatch (Figure 4.3, Figure 4.7). We display the frequency range of 0 - 15 Hz for all conditions, as no higher frequencies (above 9 Hz) were significant after our two-stage statistical criteria.

4.6.4. Statistics on spectra of first switch timing

To assess the statistical significance of behavioural spectra we used a two-stage statistical testing procedure as applied in previous investigations of attentional sampling (Landau & Fries, 2012) and electrophysiological research (Maris & Oostenveld, 2007). At the first stage, we first detected significant frequencies (at $p <$ .005 uncorrected) through a non-parametric randomization procedure. In this procedure, after obtaining the spectral amplitude for the observed data across subjects, we generated a null distribution of first switches during the same cue period by randomly shifting switch-times within each subject, while keeping the number of perceptual switches the same. We generated 5000 surrogate datasets in this way, to test the null hypothesis that there were no temporal effects on the timing of perceptual switches. We then compared the amplitude of the Fourier transform from the

observed and the surrogate data at each frequency. We regarded the spectral amplitude at a certain frequency to be significantly above chance, if the observed spectral amplitude exceeded the top 99.5% of the null-distribution of amplitudes at each frequency generated by surrogate data.

At the second stage, we applied a cluster criterion, which corrects for multiple comparisons across multiple frequencies through a permutation procedure (Maris & Oostenveld, 2007). We required that the first-level significance ($p < .005$ uncorrected) be sustained for at least two neighbouring frequencies, and retained the sum of their clustered test-statistics (amplitudes in this case) as our observed data. Then, from our surrogate dataset, we calculated the maximum cluster-based amplitudes per surrogate (maximum spectral amplitude excluding 0-1 Hz and nearest neighbour), which we retained as the null-distribution to compare against our observed data. Candidate clusters survived this second order analysis when their observed cluster-based test-statistics exceeded the top 95% of the null distribution, or corrected to $p_{cluster} < .05$ if testing across multiple clusters. The null-distributions obtained for our frequencies of interest in Figure 4.3.f are shown in Figure 4.8.

4.6.5. EEG recording and analysis

EEG was recorded and pre-processed as described in **Chapter 3**, with time-frequency analyses performed using the Chronux toolbox (http://chronux.org; Bokil, Andrews, Kulkarni, Mehta, & Mitra, 2010), and custom MATLAB scripts. To resolve our frequencies of interest (especially between 3.5 and 4.5 Hz), we used a single-taper Fourier transform with a time-window of 2 seconds, which resulted in a half bandwidth (W) of 0.5 Hz ($W = 1/2$). This half bandwidth is consequently capable of resolving differences between 3.5 and 4.5 Hz, as demonstrated in Figure 4.4.b and 4.5.b.

4.6.6. ITPC analysis

To assess the neural correlates of attentional sampling (Figures 4.4, 4.5), we analysed the inter-trial phase coherence (ITPC) within electrodes, over multiple time-frequency points (Bastos & Schoffelen, 2016). Our ITPC measure was computed as described in **Chapter 3**: For a given time, t, and frequency, f,

$$ITPC(t,f) = \left|\frac{1}{N} * \sum_{n=1}^{N} e^{i(\theta(t,f,n))}\right|$$

where N is the number of trials, and θ is the phase angle at time t, frequency f, and trial n.

Due to the stochastic nature of perceptual alternations during binocular rivalry, the number of available trials for analysis in each mismatched and matched cue type ranged from 12 to 36 trials across subjects (after averaging first across subjects, the mean number of trials was 24 (±1.5) trials across matched / mismatched and attention conditions). Because the bias level (or expected chance level for pure noise data) of ITPC is strongly influenced by the number of trials, we took additional measures to equate the number of mismatched and matched cue types for the analysis. To achieve this aim, the minimum number of trials recorded for a given cue combination was identified across subjects. Following this, subjects with greater numbers of trials had their observed number of trials supplemented by down-sampling with replacement from their recorded trials, equating them to the predefined minimum for each condition. Upon this resampled dataset, the ITPC was computed, and this process repeated 100 times. As the difference in ITPC between auditory, tactile, and combined auditory and tactile cues was not significant, we proceeded by combining crossmodal cue types within each subject.

4.6.7. ITPC statistics

To investigate the neural correlates of attentional sampling, we analysed evoked ITPC, the increase in ITPC during 0 to 2 s after onset compared to -2 to 0 s before onset. Similar to our statistical approach for the behavioural spectral analysis described above, we used a two-stage statistical testing procedure for this analysis. At the first stage, we performed a *t*-test (two-tailed) comparing whether evoked 3.5 and 8 Hz ITPC differed between mismatched and matched conditions across subjects at each electrode. At each electrode, we used the mean evoked ITPC value obtained from the down-sampling method described above. As a result of the *t*-tests, if we found a cluster of at least two neighbouring electrodes with the same polarity in the t-score with $p < .05$ (uncorrected), where inter-electrode distance did not exceed 3.5

cm, we proceeded using this cluster in the second stage of statistics. As a result of this cluster criterion, we always identified a minimum size of 2-electrode clusters (Figure 4.4a and 4.5a).

At the second stage, we first computed the absolute value of the sum of observed *t*-scores within the identified cluster, which we retained as our observed test-statistic (Figure 4.10). To create the null distribution, condition labels (mismatch and match) were randomly shuffled for each electrode within each subject, to create two surrogate datasets the same size as our original mismatch and match conditions. Then the *t*-scores were computed for each electrode based on our surrogate datasets, and the electrode with the maximum *t*-score and the maximum *t*-score of its neighbours retained. The sum of these *t*-scores were then retained per shuffle, and this procedure repeated 2000 times to obtain a null distribution of the sum of *t*-scores around the maximum electrode for each shuffle of our surrogate data. Against this distribution, the sum of observed *t*-scores for the candidate cluster was then compared. When the observed sum of *t*-scores was within the top 5% (or cluster corrected to $p < .05$) then we concluded that there was a significant difference between mismatch and match conditions. The null-distributions and observed test-statistics produced by this analysis are shown in Figure 4.10.

4.7. Supplementary Figures

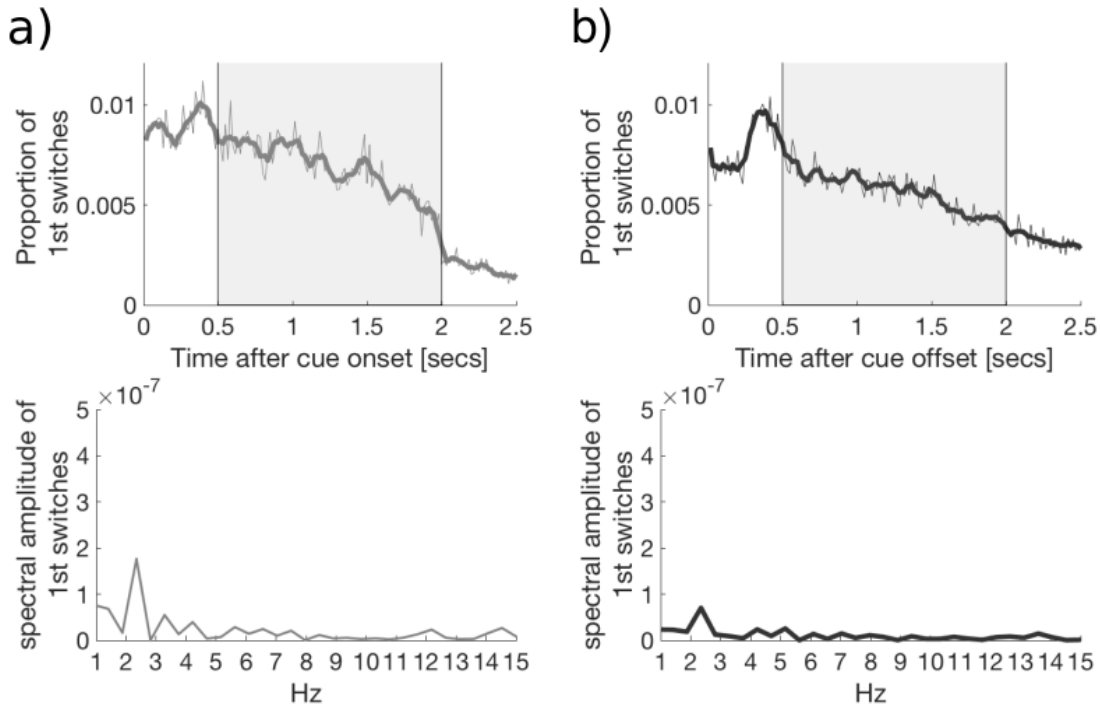

Figure 4.7, related to Figure 4.3.

The time course of the proportion of first switches made at each time point; following any crossmodal cue onset (a), and following crossmodal cue offset (b). An analysis of the time-course of perceptual switches reveals no significant spectral peaks at 3.5 or 8 Hz. Y-axis scaled as per Figure 4.3f. The presence of an early peak (0 - 0.5 s) in the proportion of first switches suggests that these changes may be due to stimulus transients, rather than the cue-conditional allocation of attention. As such this early time-window was omitted from subsequent analysis.

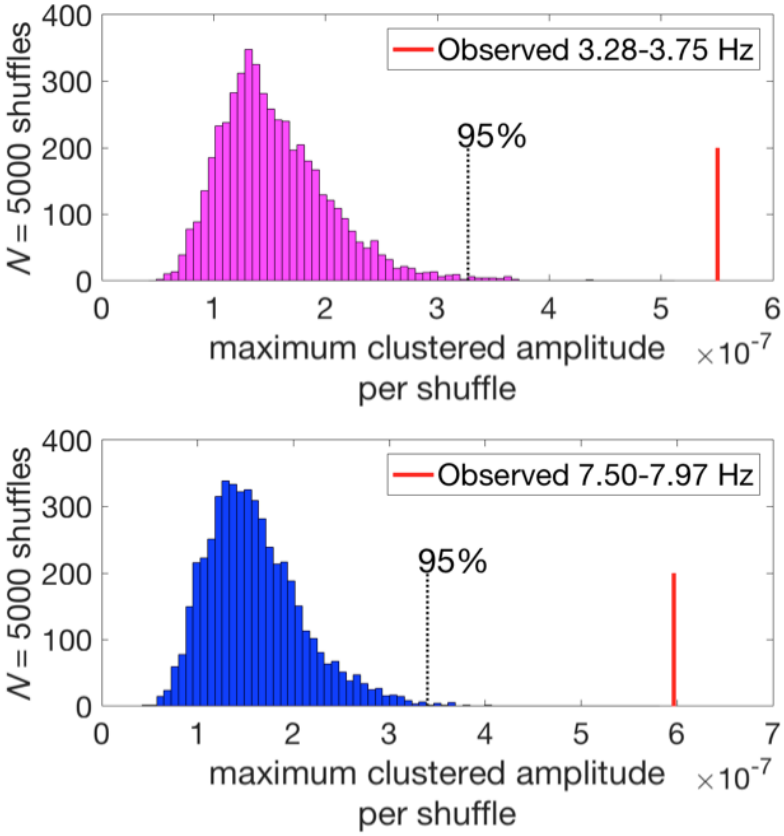

Figure 4.8, related to Figure 4.3.f.

The null-distributions for the surrogate datasets generated by the randomization procedure, and the actually observed values of second-stage statistics (i.e., maximum and its highest neighbour's summed Fourier amplitude). After satisfying first-level criteria ($p < .005$ uncorrected for two neighbouring frequencies), we proceeded to this second-stage statistical test. The observed second-stage statistics (red line) were regarded as significant after cluster corrections for multiple comparisons at $p < .05$ level; exceeding the top 95% of the null distribution.

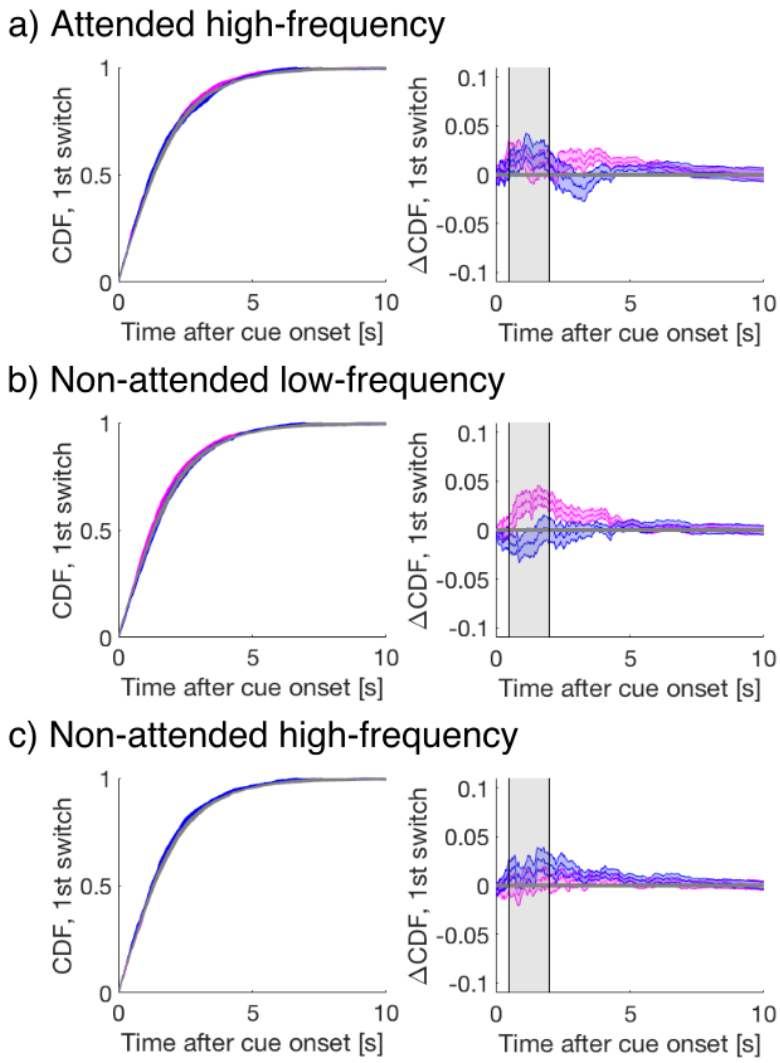

Figure 4.9, related to Figure 4.3.

Left column) The cumulative density function (CDF) of the time to first-switch for all conditions other than attended low-frequency. Mismatched, matched, and visual-only conditions are coloured in magenta, blue, and grey in all panels. Lines and shading show mean and standard error across subjects (N=34). Right column) the difference in CDF between conditions, each of which failed to exhibit any significant crossmodal effects on perceptual switches compared to visual only periods (FDR q = .05). Thus, we did not pursue further spectral or neural analyses of these conditions.

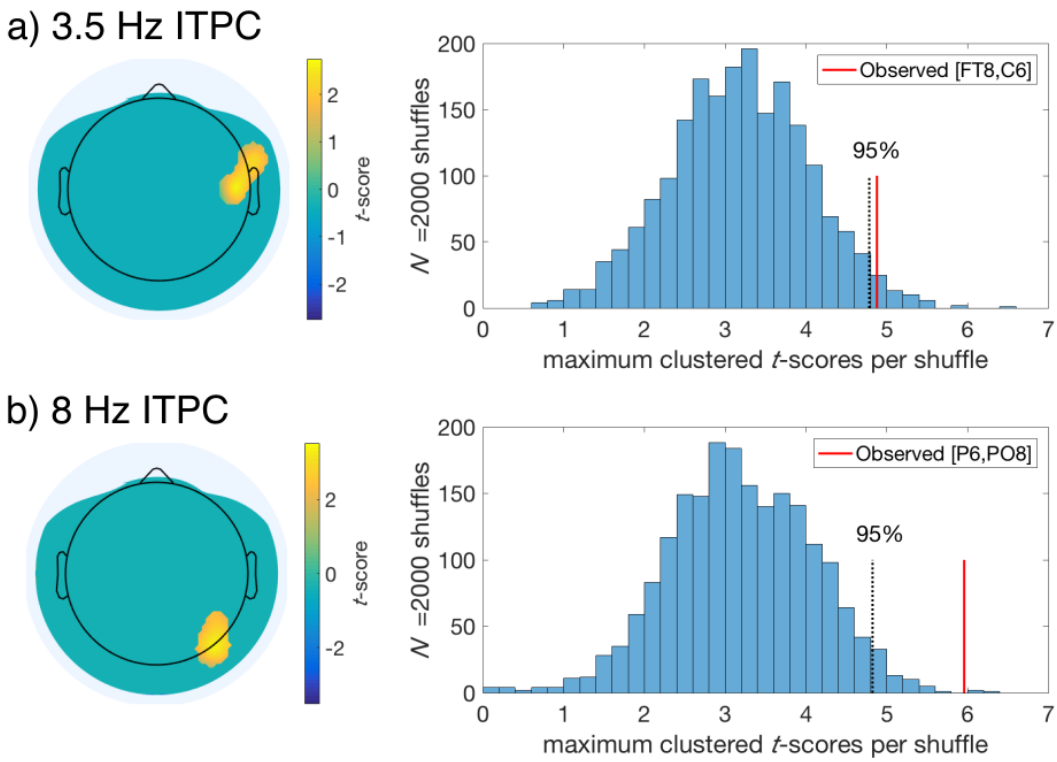

Figure 4.10, related to Figures 4.4 and 4.5.

Displayed are the regions selected for correlation analysis after satisfying our two-stage statistical tests on evoked ITPC, comparing mismatched and matched conditions for a) 3.5 Hz and b) 8 Hz. The right panels display the resulting null distributions obtained after the permutation of condition labels (mismatched vs matched) and performing t-tests across subjects on the mean evoked ITPC for each electrode after down-sampling (see Method). The maximum clustered t-scores per shuffle were retained to create the null distributions. The observed sum of t-scores is displayed as a vertical red line, while the top 95% of the distribution is marked with a vertical dotted back line.

Chapter 5: Measuring graded changes in consciousness through multi-target filling-in

Authors and affiliations: Matthew James Davidson[1, 2] *, Irene Graafsma[3], Jeroen J.A. van Boxtel[1,2,4*,+] , Naotsugu Tsuchiya[1, 2,*,+]

[1)] School of Psychological Sciences, Faculty of Medicine, Nursing, and Health Sciences, Monash University.
[2)] Monash Institute of Cognitive and Clinical Neurosciences, Monash University.
[3)] School of Psychology, The University of Amsterdam
[4)] School of Psychology, The University of Canberra.
* corresponding author, lead contacts.
+ equal contribution

5.1. Article introduction

I conclude the first empirical axis of this thesis, investigating crossmodal contributions to the contents of consciousness during binocular rivalry **(Chapters 3 and 4)**. Before returning to these results in an integrated discussion, I now examine the combinatorial advantage of combining the frequency-tagging approach with an understudied class of perceptual phenomena, known as perceptual filling-in (PFI; **Chapters 5 and 6**).

To briefly review, PFI is a remarkable perceptual phenomenon which occurs when an incomplete region of a visual scene becomes perceptually filled-in by the visual features of the background surrounds. We leveraged the ubiquity of this process by recording four simultaneous behavioural responses to four unique regions undergoing perceptual filling-in. By dynamically flickering the background of the display (rather than just target regions), we were able to record graded steady-state visually evoked potentials (SSVEPs) in the EEG.

Paradigms that capture subtle changes to the contents of consciousness are highly coveted in the fields of perceptual and cognitive neuroscience, to best quantify

the subtle changes that accompany a change in subjective experience. This line of enquiry was also motivated after identifying the utility of filling-in phenomena for application in consciousness research. In particular, the potential to collect behavioural reports for spatially separated regions of a visual scene (see also, **Appendix 1**).

Chapters 5 and 6 contribute to the second empirical question outlined in Chapter 2, on whether distinct neural correlates of attention and consciousness can be captured. As PFI was identified as a useful resource in this aim**, Chapter 5** first to establish that distinct neural correlates of changes in perceptual awareness can be captured. **Chapter 6** then refines this paradigm to provide evidence for a dissociation between attention and consciousness during PFI.

5.2. Abstract

Perceptual filling-in (PFI) occurs when a physically present visual target disappears from conscious perception, with its location filled-in by the surrounding visual background. Compared to other visual illusions, these perceptual changes are crisp and simple, and can occur for multiple spatially separated targets simultaneously. Contrasting neural activity during the presence or absence of PFI may complement other multistable phenomena to reveal the neural correlates of consciousness (NCC). We presented four peripheral targets over a background dynamically flickering at 20 Hz. While participants reported on target disappearances/reappearances via button press/release, we tracked neural activity entrained by the background during PFI using steady-state visually evoked potentials (SSVEPs) recorded in the electroencephalogram. We found background SSVEPs closely correlated with subjective report, and increased with an increasing amount of PFI. Unexpectedly, we found that as the number of filled-in targets increased, the duration of target disappearances also increased, suggesting facilitatory interactions exist between targets in separate visual quadrants. We also found distinct spatiotemporal correlates for the background SSVEP harmonics. Prior to genuine PFI, the response at the second harmonic (40 Hz) increased before the first (20 Hz), which we tentatively link to an attentional effect. There was no difference between harmonics for physically removed stimuli. These results demonstrate that PFI can be used to study multi-object perceptual suppression when frequency-tagging the background of a visual display,

and because there are distinct neural correlates for endogenously and exogenously induced changes in consciousness, that it is ideally suited to study the NCC.

5.3. Significance statement:

Perceptual filling-in (PFI) is a transient illusory disappearance of visual objects from consciousness. By holding the object constant, we can contrast neural activity during periods with and without PFI to isolate the neural correlates of conscious perception. Unlike traditional visual illusions, PFIs are subjectively crisp and simple, and can happen simultaneously at different spatial locations. By frequency-tagging the background display, we demonstrate graded neural correlates for graded changes in consciousness, and provide evidence to differentiate between the perceptual processes evoked during PFI from those evoked by the physical removal of the same peripheral stimuli.

5.4. Introduction

In perceptual filling-in (PFI) phenomena, areas of the visual environment that are physically distinct become interpolated by the visual features of the surrounding texture or background (Komatsu, 2006; Meng, Remus, & Tong, 2005; L Pessoa et al., 1998; Ramachandran & Gregory, 1991; Weil & Rees, 2011). Although PFI neatly displays how our awareness of a visual scene is shaped by (unconscious) inferential processes (Komatsu, 2006), it has traditionally been investigated to understand how our visual system compensates for retinal-blind spots (Durgin, Srimant, & Levi, 1995; Komatsu, Kinoshita, & Murakami, 2000; Ramachandran & Gregory, 1991; Spillmann, Otte, Hamburger, & Magnussen, 2006), and visual field defects (Gassel & Williams, 1963; Gerrits & Timmerman, 1969; Safran & Landis, 1996). Accordingly, a range of low-level visual attributes such as target contrast (Stürzel & Spillmann, 2001) target eccentricity (De Weerd, Desimone, & Ungerleider, 1998), and microsaccades (Martinez-Conde, Macknik, Troncoso, & Dyar, 2006) have been shown to affect the dynamics of PFI. As a result, the neural interpolation of information in lower visual areas has been implicated as one active mechanism behind PFI (De Weerd, Gattass, Desimone, & Ungerleider, 1995; Komatsu, 2006; Meng et al., 2005; L Pessoa et al., 1998).

In addition to the role of low-level visual processes, top-down attention and higher-cortical areas have also been implied to play a role in the initiation, maintenance, and termination of PFI (De Weerd et al., 2006; Weil, Wykes, Carmel, & Rees, 2012). For example, selectively attending to the location of a target (De Weerd et al., 2006), or attending to shared features among peripheral targets (De Weerd et al., 2006; Lou, 1999) has been shown to increase the likelihood of PFI to occur. This poses an intriguing puzzle, as neural responses to a sensory stimulus usually increase when prioritized by top-down attention (Harris & Thiele, 2011; Reynolds & Pasternak, 2000; Spitzer, Desimone, & Moran, 1988), and increase when the stimulus is consciously perceived (e.g. De Weerd et al., 1995; Polonsky, Blake, Braun, & Heeger, 2000, but also see Donner et al., 2008; Logothetis, 1998; Watanabe et al., 2011). As attention during PFI increases disappearance rates, insights into the mechanisms of this phenomenon may contribute to the hotly debated dissociation between attention and consciousness (Koch & Tsuchiya, 2007; Ling & Carrasco, 2006; van Boxtel et al., 2010a, 2010b). We were also motivated to explore whether the use of PFI could capture graded changes to the neural correlates of conscious perception, as PFI can occur over multiple regions embedded in the same visual-background. More specifically, we hypothesized that upon the phenomenological interpolation of target regions, an increased neural response to the surrounding visual background would be recorded, consistent with active mechanism accounts of perceptual suppression during PFI (De Weerd et al., 1995; Komatsu, 2006; Meng et al., 2005; L Pessoa et al., 1998).

We investigated the neural correlates of PFI through the use of frequency-tagging in the EEG, and focused on the background of our visual display in contrast to a previous study that looked only at a single target stimuli (Weil et al., 2007). By presenting flickering visual stimuli, frequency-tagging elicits a steady-state visually evoked potential (SSVEP), which can be analysed as a narrowband change in power at the flicker-frequency of interest (Norcia et al., 2015; Vialatte et al., 2010). This flicker effect is used to 'tag' isolated populations of neurons processing each flickering stimulus (reviewed in Norcia et al., 2015). SSVEPs have been used to track fluctuations in visual awareness between competing stimuli (Brown & Norcia, 1997; Lansing, 1964; Sutoyo & Srinivasan, 2009; Tononi et al., 1998; Zhang et al., 2011) as well as to track the allocation of attention (Andersen et al., 2006; Müller et al., 1998; Müller et al., 1998). The latter effect may be particularly strong in the second

harmonic (i.e. frequency double) of the SSVEP driving frequency (Kim et al., 2007; 2011). To investigate the neural correlates of PFI, we combined the SSVEP technique with a novel multi-target PFI paradigm, by frequency-tagging the *background* of our visual display. This approach allowed us to obtain a more graded response to the amount of change in conscious perception, by investigating how the number of targets perceptually filled-in influenced the neural responses to the shared visual-background display.

5.5. Methods

5.5.1. Participants

Twenty-nine healthy volunteers (11 male, 18-39 years of age, 24 ± 5 years) took part in the study. Participants had normal or corrected-to-normal vision. All participants were recruited via convenience sampling, provided written informed consent prior to participation, and received a monetary compensation (30 AUD) for their time. The study was approved by the Monash University Human Research and Ethics Committee (MUHREC #CLF016).

5.5.2. Apparatus and stimuli

Participants were seated in a dark room approximately 50 cm distance from a computer screen (size 29 x 51 cm, resolution 1080 x 1920 pixels, subtending 32 x 54° visual angle, refresh rate 60 Hz). To frequency-tag the background image, we prepared 100 random patterns prior to the start of each experiment. To construct each background patter, we first down-sampled the screen to 540 x 960 pixels. Then we assigned a random luminance value (drawn from a uniform distribution from black to white) to each down-sampled pixel. These background images were refreshed at a rate of 20 Hz by randomly selecting from the set of 100 prepared patterns.

On top of this background image, the display was composed of a central fixation cross (1.03° visual angle in height and width), surrounded by four counter-phase flickering 2 x 2 checkerboard targets (4.56° visual angle in diameter). A target was located in each quadrant, centred at 13.3° eccentricity from the centre of the

screen. Note that these targets were not located on a 45° diagonal, but on the diagonal axes between corners of our screen (i.e. closer to the horizontal than vertical; horizontally distance from centre 20.2°, vertical distance from centre 11.5°; Figure 5.1). Targets were smoothly alpha-blended into the background texture following a 2D Gaussian profile (*SD* = 1.06° visual angle in diameter). As small, peripherally located targets flickering above 7 Hz are more likely to disappear (Anstis, 1996; Schieting & Spillman, 1987), each target was flickered by reversing the contrast of checkerboard elements at one of four unique frequencies (8, 13, 15 and 18 Hz).

We chose to use small, peripherally located flickering targets to optimally induce disappearances across all our participants. Yet, these same parameters are sub-optimal for frequency-tagging stimuli. (Achieving perfect presentation was also not possible for all targets due to our 60 Hz refresh rate). As a result, our simultaneous background flicker remains as a relatively pure EEG measure for tracking the disappearance of multiple disappearing PFI targets, which is ideal because the measure of background SSVEP is a novel element of our experiment.

5.5.3. Task procedure

Each experimental session was composed of 25 trials, 60 seconds per trial. Between the trials, participants were able to take short self-timed breaks, resulting in a total time-on-task of approximately 30 minutes. Before the experiment, participants were instructed to fixate on the central cross, and were informed that they may sometimes experience a visual illusion where any number of peripheral targets may disappear from their field of vision. We did not monitor eye position. Participants then completed one practice trial to familiarize themselves with the corresponding button presses required for targets in each of the four visual quadrants. Specifically, they were instructed to press keys ‘A’, ‘Z’, ‘K’, and ‘M’ on a traditional QWERTY keyboard, assigning them to the upper left, bottom left, upper right, and bottom right targets, respectively. Participants were instructed to hold each button for the duration of disappearance of the corresponding target, and to release it immediately upon the corresponding reappearance. Figure 5.1 presents the basic configuration of the experimental display used (see Movie 1 for an example of the flickering display including physical removal of the targets, hereafter referred to as catch periods).

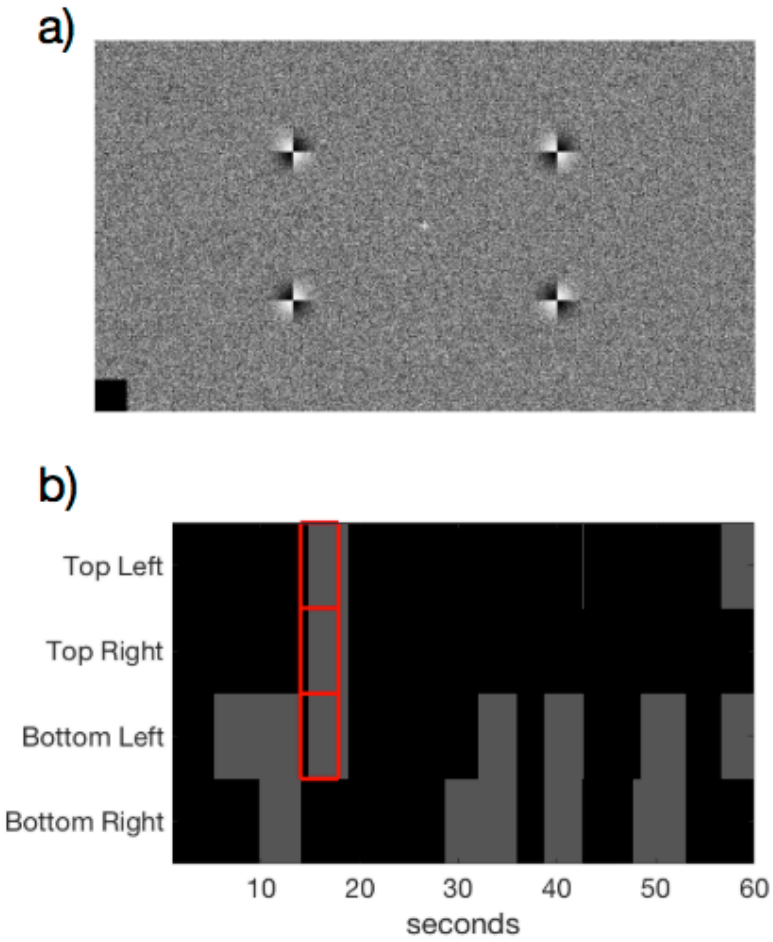

Figure 5.1: Stimulus display and example response.

a) Stimulus display containing a central fixation cross, dynamic background (updated at 20 Hz) and four target checkerboard stimuli, each reversing in luminance polarity at a different frequency (8, 13, 15 or 18 Hz). b) Example time course of behavioural responses over a 60-second trial from one participant. Participants were asked to monitor each peripheral target simultaneously, and to press and hold each button upon perceptual disappearance (PFI events shown in grey) at the corresponding location of the target. Catch periods are shown in red, for which targets were physically replaced by the flickering background texture. Note that targets often disappear and reappear together.

5.5.4. Catch periods

We introduced catch periods to check if participants were correctly reporting on disappearances. During a catch period 1 to 4 targets were physically removed from the display and replaced with the background through alpha blending. Multiple targets during PMD periods were removed with the same onset. Each catch period lasted from 3.5 to 5 seconds in duration (drawn from a uniform distribution). To mimic the

phenomenology of endogenous PFI events, we generated catch periods by linearly ramping the luminance contrast of the target up or down over 1.5 seconds. Participants were not informed of the catch periods.

These physical catch periods also served as a control condition for comparison with the neural signals evoked by PFI. Within 24 trials, catch events in which one, two, three or four targets were removed each occurred on six trials for each participant. The location of the removed targets in the case of one, two and three targets were randomized. The order of these catch events were also randomized for each experiment. A previous study showed that flickering peripheral targets tend not to disappear in the beginning of trials (Schieting & Spillman, 1987), so each catch event began no sooner than 10 seconds after the beginning of each trial to ensure that catch disappearances remained indistinguishable from PFI. Our own data also confirmed that participants reported much lower PFI in the initial 10 seconds of each trial, with PFI plateauing after approximately 10-15 seconds. We also did not include catches within the last 10 seconds. We note that for 10 of our 29 participants, four-target catch periods did not occur due to a coding error, and instead all four targets remained on screen, resulting in catch periods being presented on 92% of trials overall (over all N=29 participants).

5.5.5. Participant and trial exclusion based on catch periods

Initial screening analyses sought to confirm whether participants were able to simultaneously monitor the visibility of multiple peripheral targets using four unique buttons, and perform this task accurately and in compliance with instructions. Due to a keyboard malfunction, button press responses to three and four disappearing targets became indistinguishable in our post-hoc analysis, and have been analysed together henceforth as "3 or 4 buttons pressed". In the subsequent analyses where the number of buttons pressed mattered, we proceeded as if three buttons were pressed in these periods.

We analysed button press responses during catch periods to estimate participant attention on task. As catch periods were embedded within a trial, some catch periods occurred when participants had already pressed buttons. Such events are more frequent for those who report more frequent PFI. To estimate this baseline

button press rate per individual participant, we performed a bootstrapping analysis with replacement. For a given catch onset in trial T at time S (seconds), we randomly selected a trial T' (T=T' was allowed) and epoched the button press time course over the period of [S-2, S+4] at corresponding catch target locations in trial T'. We repeated this for all trials (T=1...24, except for the 4-catch error mentioned above) to obtain a single bootstrapped set of trials per participant. We then obtained the mean button-press time course across button-locations from each of the 200 bootstrap sets to obtain a null distribution of the shuffled button-press time course. We also obtained the mean button-press time course for observed data across button-locations, excluding catch periods when four targets were removed.

As the distribution at each time point for both observed and shuffled data was not normally distributed, we first converted the data into z-scores using the logit transformation before calculating the confidence intervals (CI). Then, we used mean z-scores (±1.96 standard deviation of z-scores) as the CI for the null distribution of shuffled data within each participant, and observed data across participants.

We excluded 3 participants whose mean button-press time course around the actual catch onset failed to exceed the CI within the first two seconds (i.e., [0, S+2]). We defined the catch-onset reaction time as the first time point after which the mean button-press data exceeded the top CI, indicating successful button presses for catch targets. Figure 5.2a shows the catch response for an example participant retained for analysis. Four further participants were removed from subsequent analyses for failing to experience PFI during most of the experimental session (i.e., only brief events on 1 or 2 trials). For the remaining participants, the mean reaction time to respond to catch onsets, and thus the disappearance of a peripheral target was 0.92 seconds (*SD* = 0.046). Figure 5.2b shows the proportion of button press responses for all catch events across participants retained for analysis (*N*=22).

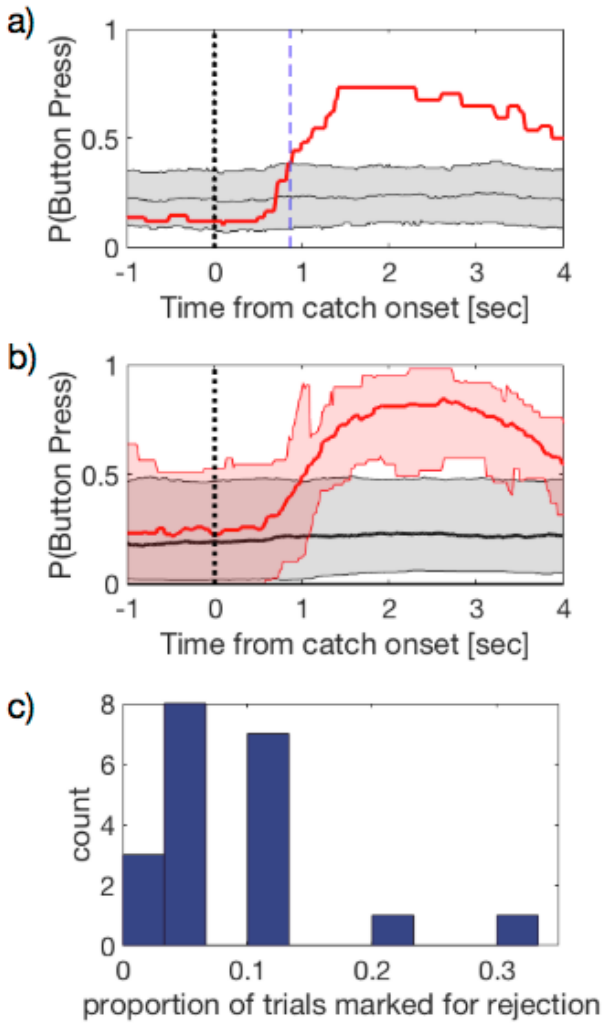

Figure 5.2. Catch period analysis and trial rejection following the physical removal of flickering targets at catch onset.

a) Example catch response for a single participant. The first time point that the observed likelihood of button press (red) exceeded the bootstrapped CI (grey) corresponds to the catch reaction time (0.87 sec for this participant, marked with a vertical dashed blue line). b) The mean time course for the likelihood of button press and its bootstrapped sets across participants, shown in red and grey respectively. Shading represent the CI (computed with logit transform and presented after reverse transform) across participants. c) Participant-level histogram of the proportion of trials rejected, based on period-by-period catch analysis.

Having identified which participants could successfully indicate target disappearance based on their button press data, we continued to identify and remove any trials from the subsequent analysis in which a catch was not correctly detected. We undertook this procedure to assure that in all retained trials participants paid proper attention on task and reported accurately on PFI . We regarded a catch period as being successfully identified if participants pressed the corresponding button for at

least 50% of the allowed response time window. For multi-target catch periods, we applied the same criteria for each button separately. If any button was not pressed at least 50% of the time, the catch was considered undetected. For four-target catch periods, we analysed it as if it was a three-target catch period. This window was from the onset of the catch plus 1 second (in consideration of the reaction time delay) to the end of catch. For example, if the catch period under consideration was 3.5 seconds in duration, we defined the allowed time window to be [1, 3.5] seconds from the catch onset. Figure 5.2c shows a participant-level histogram for the number of rejected trials ($M \pm SD$: 1.75 ± 1.48 trials or 8.96 ± 7.89% of all trials). After participant and trial exclusion, we continued by examining the behavioural dynamics of PFI.

5.5.6. Quantifying PFI and location-shuffling analysis

We analysed the number of PFI events, duration of each PFI event, and total duration of PFI per 60 second trial. Although these variables may be correlated, they have also been shown to reveal complementary data features in similar multi-target designs (e.g. (Bonneh et al., 2013; McEwen, Paton, Tsuchiya, & van Boxtel, 2018; Thomas, Davidson, Zakavi, Tsuchiya, & van Boxtel, 2017). Each of these variables were compared based on the number of PFI (nPFI; 1, 2, 3 or 4), quantified by the number of simultaneous button-press responses at each time point. We note that our reported analysis did not exclude transient button-press periods (<200 ms), as the dynamics of multi-target PFI are currently unknown. Instead, we performed a reconstruction analysis to estimate the contribution of overlapping and closely spaced PFI events on EEG responses (see ***Reconstruction analysis - Methods).*** However, when we repeated our behavioural analysis by removing these instances (< 200 ms), the results were not significantly different

To investigate whether the simultaneous multi-target PFI observed in participant data (e.g. Figure 5.1b) exceeded that to be expected by chance, we performed a shuffling analysis to create a null distribution. Specifically, we created 1000 shuffled trials for each participant, by randomly selecting the button press time course for each of the four target locations independently from any of the trials throughout their experimental session (this could include multiple locations within the same trial). As such, newly created shuffled trials allowed us to compare the effect of

multiple disappearing target events within the same trial (the observed data) to the shuffled data without the presence of a temporal correlation between target locations. If target disappearances during PFI were independent, then shuffled and experimental data should be similar. The comparison between the observed and the shuffled data is displayed in Figure 5.7.

5.5.7. Linear-mixed effect analysis – Behaviour

All statistical analyses were performed using Matlab (ver: R2016b) and jamovi (ver 0.9). We used linear-mixed effect (LME) analysis to examine whether various PFI characteristics (e.g., durations) were affected by the number of simultaneously invisible targets (nPFI; n = 0, 1, 2, 3 or 4), including intercepts for participants as a random effect. We performed likelihood ratio tests between the full model and a restricted model which excluded the factor of interest (Glover & Dixon, 2004; Pinheiro, Bates, DebRoy, & Sarkar, 2014; Winter, 2013).

We also performed LME analyses to compare the slopes of observed and shuffled data, when considering the effect of the number of simultaneously invisible targets on PFI characteristics. For this analysis, we fit a linear model (1st order polynomial) to the observed data across participants (N = 22), and retained the slope (β) as our observed test statistic. Similarly, we also fit the same linear model to each of n = 1000 sets of shuffled data, each of which was computed from the shuffled trials across N = 22 participants. We shuffled the trials within each participant within each set and again retained the β values. Then, we compared the observed β value with the null distribution of the β values from n = 1000 shuffled sets. If the observed β exceeded the top 97.5% or was lower than 2.5% of the null distribution, we considered the observed effect to be significant at $p < .05$.

5.5.8. EEG acquisition and pre-processing

Throughout each session whole-head EEG was recorded with 64 active electrodes arranged across an elastic-cap (Brain Products, ActiCap) according to the international 10-10 system. Electrode impedances were kept below 10 kΩ prior to experimentation, and recorded using the default reference (FCz) and ground electrode (AFz) via Brainvision recorder software (sampling rate = 1000 Hz, offline bandpass of 0.5-70 Hz). All EEG data was stored for offline analysis using custom MATLAB

scripts (Ver: R2016b), as well as the EEGLAB (Delorme & Makeig, 2004) and Chronux (Bokil et al., 2010) toolboxes. All EEG channels were first re-referenced to the average of all electrodes at each sample and down-sampled to 250 Hz. We further applied a Laplacian transform to improve spatial selectivity of the data, which is known to contribute minimal contamination to the SSVEP when using rhythmic-entrainment source separation (RESS; Cohen & Gulbinaite, 2017), which we used to extract SSVEP responses as detailed below.

5.5.9. SSVEP Signal-to-Noise Ratio (SNR) calculation

To estimate the topography and across-channel correlation of SSVEPs (Figures 5.5 and 5.11), we first calculated the natural log of the power spectrum via the fast Fourier transform (FFT) over the period -3000 to -100 ms before, and 100 to 3000 ms after button press. In the SSVEP paradigm, we operationally regard power at the tagged frequency as signal and power at non-tagged neighbouring frequencies as noise (Norcia et al., 2015) and compute the signal-to-noise ratio (SNR) at each frequency. In logarithmic scale, this corresponds to log of the power at each frequency subtracted by the mean log power across the neighbourhood frequencies. In this paper, all SNR results are based on this log-transformed SNR metric because without log-transform, SNR is highly skewed and not appropriate for various statistical tests. Over the 2.9 s time window (half-bandwidth = 0.35Hz), we computed the SNR at frequency f (Hz) as the mean log power over the neighbourhood frequencies for f subtracted from the log power at f. The neighbourhood is defined as [f-1.22, f-0.44] Hz and [f+0.44, f+1.22] Hz. In addition, we also computed the time-course of the SNR over a 1 second window (half-bandwidth = 1 Hz) with a step-size of 0.15 second, to enable the comparison of fluctuations in SNR over time. For this shorter time window, we used the neighbourhood as [f-3.92, f-1.95] Hz and [f+1.95, f+3.92] Hz to compute the log(SNR) time course.

5.5.10. SSVEP analysis via rhythmic entrainment source separation (RESS).

After examining the topography of log(SNR) responses, we applied rhythmic entrainment source separation (RESS) to optimally extract the time-course of frequency-tagged components of SSVEPs without relying upon electrode channel

selection (Cohen & Gulbinaite, 2017). In standard SSVEP analysis, the SNR of the target frequency is examined by averaging across the electrodes within a region of interest or selecting one electrode in a certain way (e.g., prior hypothesis, anatomical localization or separate datasets). An alternative to this classic approach is RESS, which creates a map of spatial weights across all electrodes which optimize the SNR at a particular frequency, tailored for each participant. Specifically, RESS functions by creating linear spatial filters to maximally differentiate the covariance between a signal flicker frequency and neighbourhood frequencies, thereby increasing the signal-to-noise ratio at the flicker frequency. After obtaining signal and neighbourhood covariance matrices, the eigenvector with the largest eigenvalue is used as channel weights to reduce the dimensionality of multi-channel data into a single component time course, which reduces multiple comparisons across channels in statistical testing.

After epoching all data using the time-windows -3000 to -100 ms and 100ms to 3000 ms peri button press/release, we then constructed RESS spatial filters per participant, avoiding catch periods. We constructed RESS spatial filters from 64-channel EEG, by extracting signal data following a narrow-band filter via frequency-domain Gaussian, centred at flicker frequencies (20 and 40 Hz, full-width at half maximum = 1 Hz). We proceeded by selecting broadband (unfiltered) neural activity to construct reference covariance matrices. We selected broadband EEG as our references after confirming that improvements to the SNR at both the 20 and 40 Hz signal were statistically equivalent. Comparing signal to broadband activity has previously been shown to allow the reconstruction of SSVEP signals using RESS (Cohen & Gulbinaite, 2017).

Critically, we performed the above procedure without distinguishing whether targets were disappearing or reappearing due to button press or release in order to reduce the possibility of overfitting. If we were to construct separate filters for periods around the time of target disappearance and reappearance, then any differences between these conditions could be due to differences in the obtained filters, or overfitting of the filters prior to our condition comparisons. After application of the RESS spatial filters, we reconstructed the time course of SSVEP log(SNR) from the RESS component time courses, separately for each flicker of interest as described above. The main results of our analysis also hold with more conventional SSVEP analysis, such as when focusing on only parieto-occipital electrodes.

5.5.11. SNR time-course data cleaning

Preliminary analyses revealed a sharp and consistent decrease in 40 Hz log(SNR) amplitude which was time-locked to the beginning of each catch period. Subsequent inspection of recorded screen flip-times revealed a lag in background stimulus presentation (16.7-33.3 ms duration) at catch onset, which resulted in the background pixels for one presentation frame being skipped. This caused an artefact in the spectrogram where the time window of the analysis included the problematic period. To correct for this artefact conservatively, we interpolated the 40 Hz SNR time-course from -500 to 500 ms around physical catch onset.

5.5.12. Event-by-event image analysis of button press and SSVEP-SNR

Due to variations in the frequency and duration of PFI per participant, averaging data over participants is not straightforward. To resolve this, we performed image-based event-by-event analyses (Fujiwara et al., 2017) to investigate whether the amount of PFI reported may reflect changes in log(SNR). Within each participant, all PFI events were sorted in descending order based on the sum of buttons pressed at each time point, and over a 3 second time window (detailed below) per disappearance/reappearance event. For this analysis, we counted three button presses as 3 even though participants might have tried to press 4 buttons. For PFI disappearances and reappearances, we averaged this over [0, +3] seconds and [-3, 0] seconds with respect to the button press or release, respectively. We call this sum of the number of buttons pressed over these time periods "the amount of PFI". We then resampled along the trial dimension to 100 samples to map from 0 to 1 (normalized event count) for each participant. Participant data was then smoothed along the normalized trial dimension and averaged across participants, to visualize the time-course of SNR as a function of normalized PFI. This resampling, smoothing and averaging process performed on button-press responses was repeated for the event-by-event time course of log(SNR), with the order of events predetermined by the corresponding button-press responses per participant. A schematic pipeline for this entire procedure is displayed in Figure 5.3.

To quantify the relationship between log(SNR) and the amount of PFI, we grouped events when the amount of PFI was between 0 and 1, 1 and 2, or greater than

2. A median split based on the amount of PFI resulted in similar data and subsequent conclusions.

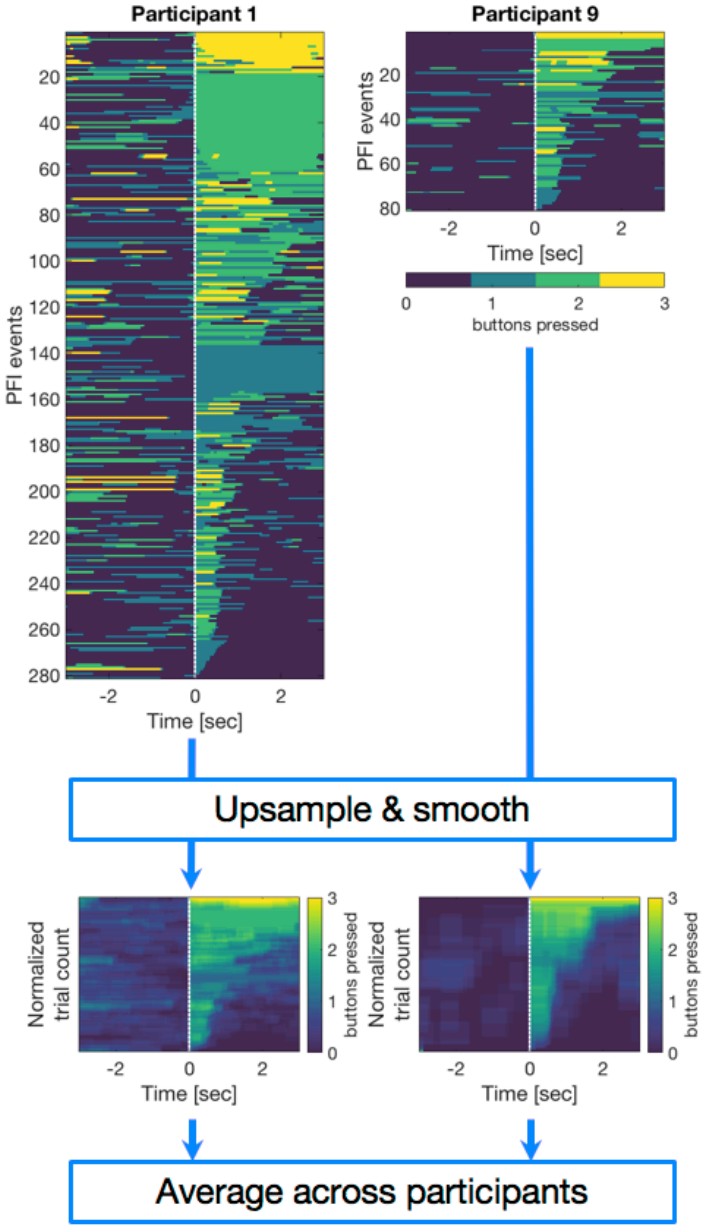

Figure 5.3. Pre-processing for event-by-event based image analyses.

PFI events were first sorted according to "the amount of PFI" (the sum of buttons pressed over 3 seconds) occurring after a button-press, or before a button-release event. Each image along the y-axis was then resampled to normalize the trial number into arbitrary units of 100 samples. A 15-sample moving average was then applied to smooth each image along the normalized event-dimension, before averaging across participants. The same process was also applied to RESS log(SNR) after sorting by the amount of PFI per event based on button-press (or -release) events. This image-based analysis enables us to compare PFI dynamics despite differences in the number of PFI events per participant.

5.5.13. Reconstruction analysis to compare the impact of multiple-target disappearances and reappearances on SNR, during PFI and catch periods

Due to our novel task design which employs multiple targets, it is necessary to account for whether the temporal dynamics of log(SNR) during PFI or catch periods differ due to the involvement of unique mechanisms, or are attributable to differences in the requirements to report on overlapping events. For example, in our data the increase of log(SNR) at PFI onset appears to be transient, while decreases upon PFI offset are more sustained (Figure 8). Could these differences be due to the overlapping influence of temporally proximal PFI events? This is particularly important as the temporal profile of PFI and catch periods may differ based on the way we programmed catch periods; PFI events can accumulate for multiple targets in close temporal proximity, yet multiple catch periods share a common temporal onset, even when spatially distributed. We approached this problem by performing a SNR reconstruction procedure, to model and compare the expected temporal dynamics of log(SNR) during PFI and catch, when accounting for accumulating disappearance and reappearance events. This analysis progressed through three steps (Figure 5.4)

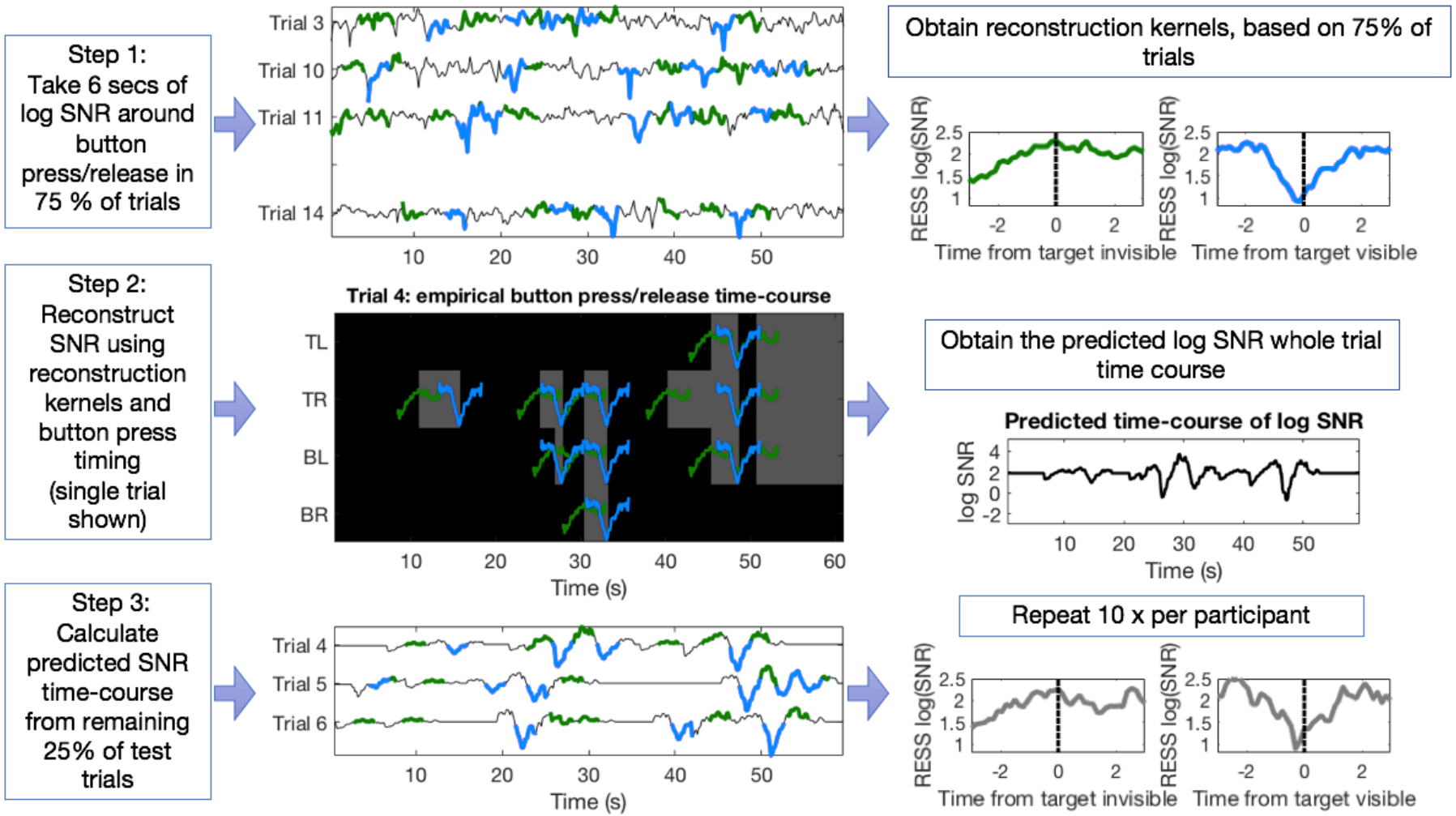
Step 1:
Take 6 secs of log SNR around button press/release in 75 % of trials
Trial 3
Trial 10
Trial 11
Trial 14
10
20
30
40
50
Obtain reconstruction kernels, based on 75% of trials
RESS log(SNR)
2.5
2
1.5
1
-2
0
2
Time from target invisible
Time from target visible
Step 2:
Reconstruct SNR using reconstruction kernels and button press timing
(single trial shown)
Trial 4: empirical button press/release time-course
TL
TR
BL
BR
10
20
30
40
50
60
Time (s)
Obtain the predicted log SNR whole trial time course
Predicted time-course of log SNR
log SNR
4
2
0
-2
10
20
30
40
50
Time (s)
Step 3:
Calculate predicted SNR time-course from remaining 25% of test trials
Trial 4
Trial 5
Trial 6
10
20
30
40
50
Time (s)
Repeat 10 x per participant
RESS log(SNR)
Time from target invisible
Time from target visible

Figure 5.4. Pipeline for SNR reconstruction analysis to estimate the impact of accumulated PFI disappearances/reappearances on the observed time course of log(SNR).

Step 1: we first calculated the reconstruction kernels in response to target disappearance and reappearance events from 75% of training trials per participant. Log(SNR) around button press/release events (epoched -3 to +3 seconds) is shown in green/blue, respectively. Reconstruction kernels are computed as the mean log(SNR) time course around button press/release events (over 18 trials, for this participant who had no rejected trials). Step 2: to predict the time course of log(SNR), we convolved the reconstruction kernels from Step 1 with recorded time of button press and release events in the remaining test trials (here only displaying 1 trial for demonstration purposes). As multiple PFI disappearances and reappearances can happen across target locations in close temporal proximity (< 1 second), this analysis enabled an estimation of the impact that consecutive PFI events have on SNR time course. The predicted time courses (grey) are computed as the mean log(SNR) during PFI events for test trials (over 6 trials for this participant). The predicted time courses are compared with the observed time courses from the same test trials (6 trials). This entire procedure was repeated 10 times per participant to obtain the mean predicted and observed time course for correlation analysis.

First, we calculated the mean log(SNR) time-course for PFI disappearances and reappearances using 75% of trials. Within these trials, we stepped through each time-point in the accumulative button-press responses (0-3 buttons pressed), and epoched the log(SNR) time-course from -3 to +3 seconds around the time of PFI events, which we defined as any change in button-press state (6-second epoch). For this analysis, we did not distinguish the number of disappearing targets at each time-point, just the direction of change (disappearing or reappearing), and obtained the mean disappearance/reappearance time courses which we subsequently used as reconstruction kernels. Second, using these reconstruction kernels, we then predicted the SNR in the remaining 25% of 60-second test trials. We did this by assuming linearity and time invariance in PFI responses, and predicted the 60-second whole-trial SNR time course by convolving the 6-second reconstruction kernels with the actual button-press or -release event times in the test trials. Outside of button press periods, we set the default SNR value as the baseline SNR value from the same trial (e.g. log(SNR) = 2.1 above). Third, from the reconstructed 60-second time course of SNR, we epoched from -3 to +3 seconds around the PFI events and obtained the mean predicted log(SNR) time course. Figure 5.4b shows this procedure for one 60-second trial. We reconstructed a mean predicted SNR from across test trials, separately for

PFI disappearance and reappearance. We repeated this reconstruction 10 times to obtain the mean predicted SNR per participant, which we then averaged across participants. We compared this measure to the observed mean log(SNR) time course from the same test trials.

We repeated the same procedure to compare the predicted SNR from PFI reconstruction kernels to the observed SNR during catch periods. This was necessary due to the embedding of catch periods within multi-target PFI, as catch periods would often overlap with ongoing button-press and -release events signifying genuine PFI. We were then able to statistically determine whether the SNR time courses during subjective and physical target disappearances/reappearances were statistically distinct, by convolving the reconstruction kernels based on (training) genuine PFI with the button-press or -release event times of (test) genuine PFI and catch periods.

To compare the predicted and the observed SNR time course, we evaluated the degree of correlation between them over the 6 seconds surrounding button-press and release, obtaining R^2 for each individual participant. For the statistical analysis, we used repeated measures two-way ANOVA, testing the main effects of background harmonics (1f = 20Hz vs 2f = 40Hz) and the nature of disappearance/reappearance (PFI vs catch) on the R^2 between the observed and the predicted SNR time course.

5.5.14. Cross-point analysis

Past research on binocular rivalry has indicated that perceptual alternations between frequency-tagged stimuli are captured in the time course of SNR, and that the time point shortly after when two SNR time courses crossover concurs with button presses to indicate a change in perception (Brown & Norcia, 1997; Jamison, Roy, He, Engel, & He, 2015; Tononi & Edelman, 1998; Zhang et al., 2011). We were interested to see whether changes in SNR could also predict button presses/releases in our multi-target PFI paradigm. At the participant level, we compared the SNR time course around the time of disappearances to those of reappearances using a paired-samples *t*-tests at each time point. Clusters of significant time points were identified which satisfied $p < .05$ (uncorrected) over a minimum of 300ms, a time window which corresponds to two adjacent time points in our moving-window SNR. Per participant, the first time point in these clusters, which occurred after the time point where the two time courses crossed each other, was taken as the earliest time point at

which the SNR differentiates between target disappearance and reappearance. We also performed the same analysis to compare the time course of SNR during physical target disappearance and reappearance due to catch periods.

5.5.15. Spatial correlation analysis

To perform the spatial correlation analysis, we calculated the time-course of a 64-channel correlation between 1f and 2f log(SNR). Due to differences in the number of PFI events and catch periods, we down-sampled (with replacement) the number of PFI events to 24, which was the maximum number of available catch periods. We then calculated the correlation for this subset of trials, and repeated this analysis 100 times to obtain a distribution of down-sampled correlation values. The mean correlation value from this down-sampled distribution was then used to compare the spatial correlation of PFI and catch periods.

5.5.16. Statistical analysis – EEG

To assess the significance of SSVEP peaks in the EEG spectra, we corrected for multiple comparisons with a False Discovery Rate (FDR) of .05 (Benjamini et al., 2006; Benjamini & Yekutieli, 2001). For corrections of multiple comparisons on the time courses, we used temporal cluster-based corrections (Davidson, Alais, van Boxtel, & Tsuchiya, 2018; Maris & Oostenveld, 2007). For this analysis, the sum of observed test-statistics (e.g., *t* scores) in a temporally contiguous cluster were retained for comparison with a permutation-based null distribution. Specifically, first, we detected any temporally contiguous cluster by defining a significant time point as $p <$ .05 uncorrected (Maris & Oostenveld, 2007). Then, we concatenated the contiguous temporal time points with $p < .05$ and obtained a summed cluster-level test statistic for the cluster. Second, we repeated this procedure after shuffling the subject specific averages within each participant 2000 times. From each of the 2000 shuffled data, we obtained the summed cluster-level test statistics at contiguous temporal time points with $p < .05$ uncorrected, which served as a null distribution. We regarded the original observed effect to be significant if the original summed cluster-level statistics exceeded the top 97.5% of the null distribution of the summed statistics (as $p_{cluster} <$ .025).

5.6. Results

5.6.1. Overview:

Our presentation of the results will be structured as follows. First, we confirmed that our overall SSVEP frequency-tagging was successful (Figure 5.5). Second, we checked if the behavioural reports during catch periods were correlated with neural activity (RESS log(SNR), Figure 5.6). Third, we investigated the behavioural reports during genuine PFI events, and focused on whether or not spatially separated PFI targets interact across visual quadrants (Figure 5.7). Fourth, we then focused on RESS log(SNR) during PFI events, testing if the amount of PFI correlated with the strength of frequency-tagged EEG activity induced by our flickering background (Figure 5.8). Fifth, we devised a SNR reconstruction analysis to estimate the influence of multiple PFI events in close temporal proximity on the RESS log(SNR) (Figure 5.9). Sixth and finally, we also found unexpected temporal (Figure 5.10) and spatial (Figure 5.11) differences between PFI events and catch periods, with respect to the first (1f) and second harmonic (2f) responses (log(SNR)) to background flicker, which we interpret in our Discussion.

5.6.2. Successful frequency-tagging of dynamic background in PFI display:

We first investigated the log(SNR) of target (8, 13, 15, 18 Hz) and background (20 Hz) flicker frequencies and their harmonics. Using a short window (2.9 second duration, see Methods), we found strong and occipitally localized responses to background flicker, but no clear responses to target flicker. To increase the chance of finding target entrainment in the EEG signal, we also analysed the data with the longest time window (60 second for one trial, including catch periods) with the highest frequency resolution. Still we did not detect reliable target-related SSVEPs (Figure 5.5a), owing to their small size and eccentricity.

While the 1f (20 Hz) and 2f (40 Hz) frequency-tagged responses to our background display were strongest at POz, the spatial topographies differed between 1f and 2f (Figure 5.5b). The 1f response was localized to midline occipital electrodes,

while the 2f response extended beyond these regions to include lateral parieto-occipital and parietal electrodes. We continued to analyse the time-course of log(SNR) for background-related 1f and 2f responses after applying rhythmic entrainment source separation (RESS; Cohen & Gulbinaite, 2017), to optimally extract the SNR per participant given these differences in source topography and to avoid multiple comparisons across electrodes (see Methods). From here, all SNR values we present are the RESS log(SNR) (except for the spatial correlations presented in Figure 5.11).

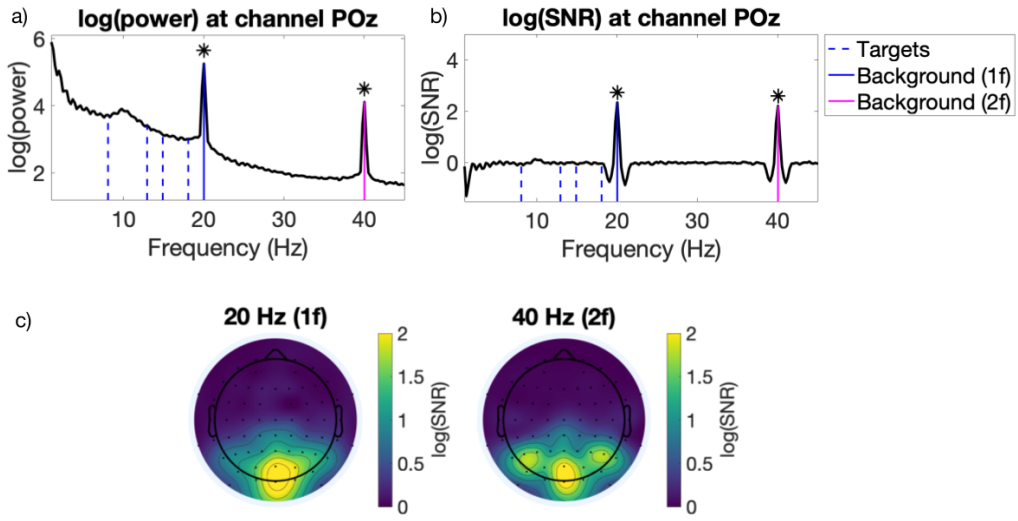

Figure 5.5. Average SSVEP responses in our paradigm.

a) The mean log(power) spectrum, and b) mean log(SNR) over all participants and periods of PFI at channel POz. Asterisks mark log(SNR) significantly different from 0, FDR-adjusted across all frequencies to p < .05. c) Topoplots for the mean log(SNR) at 1f (stimulus flicker)) and 2f (stimulus harmonic) of background-related SSVEPs. The mean is taken across participants over all epochs, excluding catch periods.

5.6.3. Frequency-tagging during catch periods

Having identified the successful entrainment of background responses (Figure 5.5), we analysed the time course of changes to the RESS log(SNR), focusing on 1f and 2f during catch periods. As SSVEPs tend to be weak for peripherally presented stimuli (Norcia et al., 2015), we checked if the physical removal of targets was strong

enough to alter the time course of the RESS log(SNR). During catch periods, we compared the mean RESS log(SNR) during -2 to -0.1 to +0.1 to +2 seconds (two-tailed paired samples *t*-tests). The SNR to background flicker increased upon target removal (1f, $t(21) = 3.80$, $p = .0011$; 2f, $t(21) = 2.21$, $p = .04$). The background SNR also decreased upon target return (1f, $t(21) = -3.51$, $p = .0021$; 2f, $t(21) = -3.50$, $p = .0021$). The increase/decrease of the RESS log(SNR) started upon button press/release, which we return to and investigate in our SNR-reconstruction analysis (Figures 5.4 and 5.9). These results are consistent with an interpretation that the background 1f and 2f SNR increases when peripheral regions are physically interpolated by the flickering background display.

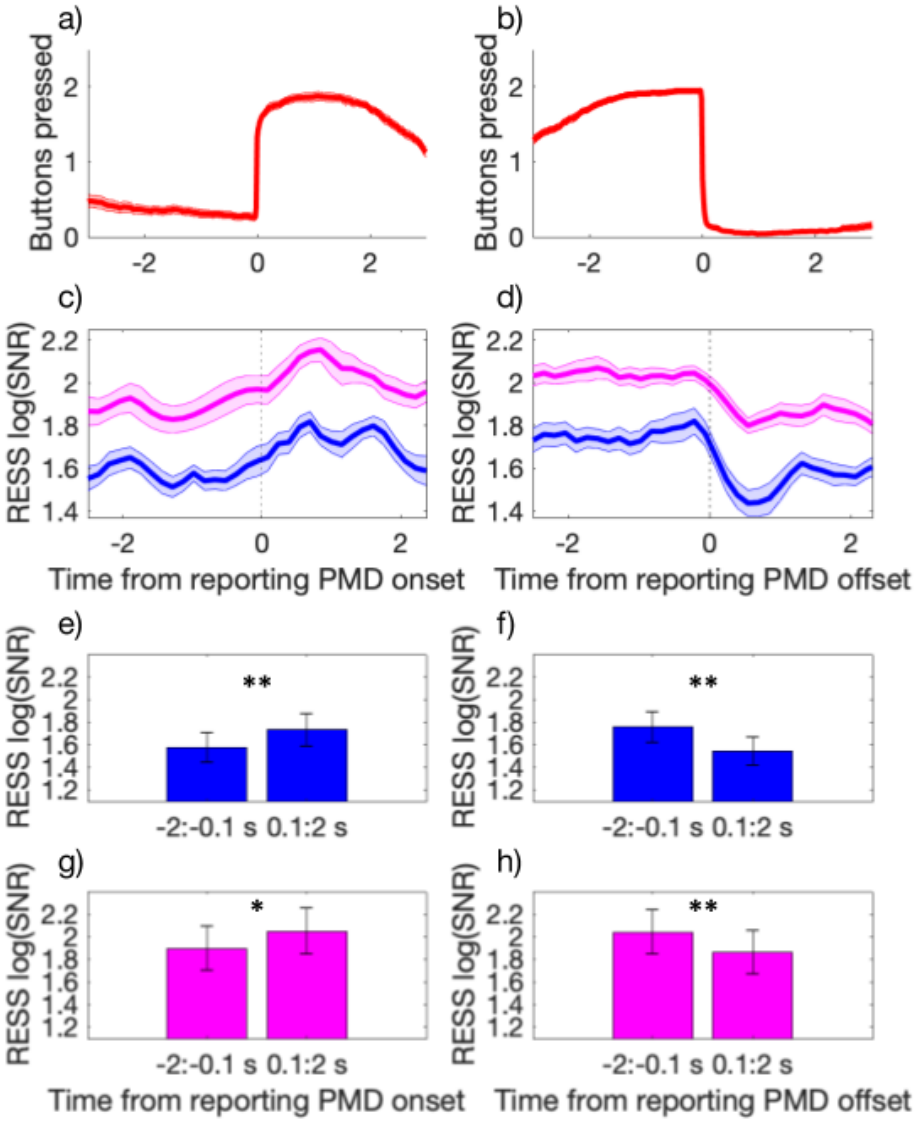

Figure 5.6. Button press time course and background RESS log(SNR) around catch periods.

a-b) mean (± 1 SEM) button-press time course across participants when responding to the physical removal of targets near the onset (a) and the offset (b) of catch periods. c-d) RESS log(SNR) for background SSVEP at 1f (20 Hz; blue) and 2f (40 Hz; magenta). Shading represents ± 1 SEM corrected for within participant comparisons (Cousineau, 2005). e-h) Bar-charts for the statistical comparisons reported in text, comparing RESS log(SNR) before and after button press during PMD periods.

5.6.4. Synergistic effect of multi-target PFI

Next, we turn to the behavioural analysis of the genuine PFI events before interpreting the EEG effects. Specifically, we investigated whether our unique multi-target design had captured an interaction between the four simultaneously presented peripheral targets. Previous research has suggested that neighbouring targets within a single visual quadrant may disappear together (De Weerd et al., 1998). Our design allowed us to examine whether much more widely distributed peripheral targets also interact. Such an interaction would be non-trivial if occurring across all four quadrants of the visual periphery, as it could imply the grouping of targets for PFI despite their disparate retinotopic locations. This would imply the involvement of potentially high-level neural mechanisms that have access to these long-range relations (Wagemans et al., 2012).

First, we analysed whether the number of targets simultaneously invisible were related to 1) the number of PFI events per trial, 2) the average duration of PFI invisibility per event, and 3) the total duration of PFI per trial (Figure 5.7, blue bars). In theory these three variables can vary independently, and in practice they can dissociate (Bonneh et al., 2013; McEwen et al., 2018; Thomas et al., 2017). While periods when all targets were visible had the longest average duration and total duration (i.e., the number of invisible targets = 0), the more interesting trends were found as the number of invisible targets increased. While simultaneous disappearances of 3 or 4 targets were rare (only 2.9 events per trial; Figure 5.7a), when they happened, the event tended to be sustained for a long duration (2.2 sec, Figure 5.7b). As a result, the total duration of 3 or 4 target invisibility (7.6 sec per trial, Figure 5.7c) is comparable to that of 2 target invisibility and longer than that of 1 target invisibility, which happened at the highest rate (8.8 events per trial, 5.8 seconds in total per trial). We formally tested this linear trend by LME analysis and likelihood ratio tests (see Methods). The number of invisible targets (nPFI; 1, 2, 3 or 4: removing 0) significantly affected 1) the number of PFI events per trial ($\chi^2(2) = 47.83$, $p = 4.1\text{x}10^{-11}$), 2) the average duration of PFI per event ($\chi^2(2) = 23.59$, $p = 7.53\text{x}10^{-6}$) and 3) total PFI duration per trial ($\chi^2(2) = 7.27$, $p = .026$).

These significant trends imply that interactions among distant targets occur in a synergistic way, and that when one target is invisible it is often accompanied by other invisible targets. To directly test if this is the case, or if these trends occur by

chance, we employed a shuffling analysis (see Methods). For this, we first sub-selected the button press time course for each location from any four trials (with replacement) and re-computed the behavioural analysis per participant. We repeated this shuffling procedure 1000 times, and from each shuffled dataset we retained the mean PFI data across participants. As the location of each button press in shuffled data could come from any independent trial (e.g. top left = trial 1, top right = trial 23, bottom left = trial 18, bottom right = trial 18), this shuffling procedure conserved the mean number of PFI events overall, while estimating the level of simultaneous invisibility between multiple PFI targets that occurs by chance, when locations are independent.

In the shuffled data, the number of PFI events per trial decreased as the number of invisible targets (nPFI) increased, which is similar to what we observed in the empirical data (11, 7, and 4 events per trial for 1, 2, and 3 or 4 target invisibility; Figure 5.7a, grey bars). However, the trend for shuffled data was quite different from the empirical data for the average durations per PFI event, which were roughly equal across nPFI in shuffled data (2, 1.8, and 1.8 seconds, respectively; Figure 5.7b), and the total duration of PFI per trial, which decreased as a function of the number of invisible targets (16, 10, and 4 seconds, respectively; Figure 5.7c) .

To statistically evaluate these trends between the observed and the shuffled data, we compared the slopes of the linear fit (LME, with random intercepts for each subject) for each of the three PFI variables as a function of the number of invisible targets (nPFI; 1, 2, 3 or 4: removing 0). For all variables, the observed slope was outside the top 97.5% of the slopes in the shuffled data (corresponding to two-tailed $p < .05$, Figure 5.7d-f). Notably, Figure 5.7e and f establish that the observed positive slope for observed data in Figure 5.7b and 7c are contrary to the expected negative slope in shuffled data. In other words, if there are no spatial interactions between distant targets, as in our shuffled data, then we should expect the simultaneous invisibility of 3 or 4 targets to be highly unlikely, and sustained for a shorter duration. By contrast, the observed data show that as more targets are involved with a disappearance event, the longer the disappearances are sustained, strongly suggesting a facilitative interaction between invisible peripheral targets. We return to this synergistic effect of multi-target PFI in our Discussion.

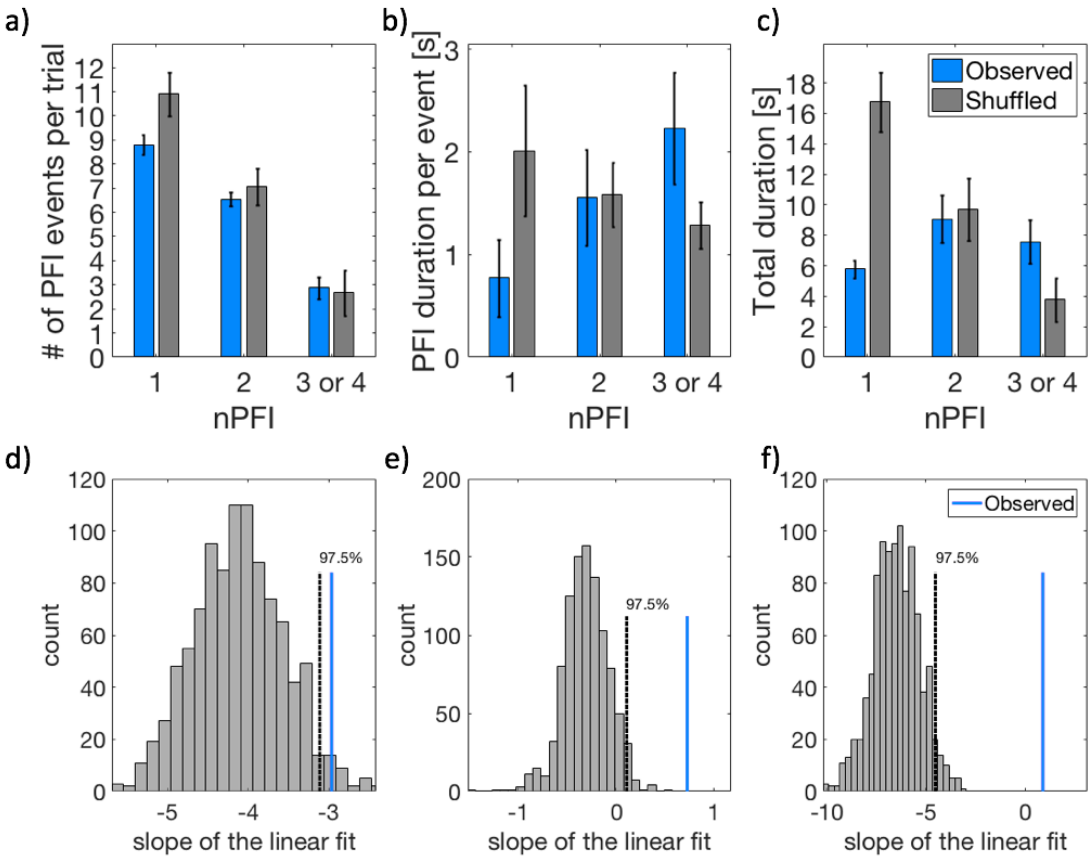

Figure 5.7. Behavioural data.

a) The number of PFI events per trial, b) the mean duration per PFI event, and c) total duration of PFI per trial, as a function of the number of invisible targets (nPFI). All panels display both observed (blue) and shuffled (grey) data. For the observed data, error bars represent 1 SEM, corrected for within-participant comparisons (Cousineau, 2005). For the shuffled data, we first computed the SEM within each shuffled data set across participants. Then, as the error bar, we show the mean of the SEM across 1000 shuffled sets. d-f) Slope of the linear fit for each of the PFI variables in a-c as a function of nPFI (excluding nPFI=0) for the observed (blue line) vs the shuffled data (1000 sets, grey histogram).

5.6.5. SSVEP time course: event-by-event image analysis reveals graded changes in conscious perception

After demonstrating that spatially distributed targets were interacting, strongly implying the involvement of high-level neural mechanisms during PFI, we turned to the neural correlates of PFI via EEG analysis of SSVEPs. We first visualized how the changes in PFI were related to changes in the log(SNR) of background flicker using an event-by-event image-based analysis. To compare the time course of button press and SNR across participants, we first sorted, per participant, all instances of PFI

disappearance (or reappearance) by the sum of the number of buttons simultaneously pressed over 3 seconds after (or before) the button press, which we define as "the amount of PFI" (see Methods and Figure 5.8). We then resampled each participants image into a uniform height, to obtain the across-participant mean despite the differences in individual PFI dynamics (see Methods and Figure 5.3). This results in the highest (and lowest) rows of the figures representing events with the highest (and lowest) amount of PFI (Figure 5.8a, 5.b). Figure 5.8c-f show the corresponding RESS log(SNR) related to 1f and 2f background SSVEP responses.

From this analysis, two qualitative insights emerged. First, that RESS log(SNR) for 1f and 2f increase just before button press when targets disappear (at time = 0), and increase with the amount of PFI (Figure 5.8c and 5.8d). Second, RESS log(SNR) for 1f and 2f decrease just before button release at target reappearance, but there is no dependence on the amount of PFI (Figure 5.8e and 5.8f).

To quantitatively compare these differences, we split SNR time courses based on the amount of PFI. Figure 5.8g-j show the mean RESS log(SNR) over each 6 second epoch, separately averaged for events with the amount of PFI between 0 and 1, 1 and 2, or greater than 2. Around the target disappearance events, we found a significant linear effect for the amount of PFI on the SNR for both 1f ($\chi^2(1) = 8.75$, $p = .003$) and 2f ($\chi^2(1) = 8.21$, $p = .004$) responses to background flicker (Figure 5.8g and h). Around target reappearance events, by contrast, the amount of PFI did not significantly affect the SNR (Figure 5.8i and 5.8j, 1f; $p = .76$; 2f; $p = .83$). Figure 5.8k-n displays the time course of the SNR separately for 3 levels of the amount of PFI around the time of button press and release.

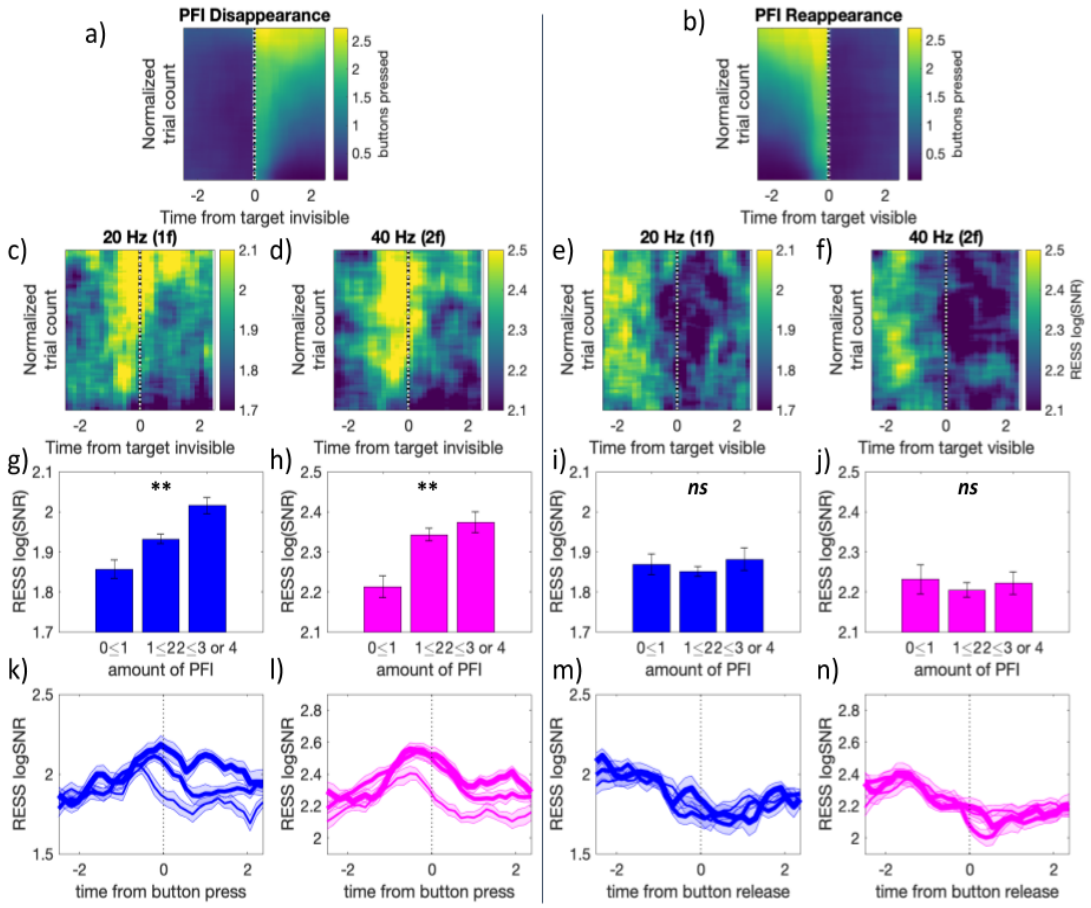

Figure 5.8. The amount of PFI is correlated with the RESS log(SNR) around PFI disappearances, but not reappearances.

Event-by-event image analysis of button press and release (a and b) and RESS log(SNR) (c-f) after sorting based on the amount of PFI per event, per participant. Background responses at 1f are shown on the left column and separated from those at 2f on the right of by dotted lines. g-j) Bar graphs for the mean RESS log(SNR) over -3 to 3 sec as a function of the amount of PFI. k-n) The time course of RESS log(SNR) around the button press or release, separated by the amount of PFI, with three levels of the amount indicated by the thin, middle and thick lines. Error bars in g-j and shading for k-n indicate 1 SEM across participants (adjusted for within-participant subject comparisons Cousineau, 2005).

5.6.6. Reconstruction analysis: SNR time courses during PFI are distinct from those in catch periods

While the previous analysis has shown that changes to the log(SNR) of background flicker were related to the amount of PFI, it does not take into account the effects of closely spaced or overlapping button responses that are required in our multi-target PFI task. Unlike other tasks that have investigated the neural correlates of multistable perception with a single target, our task design allowed graded changes in consciousness to occur in close temporal proximity (< 1 second), and even to overlap (Figure 5.1b). This was not the case for catch periods, which occurred with a simultaneous onset and offset. To account for how much of the log(SNR) time course could be accounted for by sequential responses, we performed an SNR-reconstruction analysis. This analysis quantified the stability of log(SNR) dynamics during PFI, and whether differences in log(SNR) during PMD could be predicted from the same data. In brief, we used 75% of training trials to construct reconstruction kernels, which were the changes to log(SNR) during PFI in this 'training data'. We then applied these kernels to the remaining 25% of 'test' trials, aligning each kernel to the recorded button press time-points (see Methods and Figure 5.4). We then compared the predicted time course of log(SNR) with the actual time course around the button press events in the test trials during genuine PFI and during catch periods. Figure 5.9 visualizes the high quality of prediction for the genuine PFI (Figure 5.9e and g) and the poor predictive quality for catch periods (Figure 5. 9f and h).

To quantify prediction accuracy as the degree of correlation between the predicted and the observed time course, we calculated R^2 between the respective 6-second RESS log(SNR) around button press/release events during genuine PFI and catch periods. For both 1f and 2f, the predicted SNR was correlated more strongly with genuine PFI than the catch, for both disappearances and reappearances (Table 1). Using 3-way repeated measures ANOVA (Table 2), we confirmed that the prediction accuracy is significantly better for the genuine PFI than catch periods (main effect: $F(1, 21) = 151.01$, $p = 4.7 \times 10^{-12}$). We found no or weak main effects or interactions due to other factors (i.e., 1f vs 2f, disappearances vs reappearances).

One source of the difference in the quality of prediction could be the presence of competitive (inhibitory) interactions between the background and target stimuli

during PFI (De Weerd et al., 1995; Weil & Rees, 2011), which are absent during catch periods. Unfortunately, as we could not frequency-tag the targets (i.e., 8, 13, 15 and 18 Hz), we cannot address the nature of these competitive interactions further. To uncover the nature of this interaction, future experiments may try to optimize the parameters in such a way as to frequency-tag both the target (e.g. Weil et al., 2007) and background stimuli during PFI. Next, we continue by analysing the timing of these relative changes during target disappearance and reappearance in more detail, using a cross-point analysis.

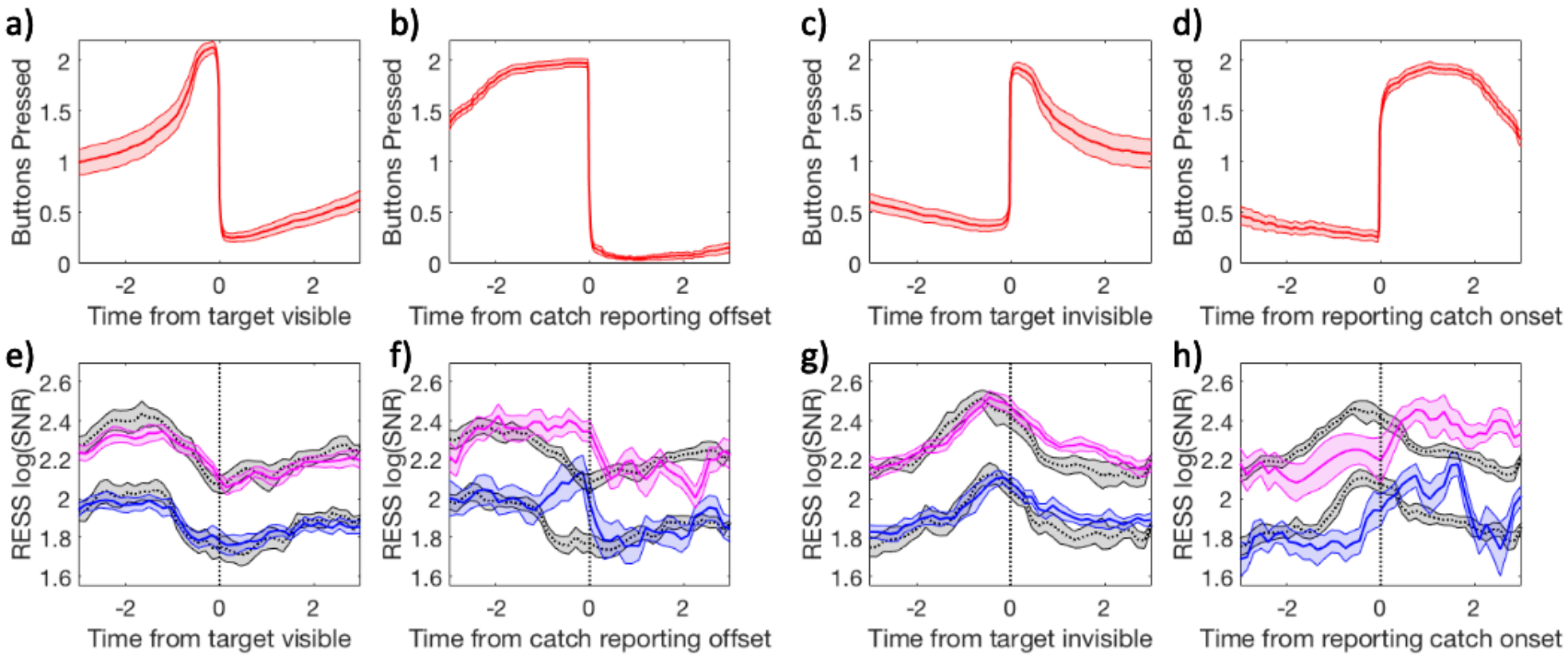

Figure 5.9. Reconstruction analysis.

a-d) mean button-press and e-h) RESS log(SNR) time course across participants around genuine PFI events for reappearance (a, e) and disappearance (c, g) and around catch reappearance (b, f) and disappearance (d, h). Note that for all panels, time 0 is always defined by a button press or release. e-h) The observed SNR time course is shown from test trials (blue for 1f and magenta for 2f), which were not used to construct the reconstruction kernels. The correlation (R^2) between the observed SNR and the predicted SNR (shown in grey) was used to quantify prediction accuracy. Shading represents 1 SEM across participants (corrected for within participant comparisons; Cousineau, 2005).

Table 1. Prediction accuracy (as R^2) across reconstruction analyses

	PFI 1f disap.	PFI 2f disap.	Catch 1f disap.	Catch 2f disap.	PFI 1f reap.	PFI 2f reap	Catch 1f reap.	Catch 2f reap.
Mean	0.50	0.54	0.08	0.13	0.45	0.49	0.12	0.13
Std. error mean	0.05	0.05	0.03	0.03	0.05	0.05	0.03	0.04
Standard deviation	0.23	0.22	0.12	0.14	0.25	0.22	0.16	0.17

Table 2. Results of 2 x 2 x 2 repeated measures ANOVA on R^2 values

	Sum of Squares	df	Mean Square	*F*	*p*	partial η^2
PFI vs. Catch	6.24	1	6.24	151.01	<.001	0.88
Residual	0.87	21	0.04			
1f vs. 2f	0.06	1	0.06	0.94	0.342	0.04
Residual	1.25	21	0.06			
Disap. vs. Reapp.	0.01	1	0.01	0.45	0.512	0.02
Residual	0.67	21	0.03			
(PFI vs. Catch) x (1f vs. 2f)	0.00	1	0.00	0.07	0.792	0.00
Residual	0.65	21	0.03			
(PFI vs. Catch) x (Disap. vs. Reapp.)	0.05	1	0.05	5.34	0.031	0.20
Residual	0.21	21	0.01			
(1f vs. 2f) x (Disap. vs. Reapp.)	0.01	1	0.01	0.47	0.499	0.02
Residual	0.39	21	0.02			
(PFI vs. Catch) x (1f vs. 2f) x (Disap. vs. Reapp.)	0.00	1	0.00	0.28	0.603	0.01
Residual	0.34	21	0.02			

Note. Type 3 Sums of Squares

5.6.7. Cross-point analysis: 1f and 2f background-related responses are temporally distinct during PFI

Our reconstruction analysis similarly predicted both the 1f and 2f components of background-related SNR during PFI events, which is not surprising given that these responses were driven by the same stimuli. Curiously, however, these harmonic responses were topographically distinct (Figure 5.5b). As there is a nascent literature suggesting that SSVEP harmonics may correspond to separate cognitive processes (Kim et al., 2007, 2011), we next investigated these spatiotemporal differences in more detail.

First, we investigated whether the RESS log(SNR) time course differed depending on the nature of disappearances/reappearances: due to physical (catch) or perceptual (PFI). We compared the time courses between target disappearance and reappearance, superimposing these time courses in the same plot and calculating the crossover points of the RESS log(SNR). For 1f (Figure 5.10a and 5.10b, blue), the RESS log(SNR) during disappearances (solid lines) became larger than that during reappearances (dotted lines). This effect occurred from -0.67 seconds prior to subjective report (paired *t*-tests, $p_{cluster} < .001$). Notably, these effects occurred 1.06 seconds later for catch periods (Figure 5.10b, from 0.39 seconds, $p_{cluster} < .001$). For 2f (Figure 5.10a and b, magenta), the RESS log(SNR) also became larger during disappearances than reappearances from -.97 seconds prior to report ($p_{cluster} < .001$), and again, were shifted roughly 1.36 seconds compared to the catch-related time course (Figure 5.10b, from 0.39 seconds; $p_{cluster} < .001$).

The observed divergence (0.3 seconds) in the crossover time for 1f and 2f seemed quite large given that both 1f and 2f were evoked from the same stimulus, using identical participants and events. As such we further investigated if this effect could be observed at the participant level. For this analysis, we calculated for each participant the first time point at which the strength of background RESS log(SNR) during disappearance exceeded that during reappearance (running paired *t*-tests). Using this criterion, we found that 2f responses crossed over at -1.02 seconds (*SD* = 0.41), 170 ms seconds earlier than 1f responses, at -0.85 seconds (*SD* = 0.37, Wilcoxon signed rank test, z = 2.13, *p* = .012). No difference was observed in cross over time for the catch-related 1f and 2f time courses (*p* = .14).

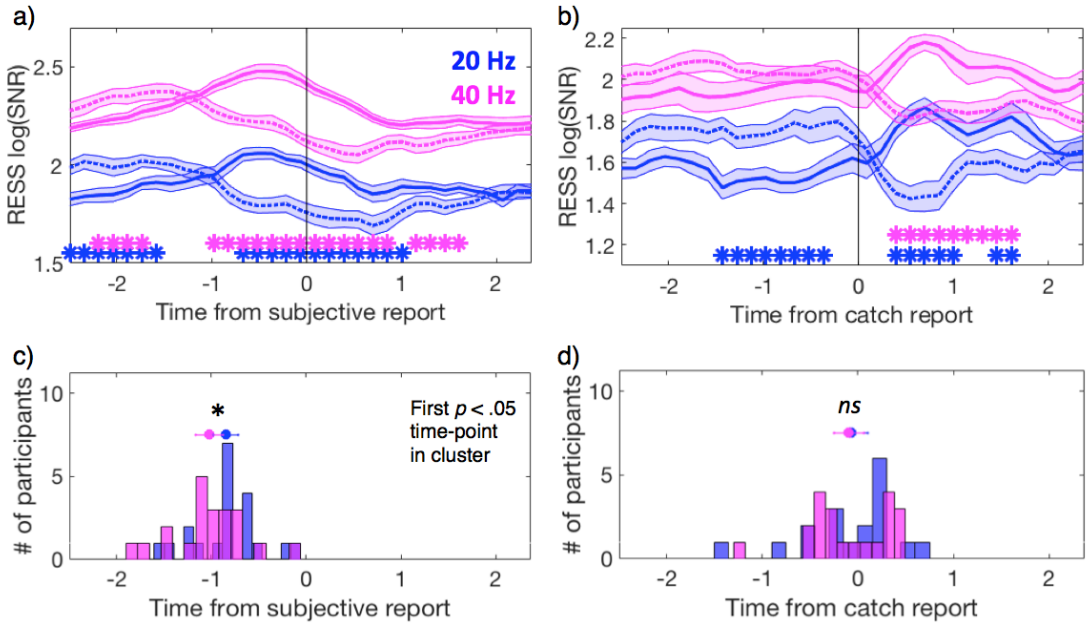

Figure 5.10. Distinct temporal profile of the harmonic responses.

a and b) Relative time course of the 1f (20 Hz, blue) and 2f (40 Hz, magenta) RESS log(SNR) during PFI events (a) and catch periods (b). Solid and broken lines represent disappearance and reappearance, respectively. c and d) Participant-level histograms for the first significant time point when comparing between the RESS log(SNR) for disappearance and reappearance during PFI (c) and catch (d). Horizontal lines indicate 1 SE about the mean corrected for within-subject comparisons (Cousineau, 2005).

5.6.8. Spatial Correlation: 1f and 2f background responses are spatially distinct during PFI

One potential factor that could have contributed to the difference in the crossover time between 1f and 2f is a difference in the spatial filters used for 1f and 2f within RESS analysis. In fact, when we focused only on the (non-RESS) log(SNR) from a single electrode (POz), the difference in cross-over times between 1f and 2f was not significant at the group or participant level. Given this, we further analysed whether the spatial characteristics for 1f and 2f were also distinct without using RESS spatial filtering during PFI.

Around the catch events, spatial correlations across 64 channels were constant (Figure 5.11b). However, when targets disappeared during PFI, the spatial correlation between 1f and 2f transiently increased (Figure 5.11a). The difference between the time courses was significant for the time-window -0.67 to 0.25 seconds around subjective report (paired *t*-tests at each time point, $p_{cluster} < .001$). The same pattern of

results was maintained when using a parietal or occipital sub-region of electrodes (but no change in correlation was seen for frontal or temporal electrodes), indicating that synchronous changes in predominantly parieto-occipital SNR were responsible for changes to the whole-head correlation over time. The same pattern was also observed when subtracting the mean log(SNR) per channel prior to calculating this spatial correlation over time.

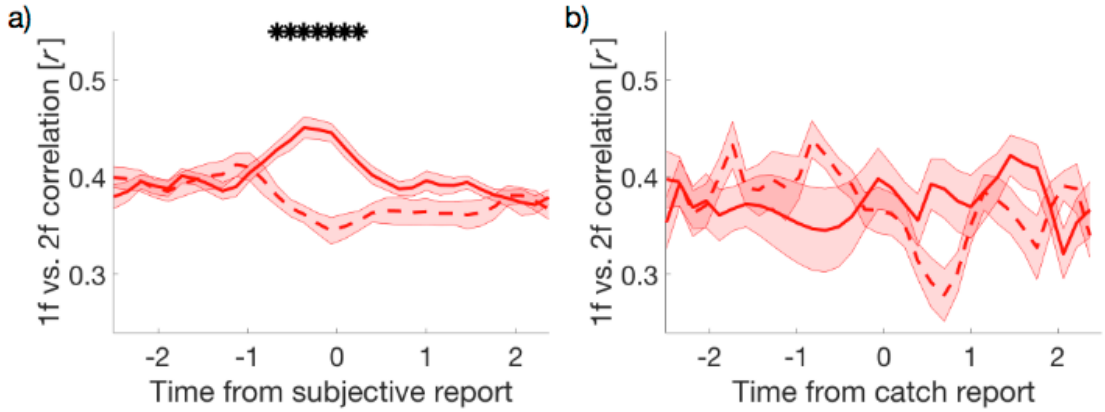

Figure 5.11. Time course of the spatial correlation coefficient (r) between 1f and 2f (non-RESS) log(SNR) across 64 electrodes.

Correlation coefficient was computed across 64 electrodes at each time point per participant. The mean time courses of correlation coefficients are shown for target disappearance (solid), and reappearance (dotted) around a) PFI, and b) catch periods. For PFI, we show the mean correlation value obtained after down-sampling PFI events to 24 (the maximum number of catch periods), over 100 repetitions of this down-sampling procedure. Asterisks denote the time points with significantly different correlation coefficients between PFI disappearances vs reappearances (paired t-tests, cluster corrected). Shading reflects the SEM across subjects corrected for within-subject comparisons (Cousineau, 2005).

5.7. Discussion

We combined a multi-target perceptual filling-in (PFI) paradigm with frequency-tagged EEG. This combination has revealed novel insights into the mechanisms of PFI phenomena, including unexpected asymmetric neural correlates for graded disappearances and reappearances (Figure 5.8), and spatiotemporal distinctions between steady-state visual evoked potential (SSVEP) harmonics (1f and 2f background responses, Figure 5.10 and 5.11). Here, we discuss these findings focusing on several advantages of our experimental paradigm.

5.7.1. Multi-target PFI to track changes in conscious perception

Frequency-tagging has been used to study the neural correlates of consciousness, mainly in combination with binocular rivalry (Brown & Norcia, 1997; Jamison et al., 2015; Katyal et al., 2016; Sutoyo & Srinivasan, 2009; Tononi et al., 1998; Zhang et al., 2011). When reporting on perceptual reversals in these paradigms, neural activity that is associated with purely perceptual processes has been entangled with the processes of attention and the act of report (Aru, Bachmann, Singer, & Melloni, 2012; Miller, 2007; Tsuchiya, Wilke, Frässle, et al., 2015; van Boxtel & Tsuchiya, 2014; van Boxtel et al., 2010a). To reduce these confounds, replays with the physical removal or alternation of stimuli have been used as a standard control condition to compare with, for example, genuine perceptual switches in binocular rivalry (Frassle et al., 2014; Lumer et al., 1998). As the requirements for both perceptual and physical reversals involve attention and report, it was hoped that contrasting these conditions would isolate the neural processes specific to endogenously generated changes in consciousness. Despite various attempts, generating catch movies (or the physical replays) that perceptually match endogenously-generated conscious changes in perception remains a significant challenge, due to highly complex phenomenal dynamics during rivalry (Knapen et al., 2011; Wilson et al., 2001). Until these report-related attentional confounds are resolved, results from such experiments, particularly binocular rivalry, need to be interpreted with caution (Blake et al., 2014; Frassle et al., 2014; Naber, Frässle, & Einhäuser, 2011).

Unlike binocular rivalry, perceptual changes during PFI are crisp and simple, suggesting PFI can prove to be a useful psychophysical tool to study the NCC. The simplicity of PFI phenomenology allowed us to 1) generate catch events that were difficult to distinguish from real PFI (see Movie 1), and 2) to ask untrained participants to accurately and simultaneously report on multiple targets, while allowing us to check the quality of their report. Equipped with this technical advance, we observed a facilitation of simultaneous target disappearances and reappearances, strongly implying long-range interactions between the distant targets.

The multi-target display also allowed us to have a more objective graded measure of differences in the contents of consciousness (i.e., the amount of PFI), which revealed an asymmetry between the neural correlates of disappearances and reappearances. At this point, we have no straightforward explanation for this. One possible explanation is the difference in saliency between PFI disappearances and reappearances, as reappearances can be predicted with higher spatial and temporal accuracy than disappearances. Increased spatial accuracy follows from the fact that reappearances can only occur at locations where a target has already disappeared moments prior. As the duration of PFI is also short compared to the 60-second trial (Figure 5.7), reappearances can also be predicted with greater temporal accuracy than multi-target disappearances. Thus, PFI disappearances may be more unexpected than reappearances, enhancing their subjective saliency. Indeed, greater phasic pupil responses to target disappearances than reappearances have been reported in motion-induced blindness (Kloosterman et al., 2015; Thomas et al., 2017) which may be closely related to PFI (Devyatko et al., 2016; Hsu, Yeh, & Kramer, 2004, 2006; New & Scholl, 2008). This difference in spatiotemporal saliency might have resulted in the asymmetric patterns of log(SNR) based on the number of disappearing or reappearing targets (Figure 5.8). To better understand the mechanisms of this asymmetry, further studies employing a paradigm that feature multi-target and graded conscious perception will be necessary. Another possibility unexplored in the present dataset is the contribution of eye-movements to PFI phenomena. Future studies could address this limitation and test whether the pattern of saccades and microsaccades differ when targets are visible or invisible during PFI.

5.7.2. Insights into PFI mechanisms

Our results are relevant to two popular models of PFI. The first is an isomorphic model. This model proposes the primary substrate of PFI are neurons in early retinotopic areas corresponding to target regions, which are activated via neurons corresponding to their target surrounds through lateral connections (De Weerd et al., 1995; Pessoa et al., 1998). The model specifically proposes a two-stage process, where a first stage of seconds-long boundary adaptation is followed by a second stage of near instantaneous interpolation of the target location by surrounding

visual features (Spillmann & De Weerd, 2003). The second is a symbolic model, whereby filling-in occurs when the visual system ignores an absence of information (Dennett, 1991; Kingdom & Moulden, 1988; O' Regan, 1992). In this model, the phenomenon of filling-in is realized at a (possibly higher) representational level, whereby a region devoid of information is symbolically labelled as 'more of the same' background, and thus is rendered invisible.

In favour of the isomorphic model, previous electrophysiological data has recorded increased spike rates in regions responding to a filled-in pattern in monkeys (De Weerd et al., 1995). Importantly, De Weerd et al.'s (1995) single-unit study did not supply behavioural reports, which we provide on an event-by-event manner, so the exact timing of their increases in neural activity could not be established relative to the onset of a perceptual disappearance. By recording simultaneous behavioural reports, we show that an increase in background SNR precedes PFI events in humans. This slow, seconds-long increase in background-related SNR prior to PFI events supports an active mechanism as a catalyst for PFI, which is central to the isomorphic model.

On the other hand, the symbolic model that suggests that filling-in happens in higher-level visual areas (Pessoa et al., 1998) is also consistent with our behavioural findings. We observed a synergistic effect among spatially distant targets, which implies the involvement of neurons that have larger receptive fields, typically found only in higher-level visual areas (Dumoulin & Wandell, 2008; Yoshor, Bosking, Ghose, & Maunsell, 2007). This across-quadrant faciliatory interaction extends a previous report of within-quadrant interactions during PFI (De Weerd et al., 1998, experiment 4). More specifically, this synergistic PFI across quadrants may point to a mechanism that facilitates perceptual grouping (Wagemans et al., 2012).

Grouping may also interact with attentional mechanisms. Indeed, attending to shared features such as temporal modulation has been shown to enhance the binding of distributed visual regions into a perceptual group (Alais, Blake, & Lee, 1998). As attending to shared features such as colour (Lou, 1999) or shape (De Weerd et al., 2006) increases the disappearance of peripherally presented targets, fluctuations in attention to the targets as a group may also have impacted on multiple-locations synergistically. Alternatively, the simultaneous disappearance of multiple targets could be due to random fluctuations of the brain's response to the background (potentially also modulated by attention). Since the background surrounds all targets,

a temporary increase in response could affect the visibility of all targets simultaneously.

Overall, our results are not compatible with the view that PFI is a phenomenon that results purely due to local adaptation processes in the retinal or low-level visual areas. Instead our results are compatible with the view that both retinotopic and contextual influences, possibly through lateral connections, determine the dynamics of PFI (Sasaki, 2007).

5.7.3. Spatiotemporal profiles of 1f and 2f background SSVEP are distinct

Another insight that arose from our application of SSVEP to study PFI regards the difference in spatiotemporal profiles of 1f and 2f responses (Figure 5.10 and 5.11). This difference was specifically modulated around the time of PFI. In the literature, 1f and 2f are traditionally considered to be similar, as they are dictated by the same stimulus input (Norcia et al., 2015). Recently, this assumption has been challenged by the finding of an attentional modulation of 2f, but not 1f, with concomitant changes in hemispheric lateralization for the topography of SSVEP responses (Kim et al., 2011; Kim & Verghese, 2012). While an increased spatial distribution of 2f compared to 1f is consistent with our results, where 1f was strongest over mid-occipital sites and 2f extended laterally (Figure 5.5), the flicker stimuli used in our experiments differ from those studies that optimized differentiating 1f from 2f (Kim et al., 2011). As such, extending this interpretation to our findings should be done with caution, but the temporal advantage of the 2f crossover compared to the 1f crossover would be consistent with a covert attentional modulation of 2f that instigates a perceptual change. Future studies with an explicit attentional manipulation will be needed to confirm whether the harmonic differences we have reported are due to the allocation of attention.

5.7.4. Conclusions

Here we extend efforts to refine NCC paradigms, by using PFI. Unlike traditional stimuli, PFI has the advantage that perceptual changes can be easily

mimicked physically, and that participants can accurately report on multiple changes in consciousness occurring in close temporal proximity without much training. While genuine PFI and physical catch periods were phenomenally similar, we revealed significant differences in their respective neural substrates through our SNR reconstruction analysis, and suggest that these differences are due to the presence of competitive mechanisms supporting perceptual disappearances, but not physical disappearances. Future studies that succeed in tagging both targets and surrounds in PFI would be able to investigate the nature of this competition. They may also reveal why there are significant differences in the dependence on the amount of PFI for disappearances, but not reappearances, which we have tentatively linked to differences in the level of expectation and saliency. These are intriguing empirical questions to be resolved in the future by capitalizing upon the peculiar effect that attention increases PFI (De Weerd et al., 2006; Lou, 1999) and/or by utilizing SSVEP-based no-report paradigms (Tsuchiya et al., 2015). We hope that our approach that combines under-utilized PFI with SSVEP techniques will inspire various novel designs to address this central question in cognitive neuroscience: the neural basis of attention and consciousness.

Chapter 6: Neural responses to an invisible target increase during perceptual filling-in

Author names and affiliations: Matthew J Davidson[1]+, Will Mithen[1], Jeroen van Boxtel[1,2]*+ Naotsugu Tsuchiya[1]*+,

[1] *School of Psychological Sciences, Monash University, Victoria, Australia.*
[2] *School of Psychology, Faculty of Health, University of Canberra, Canberra, Australia*

* Equal contribution
+ Corresponding authors

6.1 Article introduction

The previous study confirmed that participants could accurately report on the disappearance of multiple-items simultaneously, and that a dynamically updated background display could entrain strong SSVEPs. Moreover, the frequency-tagging of visual background information was used to capture graded neural responses correlated with changes in subjective awareness. A key motivation of combining frequency-tagging with PFI was to develop a paradigm capable of investigating the unique interaction between disappearances and selective attention (De Weerd et al., 2006; Hohwy, 2012; Lou, 1999).

Previously, we were unable to recover the frequency-tag of our small and peripherally presented flickering targets. This result is not surprising, as SSVEPs are strongest for larger, centrally presented stimuli (Norcia et al., 2015). Here, we refine this paradigm to frequency-tag both targets and background surrounds in a novel form of filling-in, in which the colour of centrally presented targets is filled-in by the surrounding image background. This novel PFI paradigm allowed us to ask a unique question regarding the fate of SSVEPs elicited by a filled-in region. Specifically, do

SSVEPs decrease in strength along with target visibility, or increase in support of an attentional mechanism driving PFI dynamics?

6.2. Abstract

During perceptual filling-in (PFI), salient targets in the visual periphery disappear and are interpolated by the surrounding image background, despite their continued physical presence. PFI is unique among perceptual disappearance phenomena, in that orienting attention toward target features hastens their disappearance. Contrary to claims that attention precedes conscious awareness, this suggests attention sometimes diminishes the contents of conscious perception, positioning PFI as an attractive resource for investigating the hotly debated dissociation between attention and consciousness. Here we asked participants to report on PFI by pressing one of four unique buttons corresponding to four simultaneously presented targets. During PFI, we flickered both target and background regions (15 and 20 Hz, respectively), to entrain steady-state visually evoked potentials (SSVEPs) in the electroencephalogram. Traditionally, in non-PFI paradigms, SSVEP strength increases when attention is paid to targets, and decreases when targets are rendered invisible. Thus, investigating targets disappearances, which attract attention, offers a means to track the opposing signatures of attention and conscious perception during PFI. Here we report that during PFI, frequency-tagged responses to target stimuli increased when targets were disappearing, but in direct opposition to the effect of conscious perception on SSVEP responses. This pattern was not obtained for SSVEPs from the background stimuli, nor any SSVEPs during the phenomenally matched physical replay of PFI. Endogenous changes in consciousness were also preceded by distinct neural correlates of integration, as measured by intermodulation frequencies, confirming that PFI is mediated by interactions between target regions and their surrounds. These results show that PFI can disentangle the contents of consciousness from attention.

6.3. Introduction:

Must attention and consciousness always work in parallel, or are their neural mechanisms distinct? This question has been the source of ongoing debate (Koch &

Tsuchiya, 2007; Posner, 2012; Smout & Mattingley, 2018; Watanabe et al., 2011) with an unmet empirical need for paradigms which can dissociate the neural signature of attention during changes to the contents of conscious awareness (Koch & Tsuchiya, 2007; Pitts, Lutsyshyna, & Hillyard, 2018; Tallon-Baudry, 2012; van Boxtel & Tsuchiya, 2014; van Boxtel et al., 2010a).

Perceptual filling-in (PFI) is one candidate paradigm, and represents a class of multistable stimuli that confers distinct advantages over its predecessors (Komatsu, 2006; Weil & Rees, 2011). PFI occurs when a distinct region of the visual periphery disappears and becomes interpolated by the surrounding image background, and is a relatively underexplored class of multistable phenomena (Kim & Blake, 2005; Sterzer et al., 2009). Compared to more widely studied forms of multistable stimuli that require inter-ocular suppression, PFI occurs regularly in normal vision (Anstis & Greenlee, 2014; De Weerd, 2006; Durgin et al., 1995). Multiple regions can also be rendered invisible during PFI, allowing a unique opportunity for a graded response in the amount of perceptual change to be captured (**Chapter 5**). Most critically however, covertly directing attention toward distinct targets in the visual periphery hastens their filling-in (De Weerd et al., 2006; Lou, 1999), suggesting that attention to and conscious perception of visual targets can be dissociated during PFI. This finding contradicts the claims that attention precedes conscious awareness, or that they are one and the same.

Recently, a powerful technique to track changes in neural activity has been applied to research on attention and consciousness. This technique, called frequency-tagging, increases the spatial-selectivity of EEG data, and can characterize neural processes that are unique to specific visual stimuli, even when simultaneous or overlapping (e.g. Andersen et al., 2008; Müller et al., 2006). The technique works by rapidly flickering visual stimuli, entraining neural populations processing the flickering stimulus to evoke a steady-state visually evoked potential (SSVEP): the neural populations responsive to the stimulus effectively become 'tagged' at the same unique frequency. Changes in the strength of this frequency-tag, i.e. the signal-to-noise ratio of the SSVEP response, are observed in comparison to neighbourhood frequencies in the EEG, and are assumed to reflect changes in processing resources allocated to a particular stimulus.

Frequency-tagging can be leveraged to analyse the potentially confounding contribution of attention on consciousness to perception. By combining frequency-

tagging with other types of multistable stimuli, it has been demonstrated that when flickering stimuli enter into perceptual awareness, the strength of frequency-tagged activity also increases (Brown & Norcia, 1997; Lansing, 1964; Tononi, Srinivasan, Russell, & Edelman, 1998b; P. Zhang et al., 2011). Indeed, during binocular rivalry, when two flickering stimuli compete for perceptual dominance, the contents of consciousness can be tracked by the strength of each frequency-tag, with the largest strength of the two tags corresponding to the current contents of consciousness. Apart from the main flickering frequencies, additional intermodulation (IM) components of the two competing flicker frequencies (i.e., linear combinations of multiples of frequency 1 and frequency 2) are observable in the EEG spectra (Katyal et al., 2016; Zhang et al., 2011). These IM components have been proposed to index interactions between stimulus representations (Ales, Farzin, Rossion, & Norcia, 2012; Alp et al., 2016; Boremanse et al., 2013; Gordon et al., 2017; Gundlach & Müller, 2013), and have been shown to peak prior to a perceptual change (Katyal et al., 2016), as well as increase when stimuli are attended (Gordon et al., 2017; Kim, Tsai, Ojemann, & Verghese, 2017; Kim & Verghese, 2012). Together, the strength of frequency-tagged stimuli and their IM components offer an objective means of tracking the contents of consciousness, and the interaction between competing stimuli, respectively. To date however, no study has investigated the neural markers of these overlapping processes, specifically both target *and* background stimulus representations, in the context of PFI.

In addition to increased SSVEP strength with increased stimulus saliency and awareness, selectively attending to flickering stimuli also increases the strength of SSVEP responses (Morgan et al., 1996; Müller & Hubner, 2002; Müller et al., 2003; Müller, Picton, et al., 1998; Walter et al., 2012). Notably, this increase in SSVEP strength with attention also occurs when flickering stimuli spatially overlap (Andersen et al., 2008; Müller et al., 2006; Wang et al., 2007), and when a flickering stimulus is embedded in an attended noise sequence in the absence of participant awareness (Smout & Mattingley, 2018). These results demonstrate that even invisible stimuli can evoke stimulus specific neural responses within the focus of attention.

Since both attention and awareness increase the strength of the SSVEP, it has not been possible to distinguish their effects during the presentation of ambiguous stimuli. However, as mentioned, during PFI, the influence of subjective awareness and attention predict opposing effects for the strength of frequency-tagged EEG

responses. If PFI is mediated by attention, then attending to target regions may increase SSVEP strength, even as the flickering targets themselves become invisible. As both target disappearance or reappearance may capture attention (Kloosterman et al., 2015; Naber et al., 2011), we devised a paradigm to test whether neural responses during PFI were associated with perceptual transitions per se, or to attentional capture. In a previous investigation, we showed that prior to the disappearance of these peripheral stimuli, the SSVEP response to the visual background regions is increased. We also found that multiple PFI targets disappeared together far longer than expected by chance, suggesting that invisible target regions were grouped despite their physical separation. We had tentatively linked both effects to an increase in the allocation of attention. However, we were unable to track changes in target SSVEPs, preventing us from examining whether they increase or decrease upon PFI, nor the interaction between targets and their surrounds. The current design was developed to allow, for the first time, the tracking of both targets and background. To foreshadow the results, we found that as flickering visual stimuli disappear from conscious awareness, the strength of frequency-tagged responses to target stimuli *increased.* Our results are thus consistent with PFI resulting from an increase in the allocation of attention to target regions, in support of the role of higher-order cortical areas mediating this striking phenomenon, and support the dissociation between attention and consciousness.

6.4. Method

6.4.1. Participants

19 healthy adults (12 female, 19-40 years, $M = 26.95$, $SD = 7.63$) with normal or corrected-to-normal vision participated in this study. They were recruited via convenience sampling from students at Monash University, provided written informed consent prior to taking part, and were paid 20 AUD per hour of their time (approximately 3 hours total). Participants with self-reported sensitivity to flickering stimuli were excluded. Experiment start times were restricted to 10am and 2pm to limit fatigue effects, and participants who usually drank caffeine were asked to consume a standard quantity (one beverage) in the two hours prior the experiment beginning to limit the effects of this stimulant on their performance. Ethics approval

was obtained from the Monash University Human Research Ethics Committee (MUHREC #CF12/2542 - 2012001375).

6.4.2. Apparatus and stimuli

Participants sat approximately 50 cm from a 29 cm by 51 cm computer monitor (resolution 1080 x 1920 pixels, subtending approximately 32 x 54° visual angle, refresh rate 60 Hz). The background of our stimulus (Figure 6.1) flickered at a rate of 20 Hz by randomly alternating between 100, pre-calculated, random patterns generated at the start of the experiment. The random patterns were made up by dividing the screen into squares of 10 x 10 pixels, with each square's luminance randomly set to either black or white. Visually, this procedure appeared similar to random noise, with an equal proportion of black and white patterns. At the centre was a 5 mm diameter red dot to be used as a fixation point.

One target was presented in each of the four visual quadrants (top left, top right, bottom left, and bottom right). Targets subtended 6.08° visual angle, within which the white background squares were instead blue/purple (RGB value of 205, 205, 255). There were no line-contours at target boundaries, as target boundaries were defined by the pixelated squares of the background they were overlaid on (Figure 6.1b). As a result, during PFI, the target regions were indistinguishable from the surrounding background texture. The refresh rate of texture within the target regions was set to 15 Hz, with target eccentricity individually calibrated per participant to evoke optimal PFI and SSVEP strength (Figure 6.1 b-c, **Real-time SSVEP calibration of target eccentricity**).

6.4.3. EEG acquisition

Throughout each session whole-head EEG was recorded with 64 active electrodes arranged according to the international 10-10 system (BrainProducts, ActiCap). Electrode impedances were kept below 10 kΩ prior to experimentation, and recorded using the default reference (FCz) and ground electrode (AFz) via Brainvision recorder software (sampling rate =1000 Hz, offline bandpass of 0.5-70 Hz). All EEG data was stored for offline analysis using custom Matlab scripts (Ver: R2016b), as well as the EEGLab (Delorme & Makeig, 2004) and Chronux (Mitra &

Bokil, 2008) toolboxes. Prior to experimentation, the BCI lab toolbox was also implemented to calibrate real-time SSVEP strength (see **Real-time SSVEP calibration of target eccentricity**).

6.4.4. SSVEP Signal-to-Noise Ratio (SNR) calculation

In the SSVEP paradigm, we compute the signal-to-noise ratio (SNR) at each frequency, which in logarithmic scale corresponds to log of the power at each frequency subtracted by the mean log power across the neighbourhood frequencies (Figure 6.3).

We computed the SNR at frequency f (Hz) as the mean log power over the neighbourhood frequencies for f subtracted from the log power at f (Norcia et al., 2015). Throughout this paper, the neighbourhood used in SNR calculation always excludes the frequency half-bandwidth (Hbw = (k+1) /2T). Where k = the number of tapers used in time-frequency decomposition and T= temporal window (seconds) of the data. For example, for a single taper, 2 second EEG spectrum, a 2 Hz neighbourhood used as 'noise' in SNR calculations would be defined as [f-2.5, f-0.5] and [f+0.5, f+2.5]; Hbw = 0.5 Hz. Throughout this paper, the frequency Hbw and noise neighbourhood are identified in the relevant figure captions. The length of EEG epochs used to calculate SSVEP characteristics vary (e.g. 60 seconds Figure 6.3 vs 2.5 seconds Figure 6.4), so as to compare whole trial SSVEP strength to time-varying changes in SSVEP strength epoched about participant responses.

6.4.5. Real-time SSVEP calibration of target eccentricity

Combining PFI with frequency-tagging is challenging, as the parameters which optimise each phenomenon are in direct opposition. Specifically, SSVEPs are strongest at central fixation, and for large visual targets. PFI increases with target eccentricity, and increases in likelihood as targets shrink in size. Previously, we have attempted to frequency-tag both target and background flicker using four targets widely spaced over the visual periphery, at 13.3° diagonal eccentricity from the centre of the screen (**Chapter 5**). As this procedure was unsuccessful in frequency-tagging responses to the targets, we here increased the size of our targets (from 4.56° to 6.08°), and calibrated the eccentricity of targets for each individual participant.

The purpose of the calibration procedure was to find a target eccentricity at which PFI readily occurred, while evoking observable peaks in the EEG spectra at the target flicker frequency (15 Hz). Targets were initially centred close to fixation (3° visual angle from centre). Participants were instructed to fixate a central dot, and were asked to report on PFI. If participants reported perceptual filling-in, they were then asked to describe their experience more explicitly. They were asked if targets were a) almost always visible, only occasionally disappearing, b) appearing and disappearing often, for a few seconds at a time; c) blinking in and out rapidly or d) relatively invisible, only occasionally reappearing. Responses at b) were optimal for the current experiment, to capture both periods of perceptual reversals (i.e., transition states in PFI), and stability. Responses of a) would result in target eccentricity being increased, responses of c) or d) would decrease target eccentricity. Target eccentricities were linearly spaced along the diagonal from centre (from 3 - 5.5°, in steps of 0.3°). All participants reported that perceptual filling in occurred. For six participants, optimal perceptual filling in was not achieved, as for them, the targets were mostly gone for long periods of time, and only occasionally reappeared.

Concurrently with these manipulations, the power and log(SNR) spectra at POz were displayed in real time. Due to the computational demands of presenting frequency-domain EEG spectra in real time, no inferential statistics were used to define adequate SSVEP strength. Readily observable 15 and 20 Hz peaks in the EEG spectra were taken as evidence of frequency-tagging at face value. In the absence of readily observable peaks in the real time EEG spectra, target eccentricity was reduced, and the process was repeated. If tagging was unsuccessful at all settings (n=2), the largest and most central target position at which any perceptual filling-in occurred was adopted. This was under the assumption that repetition over many trials may still result in a frequency-tag, which was confirmed in our analyses. Figure 6.1 displays the final calibrated target eccentricity across participants (N=19).

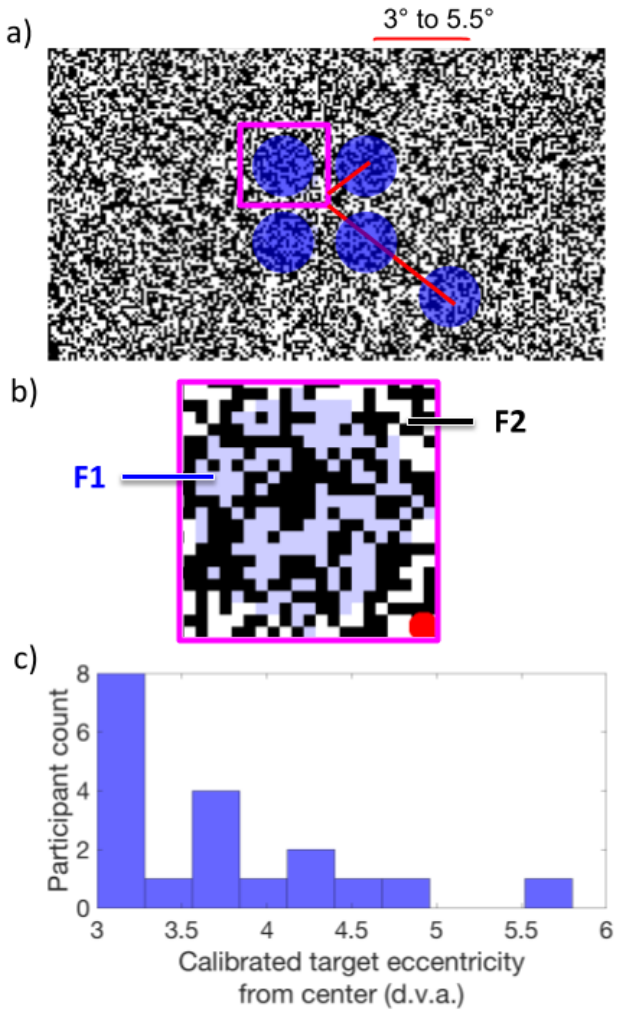

Figure 6.1. Stimulus configuration.

a) Four flickering targets were defined by a blue region, against a 20 Hz flickering background. Colour and target size is enhanced for visualization. Prior to experimentation, an online SSVEP protocol estimated the optimal target eccentricity for flickering targets to elicit a steady-state response (at 15 Hz), while still resulting in PFI. b) An enlarged representation of the target and background stimulus, F1=15 Hz, F2 = 20 Hz stimulus refresh rate. c) Across all participants, the final target eccentricity used throughout each experiment is shown.

6.4.6. Experimental procedure

Participants were instructed with the following script: "Fixate on the red dot. If you perceive that any of the four targets has completely disappeared, press the button corresponding to that target and hold it down for as long as you perceive that target to be absent. If more than one target vanishes simultaneously, try to report on them all as accurately as possible." Specifically, they were instructed to press keys 'A', 'Z', 'K',

and ‘M’ on a traditional QWERTY keyboard, assigning them to the upper left, bottom left, upper right, and bottom right targets, respectively. After calibration, each participant completed one 60-second practice trial, followed by 48 self-paced 60-second experimental trials. Participants took mandatory breaks of 3-5 minutes every 12 trials, while EEG recordings were paused and channel impedances were monitored.

6.4.7 Catch periods

Each trial included one randomly generated catch period of between 3.5 and 5 seconds (drawn from a uniform distribution), during which between one and four targets were physically removed from the screen. Participants were unaware that these physical catch periods were included, and to ensure they were minimally distinguishable from genuine PFI, they did not occur in the first 10 seconds of each trial (Schieting & Spillman, 1987). The order of all catch periods were randomized for each participant, as were the location of removed targets in the case of one, two and three targets. These physical catch periods enabled us to monitor how well participant were able to report on the visibility of four simultaneously presented targets, and additionally served as a control condition for comparison with the neural signals evoked by genuine PFI.

6.4.8 Participant and trial exclusion based on catch periods

We have previously demonstrated that the use of catch periods can identify participants who are unable to report on four simultaneous targets, as well as experimental trials which are unsuited for further analyses (**Chapter 5**). In particular, whether a participant accurately reported the catch periods (via button press) was used to estimate that participant’s attention on task.

As we embedded catch periods throughout our experimental trials, the physical removal of a target (catch onset), could occur for targets during PFI. These events were more frequent for participants experiencing greater amounts of PFI, and as such, we separated visible catch onsets (Figure 6.2a), from invisible catch onsets to estimate attention on task. To determine whether participants were responding to catches appropriately, we estimated the baseline button-press likelihood per

individual participant, by performing a bootstrapping analysis (with replacement). To perform this bootstrap, for any catch onset in trial T at time S (seconds), we randomly selected a trial T' (T=T' was allowed) and we epoched [S-2, S+4] button press data at corresponding catch target locations in T'. To accommodate for the instance of PFI at catch onset, we did not retain the button-press time course for bootstrap analysis if the corresponding buttons were already pressed for the entire 1 second prior to catch onset time of interest (between 0 < 1 second of button press was allowed). This enabled an estimate of button-press close to zero, and thus a more accurate comparison of response to the physical removal of targets during visible catch onset.

We repeated this for all available trials (T=1...48) to obtain a single bootstrapped set of 48 trials per participant. This procedure was repeated 1000 times, and the mean button press time-course of each 48-trial set was retained as the null-distribution for button-press at time of catch onset. We used z-scores of ±1.96 as the CI for the null distribution for each participant (after first converting the data using the logit transformation due to violation of normality). We defined the catch-onset reaction time as the first time point after which the mean button-press data exceeded the top CI. Figure 6.2d shows the catch response for an example participant retained for analysis. We excluded participants (n=3) with no catch onset reaction time in the first two seconds (i.e., [0, S+2]). We note that for two of these participants it appeared that buttons were consistently released at catch onset - potentially indicating buttons were released during PFI rather than pressed as per instructions. For the remaining participants, the mean reaction time to respond to catch onsets, and thus the disappearance of a peripheral target was 0.68 seconds (SD = 0.31).

Having identified which participants could successfully indicate target disappearance based on their button press data (N=16), we continued to remove any trials in which a catch was not correctly detected from subsequent analysis. This ensured that throughout each trial participants were accurately reporting on PFI, and that all retained data was indicative of complete attention to target regions. To identify individual trials for exclusion, we regarded a catch trial as failed if the corresponding button was not pressed for at least 50% of the allowed response time window. This window was the duration of the catch (3.5 to 5 seconds). Figure 6.2c shows a histogram for proportion of rejected trials per participant.

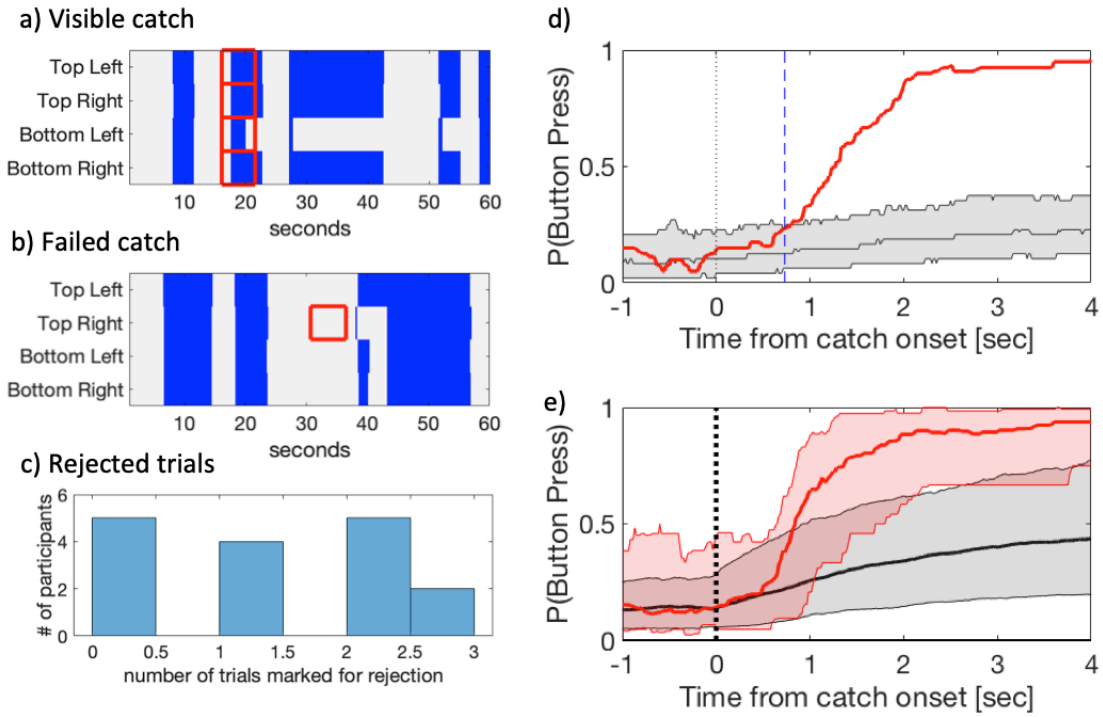

Figure 6.2. Example catch responses following the physical removal of flickering targets at catch onset.

a-b) Catch trials could occur in 1- 4 locations on screen. c) displays the proportion of trials in which catch onsets were not correctly reported per participant, marking each trial for rejection. d) Proportion of trials in which a catch was reported for a single participant. The first time point that the observed button press data (red) exceeded the bootstrapped CI (grey) corresponds to the catch reaction time to visible catches (0.73 sec for this participant, marked with a vertical blue dashed line). e) The mean time course for the likelihood of button press and its bootstrapped sets across participants, shown in red and grey respectively. Shading represent the CI (computed with logit transform and presented after reverse transform) across participants. c) Histogram of the percentage of trials rejected per participant, based on individual catch analysis.

6.4.9. PFI location-shuffling analysis

As shown in Figure 6.2 a-b, targets frequently disappeared together, rather than in isolation. To investigate whether this synergistic effect exceeded that which would be expected by chance, we performed a shuffling analysis to create a null distribution that destroyed the temporal correlation between targets (**Chapter 5**). Specifically, we created 1000 shuffled trials per participant, that could contain the button-press data for each location from any trial throughout their experimental session. The trials that each button-press time course was selected from were sampled from a uniform distribution, and thus could include multiple locations from the same trial. This allowed us to conserve the total amount of PFI recorded, while ensuring

that the button-press data at a given location could come from any independent trial (e.g. top left = trial 10, top right = trial 43, bottom left = trial 12, bottom right = trial 2). If target disappearances during PFI were independent, then destroying the temporal correlation in shuffled data should not matter, and shuffled and experimental data would be similar. We repeated our behavioural analysis (detailed below) on this shuffled data, with the results displayed in Figure 6.9.

6.4.10. Linear-mixed effect analysis – Behaviour

All statistical analyses were performed using Matlab (Ver: R2016b). To examine whether an increasing number of filled in targets (nPFI; n=0, 1, 2, 3 or 4) affected PFI characteristics, we used linear-mixed effect (LME) analysis including intercepts for participants as a random effect. We performed likelihood ratio tests between the full model and a restricted model which excluded the factor of interest (nPFI; Glover & Dixon, 2004; Pinheiro et al., 2014; Winter, 2013) These analyses were performed on both observed and shuffled data.

We also performed LME analyses to compare magnitude and direction for the slopes of observed and shuffled data. For this analysis, we retained the slope (β; 1st order polynomial) fit to our observed data for our observed slope statistic. We also fit the same linear model to each of the shuffled data sets (n=1000), and compared the observed slope statistic with the null distribution of these slope values. If the observed β exceeded the top 95% of the null distribution, we considered the linear fit for observed data to be significantly different from the shuffled data.

6.4.11. EEG pre-processing

After data collection, noisy channels were identified using a modified version of the PREP pipeline (Bigdely-Shamlo, Mullen, Kothe, Su, & Robbins, 2015). We omitted the bad-by-RANSAC criterion that identifies correlated channel groups which deviate from other channels. This was necessary as frequency-tagging elicits responses in localised, often highly correlated channel clusters. Bad channels were then spherically interpolated (channels rejected per participant M = 5.70, SD = 0.62). After channel rejection and interpolation, whole-trial EEG data were re-referenced to the average of all electrodes, demeaned using the whole-trial average and linearly

detrended, before being down-sampled to 250 Hz. We then applied a Laplacian transform to improve the spatial selectivity of our data. Figure 6.3 shows the whole trial SNR spectrum from electrode POz, as well as topographical distribution of frequency-tagged components.

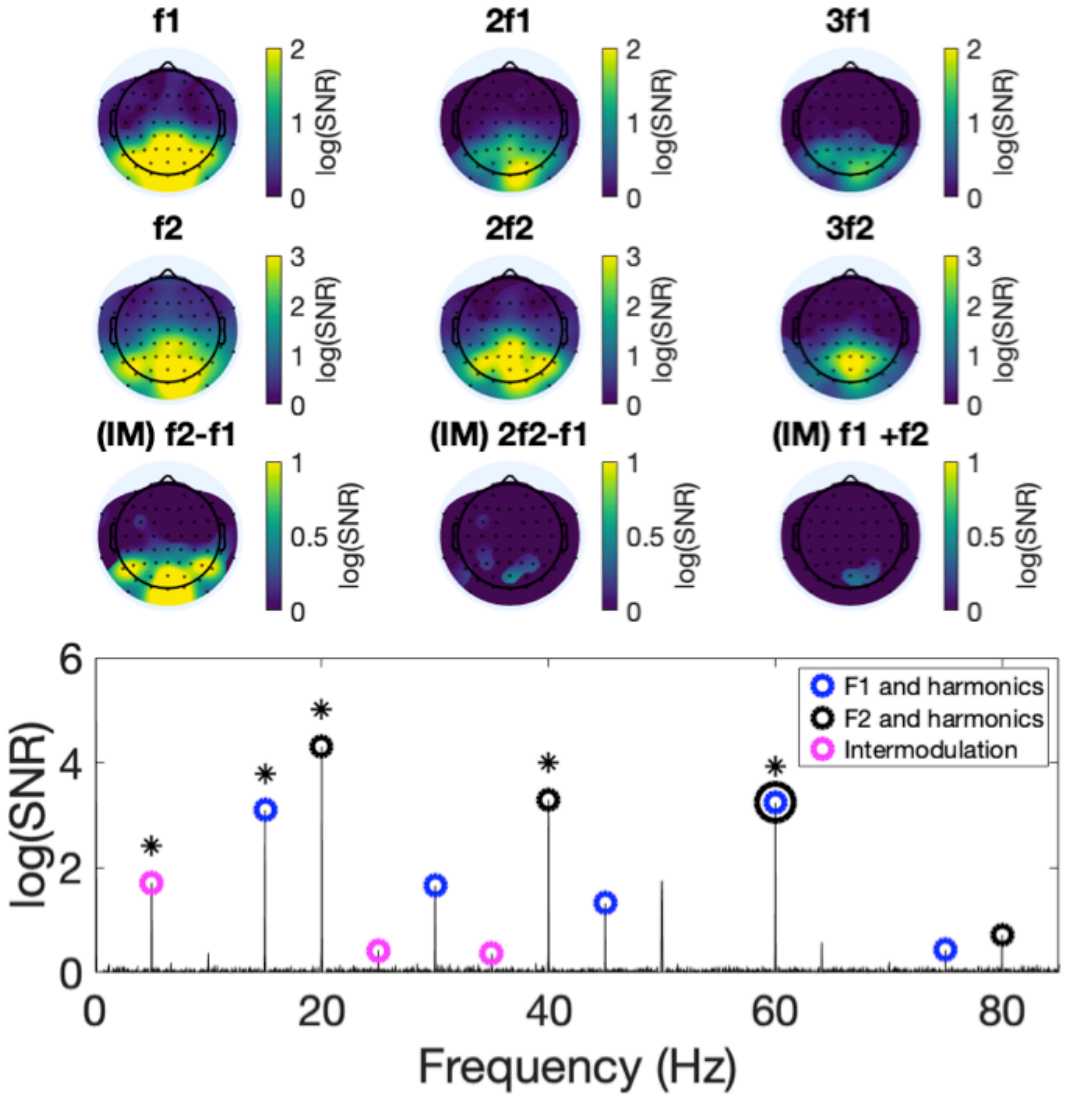

Figure 6.3. Average SSVEP responses to target and background flicker in our paradigm.

The mean log(SNR) spectrum (at POz) over all participants and periods of PFI (60 second whole-trials; hbw = .017 Hz). Target flicker and harmonics are shown in blue, background flicker and harmonics are shown in black, with intermodulation components shown in magenta. The topographic distribution of SNR across participants is shown above the SNR spectrum at each frequency. In topoplots, non-significant electrodes are masked. Asterisks mark log(SNR) significantly different to 0 in spectrum at POz, FDR-adjusted across all frequencies to $p < .05$.

6.4.12. SSVEP analysis via rhythmic entrainment source separation (RESS)

SSVEP topography can vary based on individual participant anatomy, the entrained neural network (Ding et al., 2006) as well as based on the frequency of flicker selected (as shown in the above figure). As such, we applied a spatiotemporal filter called rhythmic entrainment source separation (RESS), to reduce the distributed topographical response at SSVEP frequencies to a single component time-series (Cohen & Gulbinaite, 2017). The RESS single component is a weighted average from across all channels, which can be analysed in the time-frequency domain instead of relying upon the selection of a single or multiple channel based on post-hoc data inspection (Cohen & Gulbinaite, 2017). Specifically, RESS functions by creating linear spatial filters tailored to maximally differentiate the covariance between a signal flicker frequency and neighbourhood frequencies, thereby increasing the signal-to-noise ratio at the flicker frequency. We constructed RESS spatial filters from 64-channel EEG, by extracting signal data following a narrow-band filter via frequency-domain gaussian, centred at flicker frequencies (full-width at half maximum, the SD of filter = 1 Hz). As the frequency-neighbourhood across different signals would contain different amounts of simultaneous flicker, we proceeded by selecting broadband neural activity to construct reference covariance matrices. Comparing signal to broadband activity has previously been shown to allow the reconstruction of SSVEP signals using RESS (Cohen & Gulbinaite, 2017; **Chapter 5**).

RESS spatial filters were constructed per participant per frequency (Target flicker and harmonics; 15, 30, 45 Hz, Background flicker and harmonics 20, 40, 60 Hz; Intermodulation components; 5, 25, 35 Hz), using epoched data from all time-windows -3000 to -100ms and 100ms to 3000ms around button press/release, avoiding catch periods. Each filter was fitted without distinguishing whether targets were disappearing or reappearing due to button press or release, in order to reduce the possibility of overfitting these condition comparisons. After application of the RESS spatial filters, we reconstructed the time course of SSVEP log(SNR) from the RESS component time-series as described above (**SSVEP SNR calculation)**. With RESS, we were able to focus our analysis on a single component time-series per frequency of interest, without arbitrarily selecting a single channel or averaging channels, eliminating the need for corrections for multiple comparisons across channels.

6.4.13. Event-by-Event image analysis of button press and SSVEP-SNR

We performed image-based event-by-event analysis (Fujiwara et al., 2017), to investigate whether the amount of PFI reported may reflect changes in log(SNR). This image-based analysis is necessary due to variations in the frequency and duration of reported PFI per participant.

To accommodate these differences, we first sorted all PFI events in descending order, based on the sum of buttons pressed over a three second period. We used the period [0, +3] relative to button press for PFI disappearances, and [-3, 0] relative to button release for PFI reappearances. This integral of the number of buttons pressed we term "the amount of PFI". After sorting based on the amount of PFI, we then resampled participant data along the trial dimension (y-axis) to normalize trial counts for each participant. Participant data was then smoothed along this normalized trial dimension and averaged across participants. This process was repeated for the event-by-event time course of log(SNR), except the order of trials was predetermined by the corresponding button-press per participant. A schematic pipeline for this entire procedure is displayed in Figure 6.4. Finally, to quantify whether changes in log(SNR) occur with an increasing amount of PFI, we grouped trials when the amount of PFI was between 0 and 1, 1 and 2, 2 and 3 or greater than 3.

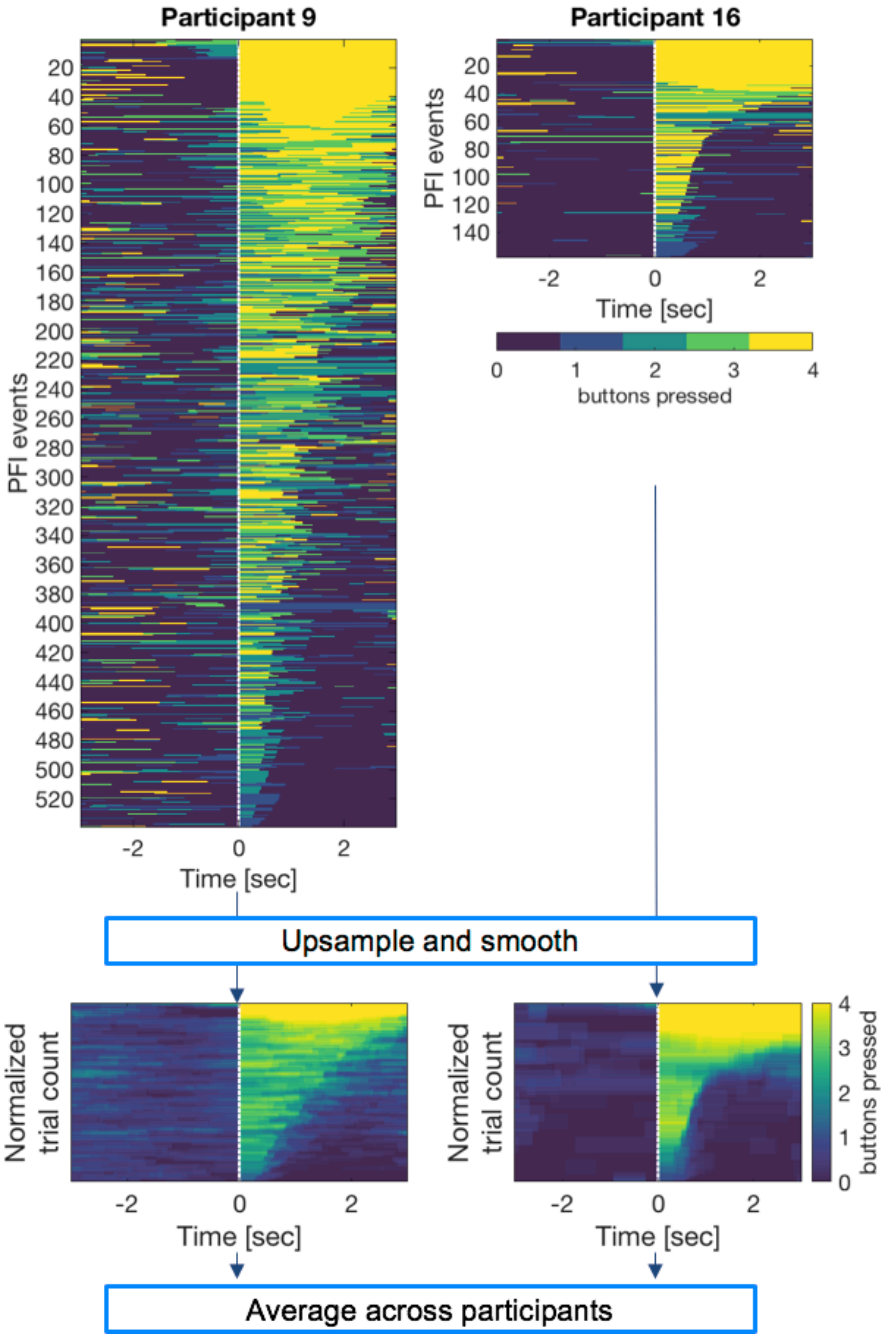

Figure 6.4. Pre-processing for event-by-event based image analyses.

PFI events were first sorted according to "the amount of PFI" (= the integral of the number of the buttons pressed over 3 seconds) occurring after button-press or before button release. Each image along the y-axis was then resampled to normalize the number of trials from 1 to 100 samples. An 8-sample moving average was then applied to smooth each image along the normalized trial-dimension, before averaging across participants. The same process was also applied to RESS log(SNR) after sorting by the amount of PFI per trial based on button press (or release). This image-based analysis enables us to compare PFI dynamics despite differences in the number of PFI events per participant.

6.4.14. Statistical analysis – EEG

To assess the significance of SSVEP peaks in the EEG spectra, we corrected for multiple comparisons with a False Discovery Rate (FDR) of .05 (Benjamini et al, 2006). We corrected for multiple comparisons in SNR time-series using non-parametric (temporal) cluster-based corrections (Davidson et al., 2018; Maris & Oostenveld, 2007). Specifically, we first detected any temporally contiguous cluster by defining a significant time point as $p < .05$ uncorrected (Maris & Oostenveld, 2007). Then, we concatenated the contiguous temporal timepoints with $p < .05$ and obtained a summed cluster-level test statistic for the cluster. The sum of observed test-statistics (e.g., *t* scores) in a temporally contiguous cluster were then retained for comparison with a permutation-based null distribution. To create the null distribution, we repeated the procedure of searching for and retaining contiguous time-points which satisfied the $p < .05$ (uncorrected) cluster criterion, after first shuffling the condition labels 2000 times. For within-subject comparisons, this amounts to randomly permuting the averages for each condition within each subject. From each of the 2000 repetitions, we obtained the maximum sum of cluster-level test statistics, which served as a null distribution. We regarded the effect to be significant if the original summed cluster-level statistics exceeded the top 95% of the null distribution of the summed statistics (as $p_{cluster} < .05$).

6.5. Results

6.5.1. Overview

We frequency-tagged both target and background stimuli during multi-target PFI. Previously, our attempts to entrain neural activity in response to background flicker have succeeded, yet we were unable to recover the target SSVEPs (**Chapter 5**) In the current report both target and background SSVEPs were observable in the across-participant EEG spectra at the highest frequency resolution (60 second time window; Figure 6.3). In addition, an intermodulation component (f2 - f1) - proposed as an index of neural integration (Gordon et al., 2017; Zhang et al., 2011) - was statistically significant. We focus our remaining analysis on these three key features,

and investigate the SNR of Target, Background, and Intermodulation component (IM) RESS log(SNR) during both PFI and phenomenally matched replay (catch periods).

In sum, we show that during catch periods, target and background SNR time-series are consistent with previous literature; increasing/decreasing when physically present/absent from the visual display (Figure 6.5). IMs peak when both Target and Background SNR are maximal, when targets are returned to the visual display. By contrast, during PFI, the same pattern of results was not observed. Both target and background SNR are shown to increase during PFI disappearance, with IM SNR peaking immediately prior to the subjective report of PFI. PFI is also preceded by a transient increase in the spatiotemporal correlation between IMs and Target flicker, as well as IMs and Background flicker. We interpret this increase as support for PFI being mediated by interactions, marked by IMs, between neural populations representing both target regions and their surrounds.

6.5.2. Target and Background SSVEP strength tracks stimulus visibility during Catch Periods

We first visualized whether the number of targets removed during catch periods would affect changes in the log(SNR) of Target, Background, and IM responses. To perform this analysis, we first sorted all catch periods by the cumulative number of buttons pressed, which we term the ‘amount of catch’ (Figure 6.5a; see **Methods**).

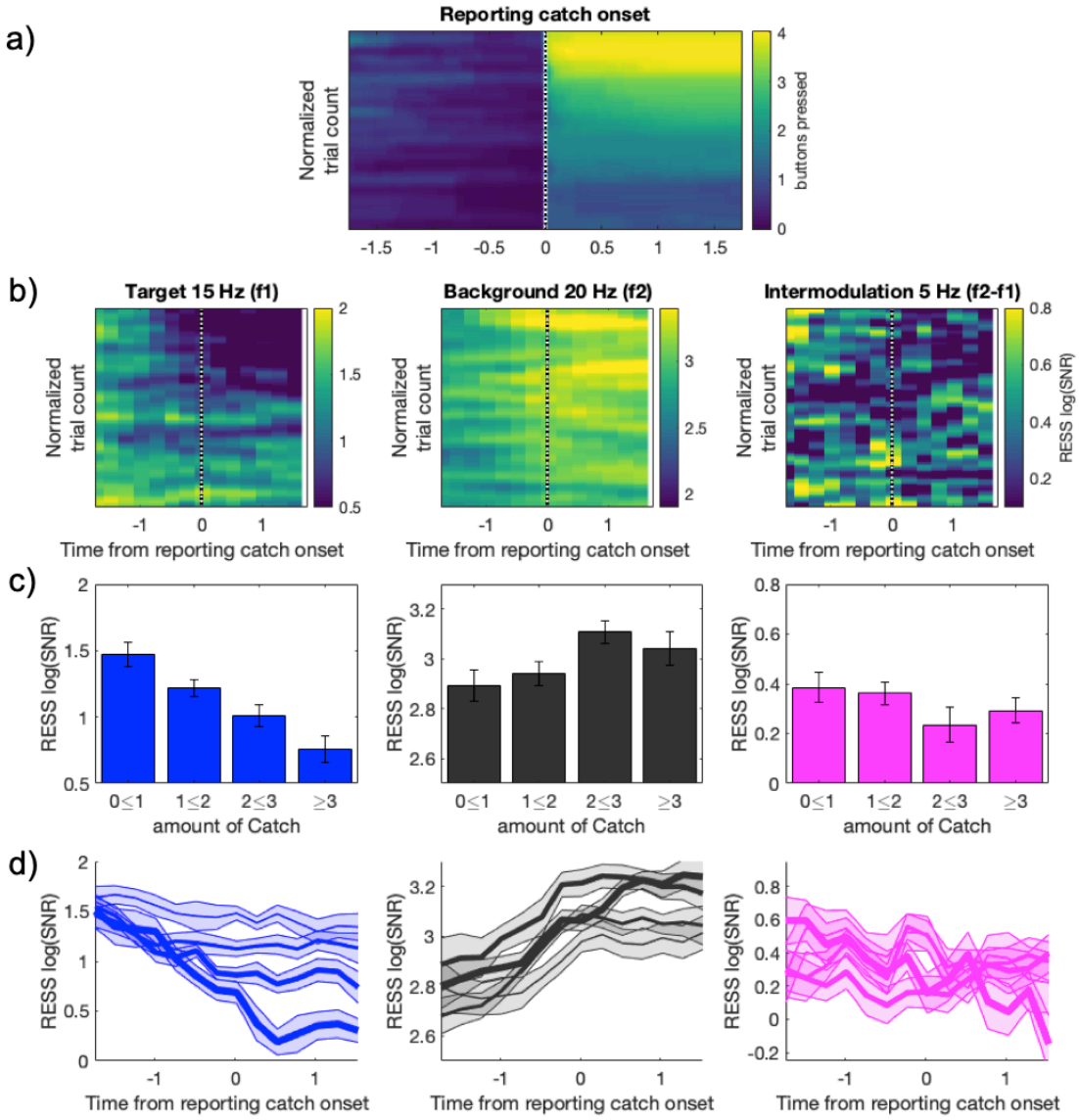

Figure 6.5. Event-by-event image analysis for the effect of the number of targets removed at catch onset on button press and RESS log(SNR).

a) Event-by-event image analysis of button press when participants report the physical removal of targets at catch onset. b) event-by-event image analysis of RESS log(SNR) after sorting based on the amount of catch per trial, per participant. Target responses at f1 are shown in the left column, Background responses in the central column, and Intermodulation response on the right. c) Bar graphs for the mean RESS log(SNR) over all time-points as a function of the amount of catch. d) The time course of RESS log(SNR) around catch disappearance separated by the amount of catch. Error bars in c) and shading for d) indicate 1 SEM across participants (adjusted for within-participant subject comparisons, Cousineau, 2005).

After sorting by the amount of catch across participants, both RESS log(SNR) at Target (f1) and Background (f2) flicker frequencies appear to be affected by an

increasing amount of catch (Figure 6.5b). Target RESS log(SNR) decreases as an increasing number of flickering targets are reported as absent at catch onset, while Background RESS log(SNR) increases. To quantitatively investigate these differences, we performed LME analyses on the amount of catch with random intercepts per participant, followed by likelihood ratio tests. We found a significant linear effect for the amount of catch on both Target RESS log(SNR) ($\chi^2(1) = 23.66$, $p = 1.5 \times 10^{-6}$), and Background RESS log(SNR) ($\chi^2(1) = 23.66$, $p = .043$). By contrast, the Intermodulation component, which is driven by both Target and Background flicker, was not significantly affected by the amount of catch ($p = .188$).

We next applied the same analysis to the time-period when participants were reporting on the return of flickering targets at catch offset. In general, we observed weaker effects for the amount of catch on RESS log(SNR) when reporting on the return of targets at catch offset. A significant effect was again found for the amount of catch on Target RESS log(SNR) ($\chi^2(1) = 16.92$, $p = 3.90 \times 10^{-5}$), however the amount of targets returning did not significantly modulate the Background ($p = .424$) or Intermodulation ($p = .117$; Figure 6.6) responses. We return to this difference between target removal and the return of targets in our Discussion.

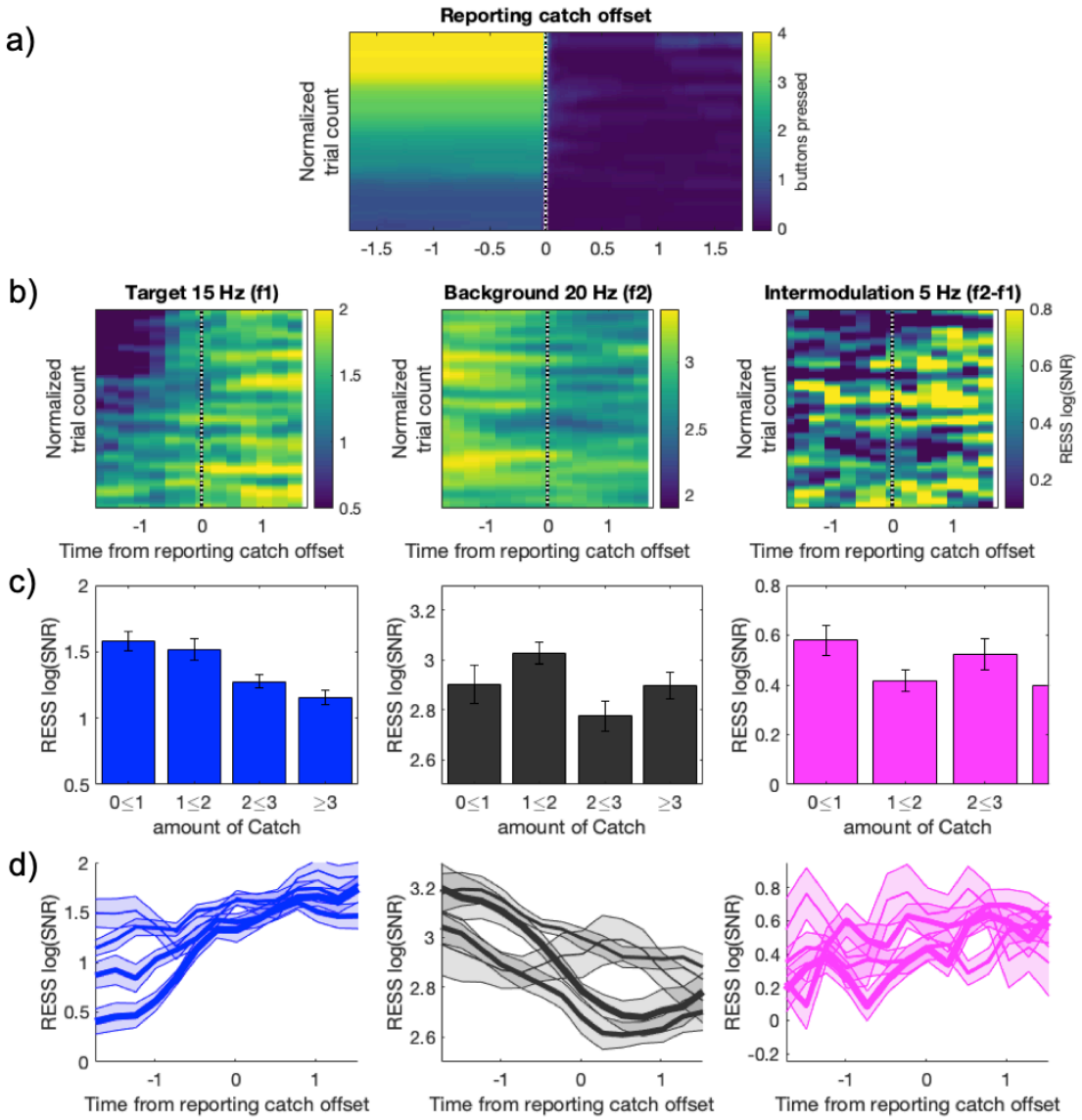

Figure 6.6. Trial by trial image analysis for the effect of the amount of targets removed at catch offset on button press and RESS log(SNR).

All other conventions as in Figure 6.5.

Overall, these results are consistent with the prevailing SSVEP literature which shows that an increase in the saliency of flickering stimuli increases SNR of frequency-tagged neural responses. Here, we have shown that the replacement of flickering Target stimuli with flickering Background regions results in parametric changes in SNR at catch onset. Next, we investigated whether the same pattern of results would be observed when flickering target regions remain on screen, yet disappear from conscious awareness due to PFI.

6.5.3. An opposite neural signature for Target-specific responses during PFI

Compared to catch periods, periods of PFI are distinct, as the presence of a 15 Hz flickering target remains unchanged. When reporting on PFI, participants experience a change in the colour of target regions, switching from purple to match the surround black and white of the background flicker. We performed the same event-by-event image-based analysis to quantify the effect that an increasing amount of PFI may have on the RESS log(SNR) of Target, Background, and Intermodulation frequencies.

Figure 6.7 displays the results of this analysis. Most strikingly, unlike the removal of targets during catch periods which decreased the strength of Target-related responses, Target RESS log(SNR) strength now significantly *increased* as an increasing number of targets disappeared during PFI ($\chi^2(1) = 10.99$, $p = 9.16 \times 10^{-4}$). This result runs counter to the intuitive assumption that with a decrease in the visibility of flickering target stimuli, a decrease in RESS log(SNR) should also be observed. Like during catch periods, a significant linear effect of the amount of PFI was also observed for the Background RESS log(SNR) ($\chi^2(1) = 11.705$, $p = 6.23 \times 10^{-4}$), as well as for the Intermodulation, which approached significance ($p = .053$).

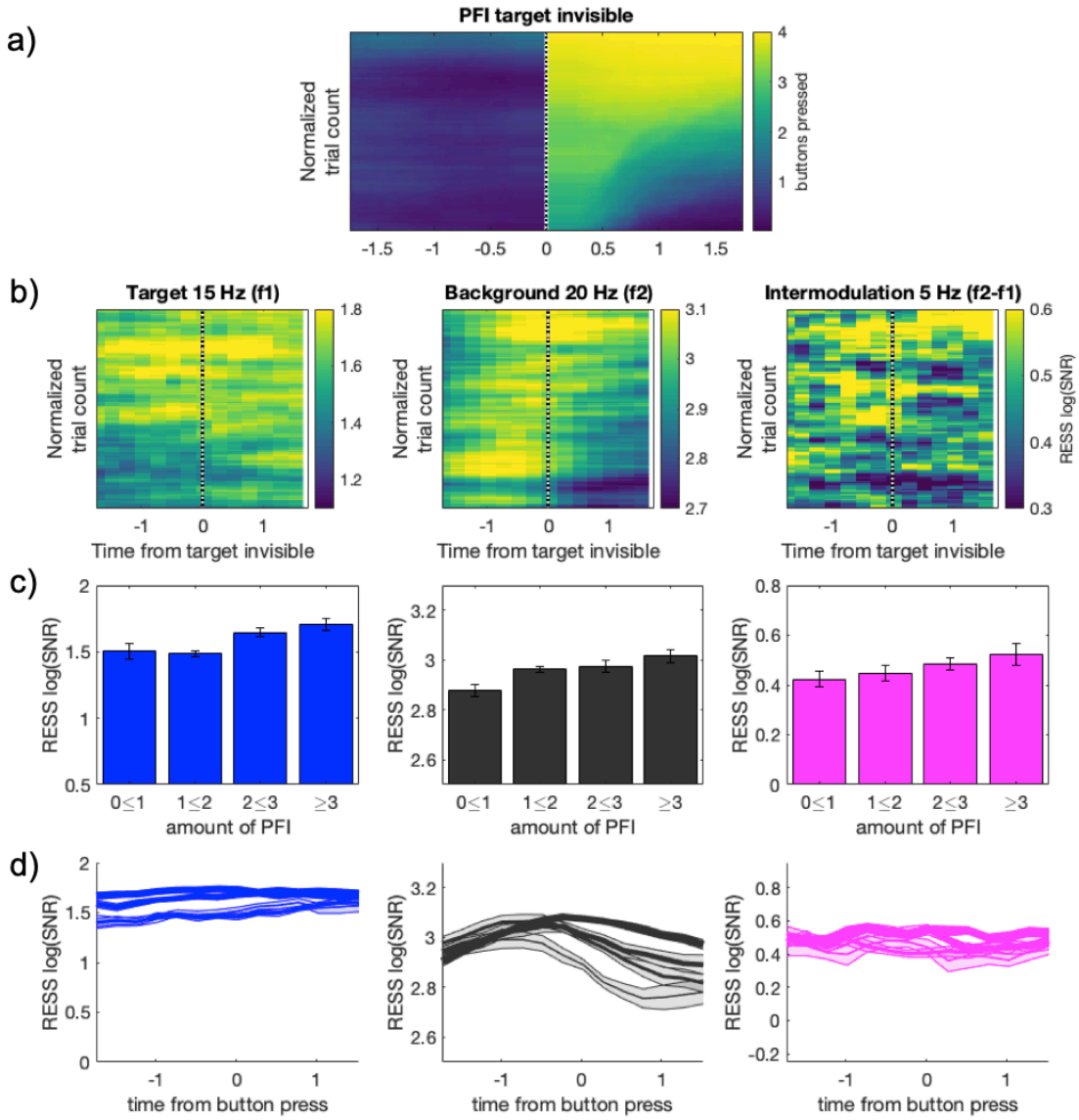

Figure 6.7. Event-by-event image analysis for the effect of the amount of targets PFI on RESS log(SNR) during Target disappearance.

a) event-by-event image analysis of button press when participants report the perceptual disappearance of targets at catch onset. All PFI events are organised so that trails with the greatest number of disappearing targets appear at the top, and those with fewer disappearing targets appear at the bottom. b) Event-by-event image analysis of RESS log(SNR) after sorting based on the amount of PFI per trial in a), per participant. Target responses at f1 are shown in the left column, Background responses in the central column, and Intermodulation response on the right. c) Bar graphs for the mean RESS log(SNR) over all time-points as a function of the amount of PFI. d) The time course of RESS log(SNR) separated by the amount of PFI. Error bars in c) and shading for d) indicate 1 SEM across participants (adjusted for within-participant subject comparisons, Cousineau, 2005).

We also investigated the effect that the amount of PFI would have on RESS log(SNR) at PFI reappearance, when targets are subjectively returning to their

original colour. Similar to our analysis of the amount of catch at catch offset, we observed a reduced strength for the effect of the amount of PFI at PFI reappearance. Neither Target, Background, or Intermodulation frequencies displayed a significant linear effect ($p = .229$, $p = .671$, $p = .576$, respectively). The results of this analysis are displayed in Figure 6.8.

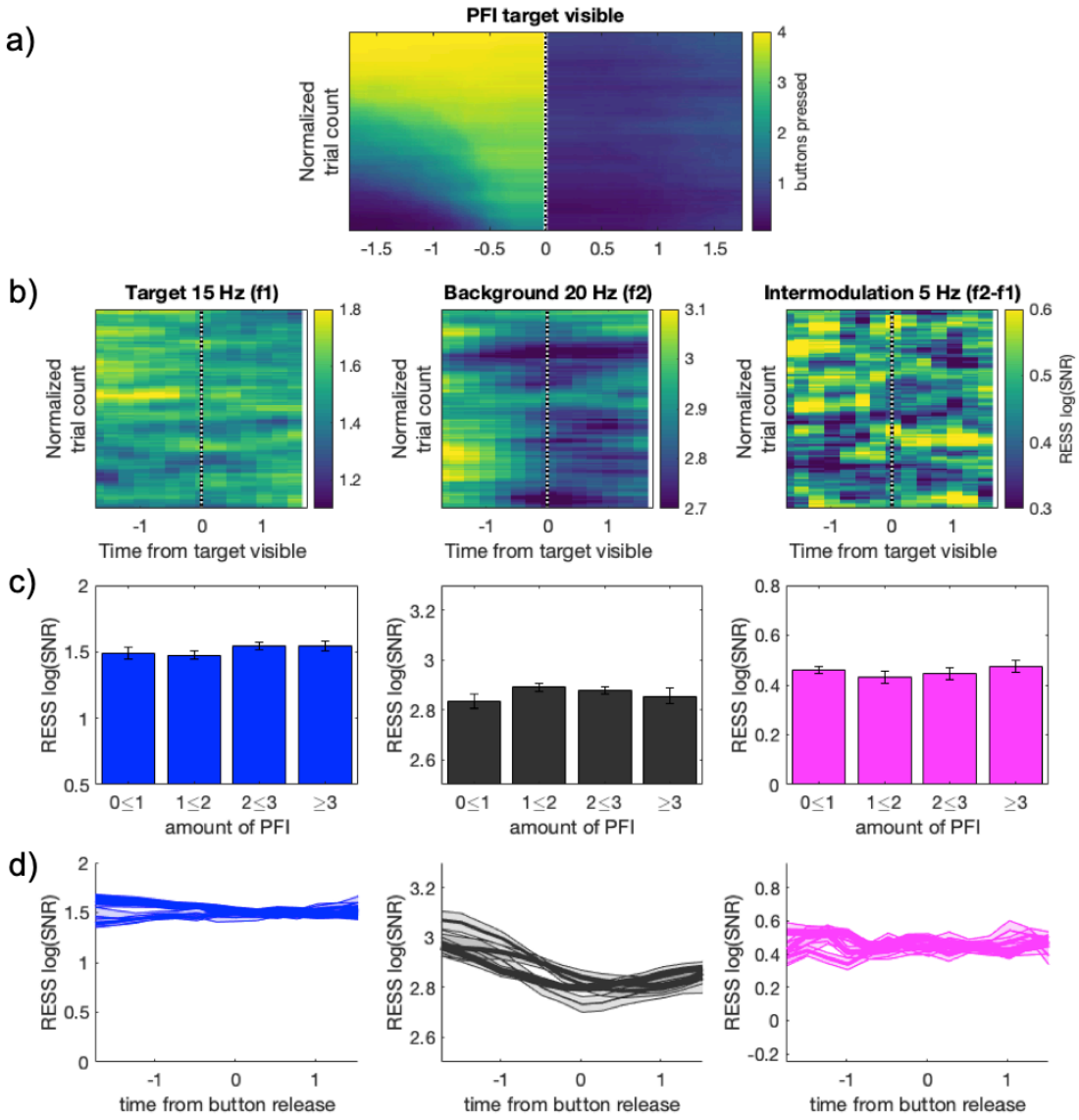

Figure 6.8. Trial by trial image analysis for the effect of the amount of targets PFI on RESS log(SNR) during target reappearance.

a) Event-by-event image analysis of button press when participants report the physical removal of targets at catch onset. All PFI events are organised so that the greatest number of disappearing targets appear at the top, with fewer disappearing targets at the bottom. b) Event-by-event image analysis of RESS log(SNR) after sorting based on the amount of catch per trial in a), per participant. Target responses at f1 are shown in the left column, Background responses in the central column, and Intermodulation response on the right. c) Bar graphs for the mean RESS log(SNR) over all time-points as a function of the amount of PFI. d) The time course of RESS log(SNR) separated by the amount of PFI. Error bars in c) and shading for d) indicate 1 SEM across participants (adjusted for within-participant subject comparisons, Cousineau, 2005).

6.5.4. A spatial-interaction between target locations increases PFI duration

We have so far revealed a distinct neural correlate for endogenously generated target disappearance compared to physical target removal, showing an *increase* in target specific neural activity with a decrease in visibility. We next investigated behavioural responses during PFI, and the effect of the number of targets simultaneously invisible.

We compared the amount of PFI per trial, duration per PFI, and total duration of PFI during either 0, 1, 2 3, or 4 target PFI periods (Figure 6.9, blue bars). We found that while zero target PFI periods were most frequent (occurring 6 times per trial), when targets did disappear, simultaneous 4 target disappearances were the most common (4 times per trial, compared to <2 times per trial for 1, 2, and 3 target disappearances). This interesting trend continued for both the duration per PFI (7 s per 0 target, <2 s for 1-3 targets, 5.5 s for 4 targets) and total duration of PFI (32 s for 0, <2 for 1-3, and 20 s for 4, respectively), showing that 4 target disappearances were the most common, disappeared for a longer duration, and greatest total duration per trial. We tested these trends regarding the number of targets filled-in (nPFI; 1, 2, 3, or 4, removing 0) formally by LME analysis, and found significant linear effects for all measures (PFI per trial; $\chi^2(3)= 58.61$, $p = 1.16 \times 10^{-12}$; PFI duration; $\chi^2(3)= 54.58$, $p = 8.43 \times 10^{-12}$; and Total duration; $\chi^2(3) = 70.56$, $p = 3.22 \times 10^{-15}$).

To investigate whether these trends obtained over four-target locations were likely to occur by chance we performed a shuffling analysis (see methods, and Figure 6.9b) and recalculated PFI characteristics (PFI per trial, PFI duration, and total duration), as a function of nPFI.

Figure 6.9 (c-h, grey bars), display the results of this analysis, displaying the mean across all 1000 shuffled sets of data. In contrast to observed data, the shuffled data showed PFI for 1, 2, and 3 targets being more common (6, 7, and 6 times per trial respectively), than for zero and four targets (each occurring less than 4 times per trial). For PFI duration, especially durations for 4 target PFI were shorter in the shuffled data. Strikingly, and in direct opposition to our observed data, the total duration of 0, 1, 2, and 3 target PFI per trial was roughly equivalent in shuffled data (each ~14 s duration), with 4 target PFI occurring for the least amount of time (<10 s).

To statistically evaluate the differences in these trends, we compared the slope of our observed data to all the slopes in our null distribution for shuffled data (from 1-3 nPFI, excluding 0). For all PFI measures, the observed slope was outside the 95 % of the null distribution (corresponding to p <.05). For the PFI per trial and total duration per trial, the positive slope for the effect of nPFI in our observed data is in direct opposition to the distribution of negative slopes in our shuffled data. As a result, the observe positive trend for increasing PFI per trial, duration per PFI and total duration of PFI with an increasing number of PFI targets cannot be expected due to chance, and represents a synergistic spatial-interaction between multiple PFI targets.

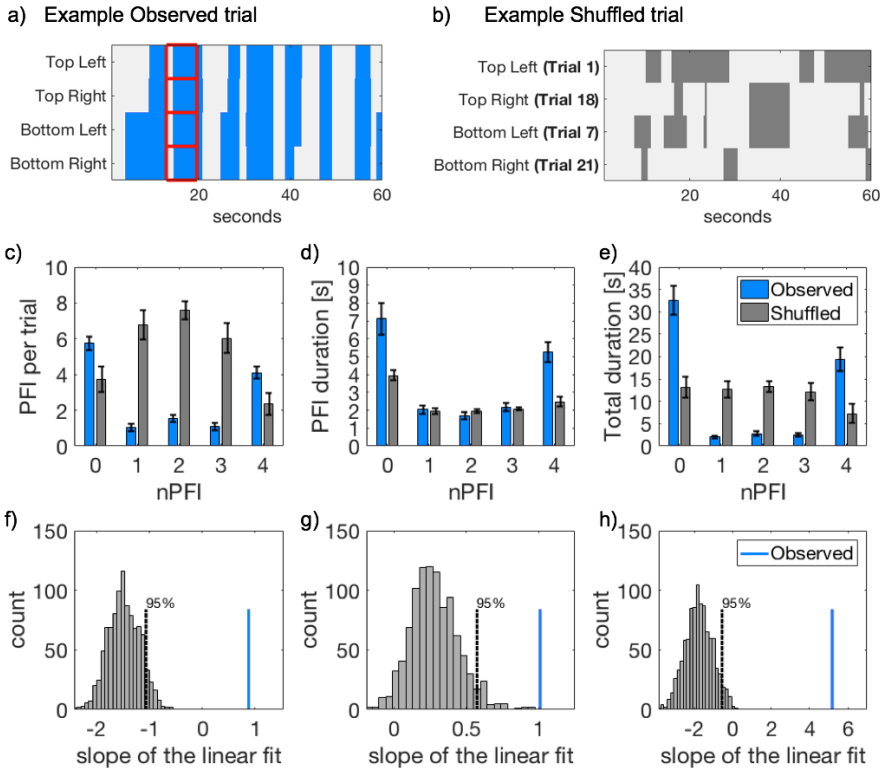

Figure 6.9. Behavioural data comparing PFI characteristics based on nPFI.

a) Example PFI data from one participant, displaying synergistic PFI across multiple locations. b) Example of Shuffled data to test whether synergistic PFI occurs by chance. c) Instances of PFI per trial, d) mean duration per PFI, and e) total duration of PFI as a function of the number of targets reported as simultaneously invisible (nPFI). All panels display both observed (blue) and shuffled (grey) data. For the observed data, error bars represent 1 SEM, corrected for within-participant comparisons (Cousineau, 2005). For the shuffled data, we first computed the SEM within each shuffled data set across participants. Then, as the error bar, we show the mean of the SEM across 1000 shuffled sets. f-h) Slope of the linear fit for each of the PFI variables in c-e as a function of nPFI for observed (blue line) vs shuffled data (1000 sets, grey histogram).

6.5.5. An overall reduced response magnitude during PFI compared to Catch periods

While our previous analysis has shown that the amount of Target disappearances during both catch onset and PFI impact on the amount of change to log(SNR), we also wanted to investigate the time-course of SNR based on the direction of these perceptual changes - i.e., comparing Target disappearance to reappearance. To investigate these differences, we superimposed the SNR time-series for disappearance (Figure 6.10, dotted lines) and reappearance (solid lines), and used temporal cluster-based paired *t*-tests to determine if and when SNR strength might differentiate between these cases. As well as illustrate whether a target disappearing/reappearing would be reflected in the log(SNR), we were also motivated to compare the magnitude of these responses during PFI and Catch periods. If PFI can be distinguished from Catch periods, then the neural correlates of endogenously generated perceptual changes can be investigated to complement efforts using MIB (Donner et al., 2008) and binocular rivalry (Brascamp et al., 2018; Lumer et al., 1998).

Overall, responses in terms of SNR magnitude were reduced, and temporally advanced, during PFI compared catch periods (Figure 6.10). For Target SNR during PFI, as detailed above, RESS log(SNR) unexpectedly increased during target disappearance compared to target reappearance. This effect was significant from -0.74 s before, through to 1.53 s after subjective report ($p_{cluster}$ <.001). By contrast, the Target SNR-time series during catch periods closely followed the visibility of target stimuli, decreasing during target disappearance and increasing during reappearance, respectively. These differences were significant -1.75 to -.99 s ($p_{cluster}$ <.001) and from -.49 to 1.53 s ($p_{cluster}$ <.001) relative to subjective report.

Background SNR during both PFI and Catch periods increased when targets disappeared, and decreased when they reappeared, consistent with the SNR reflecting the amount of Background flicker in the stimulus. During PFI, the difference between target disappearance and reappearance was significant from -.99 to 1.28 s around report ($p_{cluster}$ < .001). During Catch periods, these differences began from -1.75 to -1.24 s prior to ($p_{cluster}$ < .001), as well as -0.02 to 1.53 s after subjective report ($p_{cluster}$ < .001).

We also examined the differences in PFI or Catch direction on the Intermodulation (IM) frequency, produced by the neural interaction between Target and Background flicker. A single time-point showed significant differences in IM SNR based on whether targets were disappearing or reappearing during PFI (-0.73s; $t(15) = 2.19$, $p = .045$, *uncorrected*). However, when comparing the IM SNR evoked during the physical removal and return of targets during catch periods, IM SNR was significantly affected from 0.78 to 1.53 s ($p_{cluster} < .001$) *after* subjective report, when flickering targets had been physically returned to the screen.

As well as revealing a difference in the direction of effects (for target and IM responses) and timing of effects (for all responses), we also compared the magnitude of effects during PFI and Catch periods, in a summary time-period 0:1s after report. We chose this window as the average PFI duration for all periods of PFI was greater than 1 second (Figure 6.9), and we could therefore assume a relatively stable percept would be experienced before subsequent PFI events occurred. As well as the Target, Background, and IM frequencies of interest, we additionally explored the harmonics of Target and Background responses that were present in the EEG spectra (Figure 6.3). Figure 6.10d displays a summary of these results, and shows that the pattern of results observed for Target and Background responses persists for their linear harmonics. In general, the size of these effects decrease at higher frequencies, an observation we return to in our Discussion.

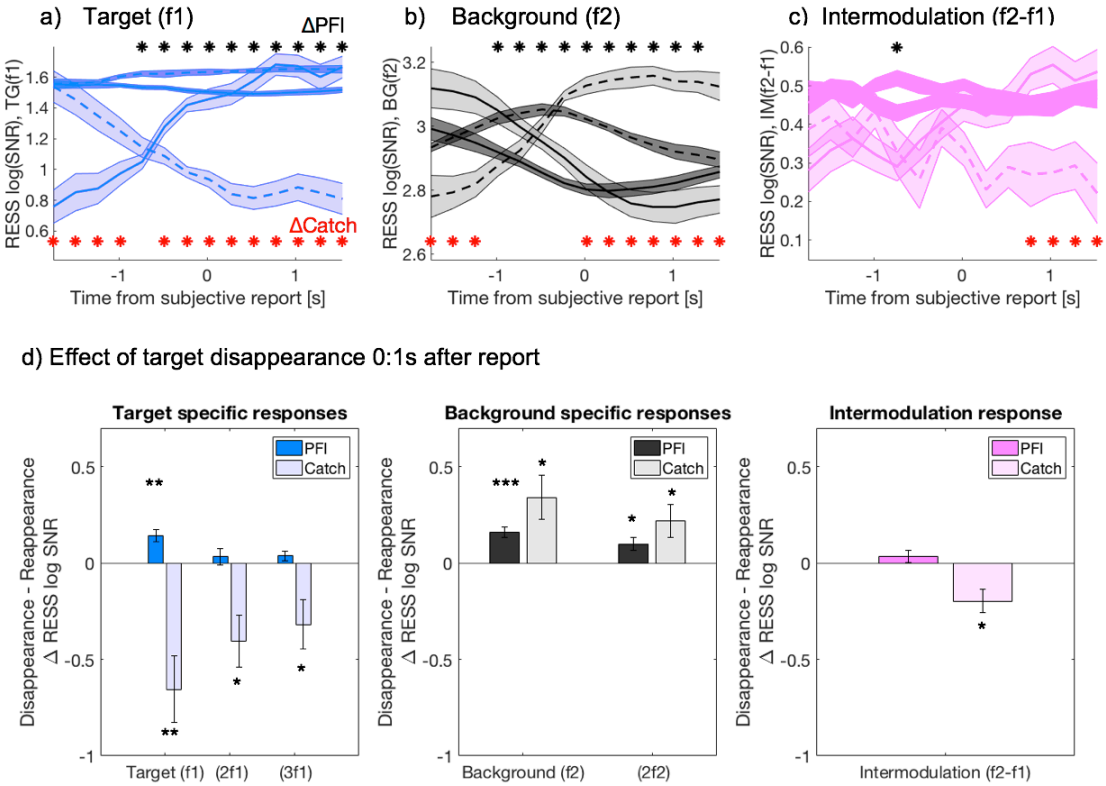

Figure 6.10. PFI vs. Catch RESS log(SNR) time-series. Comparing the direction of evoked changes in RESS log(SNR) during PFI and Catch periods.

*a-c) Separately show the RESS log(SNR) for a) Target (blue), b) Background (grey), and c) Intermodulation frequencies (magenta). In all panels, dotted lines represent target disappearance, solid lines represent target reappearance. Darker shading represents PFI, lighter shading represents catch periods. The significant differences between PFI disappearance and reappearance are marked with black asterisks, and differences between Catch onset and offset are marked with red asterisks. d) Summary of the change in RESS log(SNR) for over the period 0 to 1 s after subjective report, when comparing SNR during target disappearance and reappearance during PFI and catch periods. Note that linear multiples of target flicker (2f1, 3f1) and Background (2f2) responses are displayed in d) that are absent in a-c to show the consistency of these effects. ***p < .001, **p < .01, * p <.05, single sample t-tests compared to zero, FDR corrected.*

6.5.6. An interaction between Target and Background representations during PFI, revealed using steady-state topographical probes (SSTPs)

The importance of interactions between target and background regions during filling-in is a matter of ongoing debate (Weil et al., 2011), yet the absence of these interactions during catch periods - when targets were physically removed from the screen - may have contributed to the discrepancy shown in SNR time-series. For

example, PFI may result from an active process whereby retinotopic cortical representations of a regions surrounds are propagated into the filled-in region via lateral connections (De Weerd et al., 1995; Fiorani Jr., Rosa, Gattass, & Rocha-Miranda, 1992; Weil & Rees, 2011). Alternatively, filled-in regions may be ignored as 'more of the same', with neural changes not occurring at early visual areas during filling-in, yet possibly occurring at higher representational levels (Dennett, 1991; O'Regan, 1992). In contrast, during catch periods, flickering target regions were physically removed and replaced, destroying any interactions between the target and its surrounds.

Changes to SSVEP-topography have been demonstrated to occur with the allocation of attention (Kim et al., 2007; 2011) and changes in task demands (Itthipuripat, Garcia, & Serences, 2013; Silberstein et al., 1995; Silberstein, Harris, Nield, & Pipingas, 2000; Silberstein et al., 1990b). By using SSVEP as a steady-state topographical probe (SSTP; Vialette et al., 2010), we explored the spatial correlation between relative SSTPs for each frequency during PFI. We hypothesized that as the phenomenology of target and background visual features became more similar during PFI onset, that the spatial topography between target and background responses would also increase in similarity. To test this hypothesis, we performed a spatial correlation analysis based on log(SNR) strength across all 64 electrodes, by correlating the SSTPs of Target and Background frequencies over time. For completeness, we also explored the spatial correlation between these frequencies and the IM response. When targets disappeared during PFI, we found a transient increase in the spatial correlation between Target and IM SSTPs from -0.74 before to 0.02s after subjective report (paired *t*-tests, $p_{cluster} < .01$; Figure 6.11a). A sustained increase in SSTP correlation was also observed between the IM and Background log(SNR) (-0.98 to 1.03s; $p_{\text{cluster}} < .001$), yet no change was observed for the correlation between Target and Background SSTPs themselves. The absence of any relative change to the spatial propagation of Target and Background responses, yet significant increase in correlation between IM and other frequencies, supports the active integration of Target and Background representations as a catalyst for PFI. We return to these results in our discussion. Figures 6.11 d-f also display snapshots of these SSTPs over time, for both disappearances and reappearances during PFI. While Target and Background SSTPs remain qualitatively stable, a transient change in IM topography

is present -0.5 s prior to PFI disappearance - extending from midline occipital regions to include bilateral parietal electrodes.

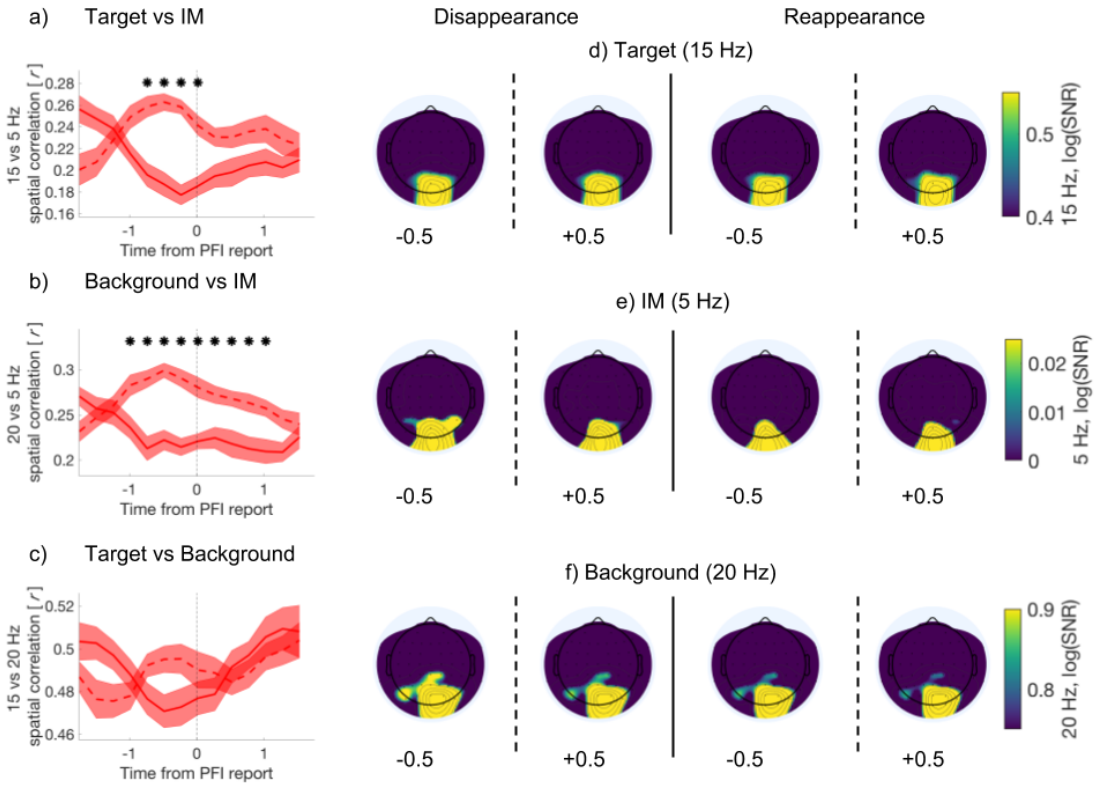

Figure 6.11. Time-course for steady-state topographical probe (SSTP) correlation between Target, Background and Intermodulation (IM) frequencies during PFI.

a-c) Correlation coefficient across 64 electrodes was computed at each time point per participant. The mean time courses of correlation coefficients are shown for target disappearance (dotted), and reappearance (solid) for different frequencies during PFI. We show the mean correlation value obtained after down-sampling PFI events to 48 (the maximum number of catch periods), over 100 repetitions of this down-sampling procedure. a) Correlation between Target and IM during PFI, b) Between Background and IM during PFI, and c) between Target and Background during PFI. Asterisks denote the time-points with significantly different correlation coefficients when comparing disappearance to reappearance (paired t-tests, cluster corrected). Shading reflects the SEM across subjects. d-f) Show SSTPs for Target (d), IM (e) and Background (f) frequencies before and after subjective report for PFI disappearance and reappearance.

6.5.7. An opposite, decrease in correlation strength during catch periods

During catch periods, the SSTPs for Target, Background and IM frequencies did not show the same pattern of results as during PFI. A significant change in the SSTP correlation was observed between all frequencies, yet critically, a transient

increase in spatial correlation was not observed prior to target disappearance, but during catch offset - when targets were physically replaced on screen (Figure 6.12). These changes were significant for Target and IM frequencies from 1.28 to 1.53 s after catch report (paired samples t-tests, $p_{cluster} < .05$), from -0.23 to 1.28s after catch report for Background and IM frequencies ($p_{cluster} < .01$) and from -0.23 to 0.52s ($p_{\text{cluster}} < .01$) around catch report for Target and Background frequencies.

Unlike during PFI, the physical removal and replacement of flickering targets on screen resulted in transient decreases and increases in the spread of frequency-tagged topography, with background SSTPs showing the opposite pattern. The SSTP for the IM frequency increased from primarily midline occipital sites to include more widespread bilateral parietal regions when subjects have reported on the return of flickering targets at catch offset.

Side by side, these differences in SSTP correlations between PFI and catch periods reveal that a markedly different propagation of target, background, and IM responses accompanies each reported change in awareness. Only PFI was preceded by an increase in the correlation between IM and target/background frequencies, which we view in support of accounts that an active integration between target-and-background neural representations is a prerequisite to PFI.

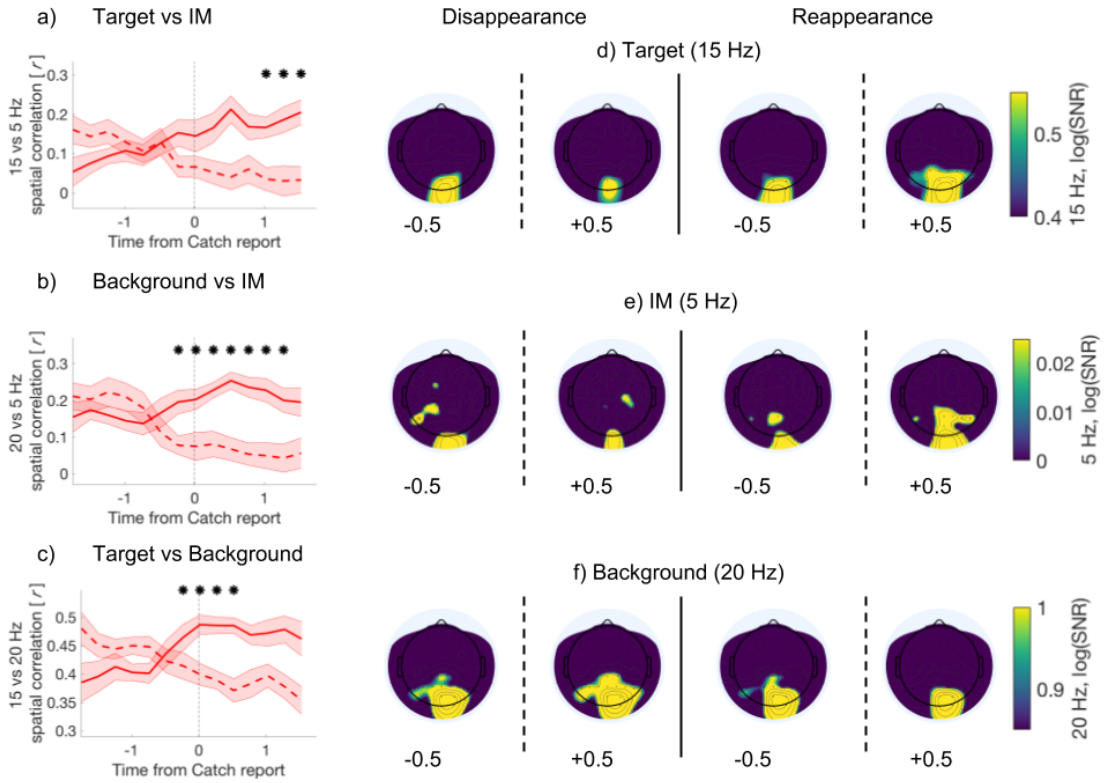

Figure 6.12. Time-course for steady-state topographical probe (SSTP) correlation between Target, Background and Intermodulation (IM) frequencies during Catch periods.

a-c) Correlation coefficient across 64 electrodes was computed at each time point per participant. The mean time courses of correlation coefficients are shown for catch onset, when targets were physically removed (dotted), and their return at catch offset (solid). a) Correlation between Target and IM, b) Between Background and IM, and c) between Target and Background SSTPs during catch periods. Asterisks denote the time-points with significantly different correlation coefficients when comparing catch onset to offset (paired t-tests, cluster corrected). Shading reflects the SEM across subjects. d-f) Show SSTPs for Target (d), IM (e) and Background (f) frequencies before and after catch report.

6.6. Discussion

We investigated multi-target perceptual-filling in (PFI) in the EEG by flickering both targets (f1) and their surrounds (f2). This simultaneous flicker evoked strong SSVEP responses at target and background frequencies, their harmonics (linear multiples of f1 and f2), and an intermodulation component (f2-f1). During PFI, we observed an increase in frequency-tagged responses to flickering targets when they became invisible. The spatiotemporal profile of these frequencies was also shown to

be distinct from the phenomenally matched replay when targets were physically removed from the screen (catch periods). Here, we focus our discussion on the implications of these results, and implications for the relevance target-background interactions in PFI dynamics.

6.6.1. An opposite neural signature for Target SSVEPs during PFI and Catch periods

During catch periods, we observed that as 1- 4 flickering targets were removed from the display, the SNR of Target SSVEPs linearly decreased. At the same time, as the flickering target regions were replaced with background flicker, a linear increase in the SNR of Background SSVEPs was also observed. These results demonstrate that the physical alteration of flickering targets was strong enough to evoke changes in the SNR time-series of SSVEP responses.

During PFI, when flickering targets remain on screen yet subjectively disappear from awareness, a different pattern of results was obtained. As an increasing number of targets became invisible, the SNR of target SSVEPs linearly increased. This counterintuitive result positions frequency-tagged PFI as a useful resource to dissociate the mechanisms of awareness from attention, as here attending to an invisible target increased SNR strength. While we did not directly manipulate attention, we ensured that all the trials we retained were indicative of full attention on task by removing any containing failed catch periods. In our experiment, when covertly attending to targets and responding on their disappearance, target responses increased. This is in direct opposition to the result we obtained during catch periods, as well as the decrease in SNR strength which normally accompanies the disappearance of flickering stimuli (Brown & Norcia, 1997; Jamison, Roy, He, Engel, & He, 2015b; Katyal et al., 2016; Tononi, Srinivasan, et al., 1998b; Zhang et al., 2011).

Previous research has shown that spatial attention can enhance the processing of subliminal visual stimuli which do not enter conscious awareness (Bahrami et al., 2008a; Wyart, Dehaene, & Tallon-Baudry, 2012; Wyart & Tallon-Baudry, 2008). To isolate these phenomena, stimuli were either presented at detection threshold (Dehaene et al., 2006; Schurger, Cowey, Cohen, Treisman, & Tallon-Baudry, 2008; Wyart et al., 2012), or remained completely invisible (Smout & Mattingley, 2018).

Our results complement and extend this body of research, as our participants continuously reported on subjective visibility, rather than gave a retrospective classification that a brief stimulus was either visible/invisible. As a consequence, we have provided the first example for effects of attention during continuous changes in subjective awareness, under conditions of reported invisibility. More specifically, our findings demonstrate that when participants attend to peripheral targets, the neural representation of a stimulus increases as it leaves the contents of conscious awareness.

Intriguingly, one previous study found target SSVEP strength to be reduced during PFI (Weil et al., 2007). There are a number of features that could explain this discrepancy with our data. Weil et al (2007) used MEG to measure target-specific SSVEPs from a single target flickering at 7.5 Hz, located in the lower-left quadrant over a dynamic random noise background. Their goal was to examine activity entrained to the target in retinotopic areas of the visual cortex, and an above-baseline SSVEP was used as evidence of reduced target-specific responses during PFI. In contrast to our findings, Weil et al.'s target was located at greater eccentricity (9.5° compared to max 5.5° diagonally from centre), and was presented alone, with attention maintained at the central fixation. We propose that any effects of attention would have been reduced in the study of Weil et al (2007) due to this increase in target eccentricity. Additionally, there are no shared features among multiple-targets, which would increase the occurrence of PFI when attended (see below; also, De Weerd et al., 2006; Lou 1999). In our study, participants were explicitly told to covertly distribute their attention among all targets to monitor for any instance of PFI. It is also noteworthy that the opposite direction for target-specific responses was not mirrored in background SNR strength during PFI. During both PFI and Catch periods, Background SNR linearly increased with an increasing number of perceptually invisible targets. As a result, our findings position frequency-tagged PFI as a powerful method in the study of consciousness, allowing the independent influences of attention and consciousness to be tracked. We speculate that target-related activity is indicative of attention, while background related activity reflects the contents of consciousness, a proposition which will need to be tested by future research.

6.6.2. An interaction between Target and Background neural representations mediates PFI

The success of our paradigm relied upon the simultaneous flicker and report of multiple targets during perceptual filling-in. Previously, attending to shared stimulus features such as shape and colour have been shown to increase PFI using multi-target designs (De Weerd et al., 2006; Lou, 1999). By grouping multiple visual targets in our display by their colour, we have extended these results to show that targets also simultaneously disappear for longer than expected by chance (Figure 6.9). We have previously argued that this synergistic effect of PFI may imply the perceptual grouping of invisible stimuli, or the result of increased attentional allocation to the background surrounding all targets (**Chapter 5,** see also **Supplementary Material**). While we have not explicitly compared the data in these cases, the strength of this synergistic effect is far stronger in the present data, which we note could be due to an enhanced grouping of PFI targets due to their shared frequency (Alais et al., 1998), larger size and/or closer location (cf. Wagemans et al., 2012).

One theoretical mechanism for filling-in proposes that active processes of inhibition and excitation in early visual networks mediate target disappearance (Sakaguchi, 2001, 2006; Weil & Rees, 2011). Due to adaptation of the target network the signal from target regions weakens, allowing excitatory signals from networks representing the background to spread through that region of visual space, likely via lateral propagations (De Weerd et al., 1995; Spillmann & De Weerd, 2003; Weil & Rees, 2011). In contrast to this model, filling-in has also been proposed to be realised at higher representational levels of the cortical hierarchy, whereby to-be filled-in regions are simply labelled as 'more of the same' (Dennett, 1991; 2003; Kingdom & Moulden, 1988; O'Regan, 1992). According to this representational account, PFI does not result from active neural processes in lower visual areas, and instead results after updating object representations (Weil & Rees, 2011).

We were able to test for the presence of interactions between target and background stimulus representations as a consequence of using the SSVEP paradigm. Because of nonlinearities in visual processing, interactions between flickering stimulus representations produce nonlinearities in the frequency domain of EEG spectra, known as intermodulation frequencies (e.g. f2-f1; IMs). Indeed, in our data

we found significant increases in IM strength, and changes in topography that were distinct for subjective PFI compared to catch periods.

During PFI, a transient increase in IM SNR was observed prior to a reported disappearance of flickering targets (Figure 6.10c). By contrast, IMs during catch periods peaked when targets were replaced on screen at catch offset. Previous research investigating IM components has demonstrated that IM strength is strongest when attending to visual competition (Zhang et al., 2011), as well as prior to a reported change in percept during binocular rivalry (Katyal et al., 2016). The increased IM strength demonstrated here, prior to PFI, may also be indicative of active processes mediating between target and background representations in concert with the active-process model. We interpret the peak in IM strength at catch offset as a consequence of the highly salient return of flickering targets.

As IM strength can also represent bottom-up and top-down interactions in the cortical hierarchy (Gordon et al., 2016), we chose to investigate these effects further by performing a steady-state topographical probe (SSTP) analysis. SSTPs are used to infer the propagation of EEG signals in the SSVEP paradigm (Vialatte et al., 2010), and here we compared this signal propagation by focusing on the spatial correlation between frequency-tagged signals. When comparing the 64-channel spatial correlation in SNR values at each frequency (Target, Background, and IM), again, PFI and Catch periods were distinct. During PFI, we observed a transient increase in the SSTP correlation between Target and IM SNR, as well as Background and IM SNR prior to target disappearance (Figure 6.11). As the IM component is driven by nonlinear interactions between Target and Background stimuli, this result strongly demonstrates that PFI is mediated by interactions between a target region and its surrounds. By contrast, during catch periods when targets were physically removed from screen, IMs did not peak prior to disappearance. Instead, a transient increase in SSTP correlation between all frequencies was observed when participants reported on the return of flickering targets at catch offset. Taken together, the opposite direction of these results supports the use of IMs as a capture of neural interaction - separately heralding the disappearance of targets during PFI, and return of flickering targets at catch offset.

6.6.3. Conclusion

By simultaneously flickering both target and background surrounds during PFI, we have provided evidence that the neural representation of a stimulus can be enhanced as it transitions from a visible to invisible state. This contrast between neural representation and stimulus visibility counters contemporary arguments that increased neural activity precedes conscious awareness. In doing so, we have also revealed direct evidence for interactions between a target region and its background surrounds during PFI, which were absent during the phenomenally matched replay of target disappearances. As such, PFI does not solely result from updated object representations, but is a consequence of nonlinear interactions in early visual areas.

6.7 Supplementary Material

6.7.1. Log(SNR) strength as a function of epoch window duration

We were motivated to investigate fluctuations in log(SNR) strength in our time-frequency analyses. While the whole-trial (60 second) EEG spectra revealed significant frequency-tagging (Figure 6.3), SSVEP strength generally decreases as the duration of neural data analysed also decreases. As such, very short windows may be unable to capture time-varying fluctuations in SSVEP responses. To estimate the effect of time-window length on log(SNR) strength, we calculated the across participant log(SNR) at increasing window lengths, between 1 and 60 seconds. Specifically, we compared the log(SNR) when averaging windows of 1 ,2, 4, 6, 10, 12, 15, 30 and 60 seconds. Then, we plotted the log power, log noise, and log(SNR) at each time window duration, focusing on the frequency-tags of interest in the present study (IM = 5 Hz, TG f1 = 15 Hz, BG f2 = 20 Hz, TG 2f1 = 30 Hz, BG 2f2 = 40 Hz). Figure 6.13 a-c shows the change in these metrics as a function of window length, in the range from 1- 10 seconds. We also calculated the across participant EEG spectra at each window length. Figure 6.13d displays the EEG spectra between 0 and 40 Hz at increasing window lengths.

These results show that for our weakest frequency of interest (IM = 5 Hz), a window length of 2 seconds is insufficient to capture a significant peak in the EEG spectrum. As a result of this analysis, we elected a moving time-window of 2.5

seconds for our time-frequency analysis. While the SNR at longer windows (≥4 seconds) is stronger, and the results comparable to those we have presented, we found that with longer windows considerable temporal smearing tended to obscure the dynamic nature of PFI, which displays short, transient changes in log(SNR) strength during changes in visual consciousness.

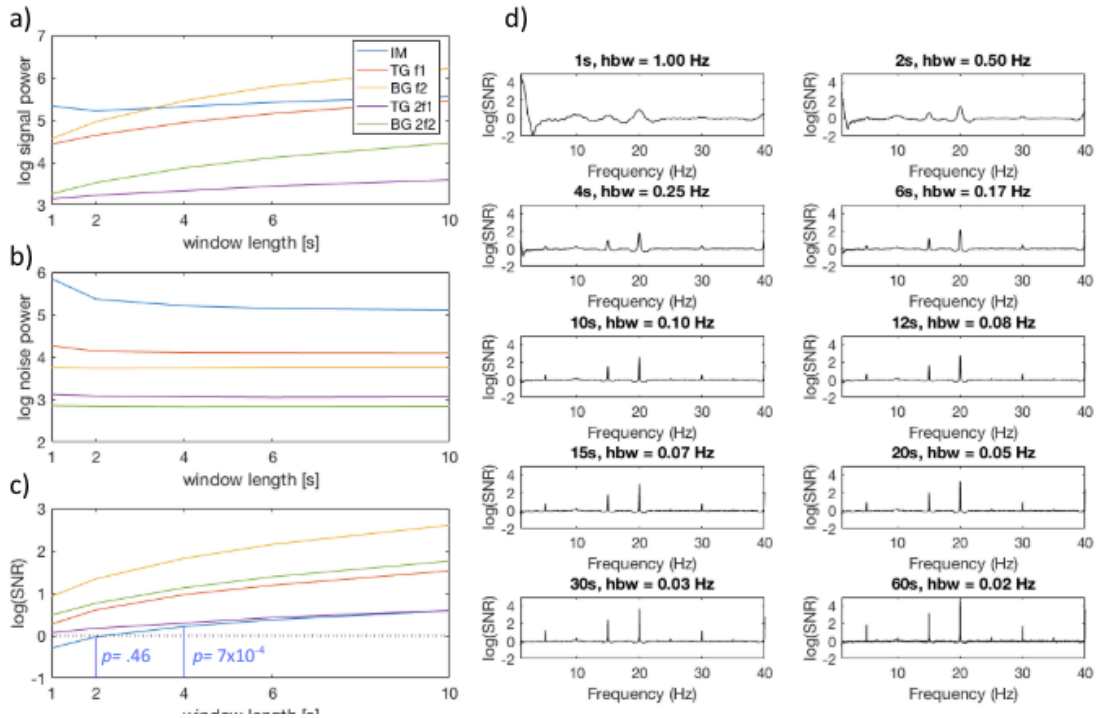

Figure 6.13. Power and SNR characteristics as a function of epoch window duration.

Spectra were calculated after averaging epochs of different duration (1, 2, 4, 6, 10, 12, 15, 20, 30, and 60 sec). a) log power at signals of interest. IM = intermodulation component (5 Hz), Target f1 = 15 Hz, Background f2 = 20 Hz; Target 2f1 = 30 Hz; Background 2f2 = 40 Hz. b) Log noise over 1 Hz, skipping the relevant frequency half bandwidth at each frequency of interest. Such that Noise (N) =([f - hbw -1 Hz], [f - hbw] and [f + hbw]: [f + hbw + 1Hz]. At each increase in epoch duration, the half bandwidth also decreases. c) Log(SNR) as a function of epoch duration. At 2 seconds, across participant spectra demonstrate non-significant frequency-tagging at IM of 5 Hz. By 4 seconds in duration, SNR at 5 Hz is significantly greater than zero across participants. d) Across participant average spectra by epoch window duration.

Chapter 7: General Discussion

7.1 Summary

7.1.1. Aims of the thesis

The primary aim of this thesis was to leverage the methodological advances of frequency-tagging to investigate the relationship between consciousness and attention. New spatial and temporal characteristics of this relationship have been unveiled, paving the way for further investigation and validation within the field.

An analysis of the previous literature identified two promising avenues for investigation. In the first empirical axis it was identified that when investigating the link between attention and consciousness, an over-reliance on binocular rivalry may be suboptimal, given the effects of attention during rivalry are modest (Dieter, Brascamp, et al., 2016; Dieter et al., 2015; Paffen & Alais, 2011), and difficult to disentangle (Blake et al., 2014; Brascamp & Blake, 2012; Koch & Tsuchiya, 2007; P. Zhang et al., 2011). Crossmodal rivalry was identified as a relatively new and promising avenue, in that the perceptual alternations usually resistant to attentional effects respond strongly to attended non-visual stimulation. By adapting a previous psychophysical experiment to manipulate attention (Lunghi et al., 2014), while evoking simultaneous visual, auditory, and tactile steady-state responses, the binding of multisensory information into a coherent percept was shown to critically depend upon the allocation of attention **(Chapters 3 and 4).**

In the second avenue of exploration, the comparatively overlooked visual-perceptual phenomenon known as perceptual filling-in was developed into a frequency-tagging protocol to explore neural markers of attention and consciousness. **Chapter 5 and 6** were motivated by previous observations that attention toward peripherally distinct visual features can hasten their interpolation by the surrounding image background (De Weerd et al., 2006; Hohwy, 2012; Lou, 1999). For the ongoing debates regarding the isomorphism between attention and consciousness, PFI remains an attractive resource - as here attending to visual objects removed them from awareness, suggesting attention and conscious visibility may not be one and the same.

Chapters 5 and 6 refined this paradigm, with distinct neural markers of perceptual change and attention provided in the EEG.

7.2. Main Results

An exploration of specific results is contained in the discussion section of each chapter. Here however, an integrated contribution from all chapters will be considered in relation to the driving empirical questions of this thesis.

7.2.1. Implications for the relationship between attention and consciousness

Several empirical contributions of this thesis support the dissociation between attention and consciousness. These have primarily been revealed through the frequency-dependent dynamics of frequency-tagged binocular rivalry, and fine-grained temporal analysis of behaviour:

- **Frequency-tagging during binocular rivalry can negatively correlate with the contents of consciousness.** In traditional accounts of frequency-tagged activity during rivalry, the SNR of occipital responses increases and decreases with the successive dominance and suppression phases of a relevant image (e.g. Kamphuisen et al., 2008; Lansing, 1964; Tononi, Srinivasan, et al., 1998a; Zhang et al., 2011). This effect was replicated for low-frequency flicker in **Chapter 3,** yet the inclusion of competing high-frequency flicker suggests that these neural responses can become decoupled from the contents of consciousness. For a subset of participants, the strength of high-frequency flicker was shown to negatively correlate with the contents of consciousness. In other words, when a low-frequency flickering image entered awareness the SNR of a nonconscious high-frequency image increased in strength. As discussed in **Chapter 3,** this result may suggest that attention can enhance the SNR of visual flicker even at nonconscious locations, in contrast to an isomorphic mapping between the effects of attention and current contents of consciousness.

- **Divided attentional sampling between a conscious and nonconscious image.** The presentation of intermittent cues to reorient attention, and continuous record of perceptual contents enabled the first investigation of attentional sampling during binocular rivalry to be reported in **Chapter 4.** In this paper, we supplied evidence of a novel hypothesis for the field of perceptual neuroscience: that attention can be flexibly divided between sensory modalities, sampling sensory representations in an oscillatory manner. More intriguing still, by using binocular rivalry to suppress an image in one eye from conscious awareness while another image is visible, we found evidence that attention can also divide its rhythmic sampling between a consciously perceived image and an image suppressed from visual awareness. Our results suggest that focused attention is an inherently rhythmic process, and that distinct neural processes support attention and consciousness.

Together these pieces of evidence provide support for the idea that both the effects and locus of attention can be distributed to a non-conscious image, in support of the dissociation between attention and consciousness.

7.2.2. How does attention interact with multisensory integration to mediate the neural correlates of conscious experience?

Chapters 3 and 4 extended previous psychophysical crossmodal rivalry investigations to investigate the neural correlates of these changes in conscious awareness. The change in perceptual contents following mismatched cues are of particular note.

- **Attention interacts with stimuli outside of conscious awareness.** In the first empirical axis of this thesis, switches between conscious and nonconscious alternatives during binocular rivalry were shown to rely upon the unique combination of low-level stimulus properties (temporal frequency) and the allocation of attention. When a nonconscious low-frequency flicker was congruent with an attended crossmodal stimulus, the contents of consciousness rapidly shifted to bring the previously suppressed image into awareness. This interaction is noteworthy, as it suggests a mediating role of

attention in the binding of non-conscious visual information into a unified percept. More specifically, our demonstration that attention increases the probability of a previously suppressed image reaching awareness strongly suggests that attention can indeed influence the suppression phase of visual stimuli. This result extends accounts of attentional context affecting dominance *only* (e.g. (Dieter et al., 2015; Sobel & Blake, 2002), to include the suppression of a visual stimulus when paired with congruent stimuli (Deroy et al., 2014; Hartcher-O'Brien et al., 2016; Klink et al., 2012).

In the EEG, the neural correlates of these effects were not captured in the narrow frequency-band elicited by stimulus-specific steady-state responses **(Chapter 3),** but indeed by changes to the ITPC at attentional sampling frequencies **(Chapter 4).** The relationship between oscillatory markers of multisensory integration is discussed in more detail below **(Chapter 7.2).**

7.2.3 Can distinct neural correlates of attention and consciousness be captured?

As detailed in the literature review, the precise relationship between attention and consciousness has been difficult to dissociate. A chief concern regards separating neural markers for each constituent process, in order to isolate the true neural correlates of consciousness from the neural correlates of attention, or subjective report.

- **A tentative marker of increased attention during a decrease in conscious visibility.** The second empirical axis of this thesis saw the development of a novel frequency-tagging paradigm combining PFI with simultaneous, multi-target behavioural responses (**Chapters 5 and 6**). In **Chapter 6** this paradigm was refined, to include frequency-tagging of visual objects as they disappear from conscious awareness. During genuine PFI, as flickering target regions disappeared from conscious awareness their consequent SNR was shown to increase, in support of an attentional-gain on invisible objects, and of a distinct neural correlate of attentional effects in opposition to visual awareness.

This work is still in progress, and several important questions are still to be answered before attributing the increase in target-SNR during PFI disappearance to attention unequivocally. Future developments and research are necessary to increase the validity of these claims, and accumulate evidence in favour of this potential dissociation between neural markers of attention and consciousness.

7.3. General limitations and speculations

Each piece of empirical evidence presented throughout this thesis is subject to unique limitations, but overarching and more general criticisms can be highlighted. These include the correlative nature of any contrastive approach in neuroscience, the nature of unconscious information when treating perception as a form of unconscious inference, and the limits of frequency-specific analyses.

7.3.1. To see or not to see

The present thesis has focused on the presence/absence of a subjective quality, and contrasted neural activity during the subjective report of each unique instance. As the contents of consciousness can only be inferred from reports - the fact remains that neural markers associated with a change in qualia may reflect other processes, rather than consciousness per se. This limitation pervades the contrastive approach so dominant and central to NCC investigations (Hohwy, 2009). For example, in binocular rivalry experiments, the difficulty in capturing and contrasting the phenomenal nature of experience is well known **(Chapter 1.3)**, particularly given the often-complex perceptual transitions which occur between each percept. In this thesis I employed two tactics in an attempt to compensate for these issues.

The use of frequency-tagging in the first instance was intended to decrease the search space of neural responses to narrow-frequency bands, and thus neural activity entrained by an identifiable stimulus. However, as shown in **Chapter 3 and 4** the neural correlates of a change to the content of consciousness were not so well-restricted. Activity within canonical frequency-tags either positively or negatively correlated with consciousness **(Chapters 3 and 6**), while activity outside of frequency-tags was shown to mediate the contents of consciousness **(Chapter 4)**. As a second tactic I incorporated a graded measure of perceptual change, rather than the usual dichotomous 'seen-unseen' distinction present in NCC research **(Chapters 5**

and 6). Clearly however, rather than a graded response, these multi-target responses can also be viewed as sets of dichotomous, 'seen-unseen' examples, in which case the same limitation has just been compounded four-fold. Even inheriting these limitations, a deeper assumption is worth identifying. The premise of any contrastive approach is that the perception of a particular quale consistently recruits a particular subset of cortical regions, networks, or other neuronal activity. Such activity would be both *necessary and sufficient* for a conscious percept to occur, yet brain responses even to the same stimuli are inherently variable (Faisal, Selen, & Wolpert, 2008; Tomko & Crapper, 1974). Recent evidence also suggests that complex patterns of neural activity which dynamically unfold over time support conscious vs unconscious *states* (e.g. (Chow et al., 2013; Demertzi et al., 2019; King et al., 2013; Tagliazucchi et al., 2013). Such time-resolved patterns of activity likely account for changes in conscious *contents* also (Demertzi et al., 2019), which were unexplored in the present thesis (yet see **Chapter 7.3.3.** below)

7.3.2. How 'unconscious' can perceptual inference be?

An additional reproach to the contrastive analysis regards the fate of unconscious information if perception is built on unconscious inferences. The findings presented in **Chapters 3 and 4** are intriguing, albeit preliminary. Whether or not the facilitative role of attended low-frequency cues were driven via interactions with a suppressed visual image outside of visual awareness, or primarily through conflict with the dominant visual flicker is difficult to disentangle - and speaks to a broader issue regarding the nature and depth of inter-ocular suppression during rivalry (Blake, 2001; Lin & He, 2009). For example, it is entirely possible that a suppressed image is not well conceived of as 'unconscious', since participants are entirely aware of the presence of a competitive image, as well as the match in stimulus features. It thus remains unclear whether attention can bind features of an invisible (multisensory) object in the absence of conscious awareness.

A truer test of the claim that attention can bind invisible, and unconscious features into a multi-sensory object would require a paradigm in which multi-sensory integration occurs in the *absence of awareness,* yet is conditional upon the allocation of attention to specific visual locations. I will briefly sketch one possible paradigm here, based on previous observations that crossmodal cues can also increase the

visibility of a stimulus masked via CFS (Aller, Giani, Conrad, Watanabe, & Noppeney, 2015; Alsius & Munhall, 2013; Palmer & Ramsey, 2012).

I envision an adaptation of the 2 x 2 factorial comparison which was introduced by van Boxtel et al (2010), to now contrast attention, invisible feature information, and multi-sensory integration. A rough sketch of this paradigm is presented in Figure 7.1.

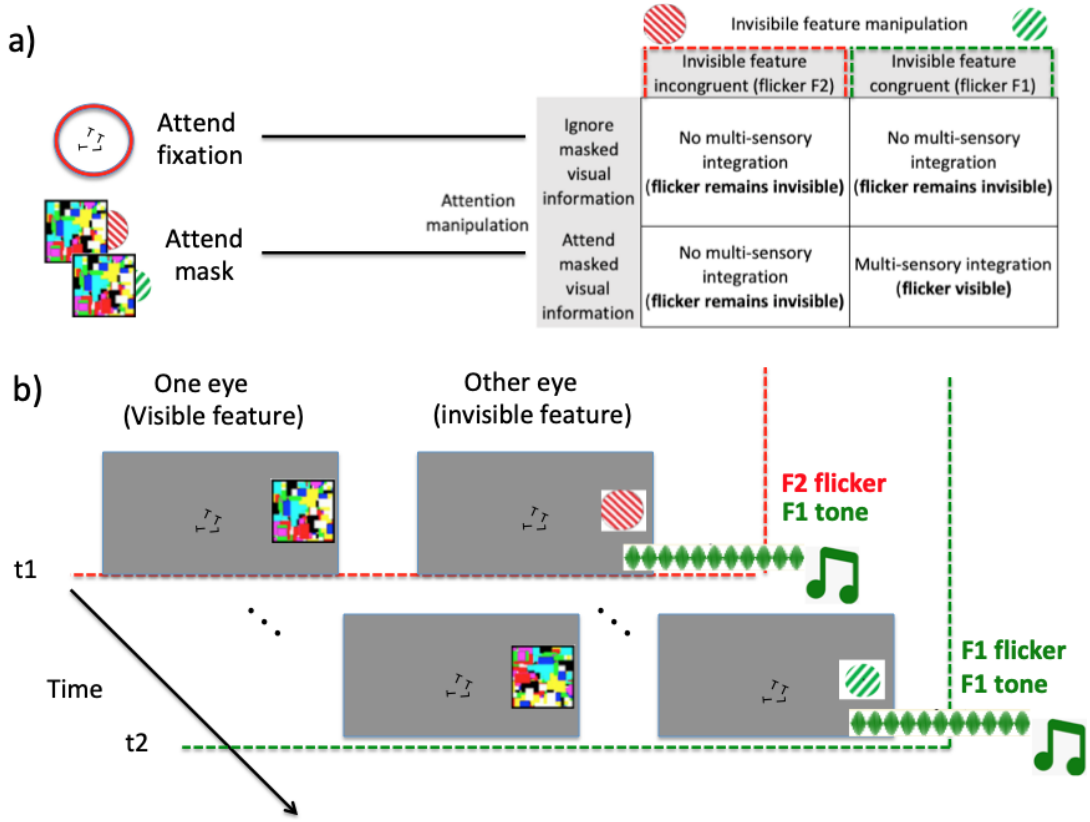

Figure 7.1. Proposed paradigm to dissociate attention, invisible feature information, and multi-sensory integration.

a) Similar to van Boxtel et al. (2010), attention can be manipulated via the presence/absence of a task at fixation. In this iteration, the flicker frequency (F1/F2) of the masked visual stimulus changes without participant awareness (behind the mask). b) At different times during each task, crossmodal cues are played (Here tones at F1 are displayed). Invisible visual features can either be incongruent (e.g. at t1 above) or congruent (e.g. at t2 above) with crossmodal cues. Participants report any instance of a break in CFS (b-CFS), where visual flicker becomes visible. An increase in b-CFS attending to the CFS mask, when the congruent visual feature changed while invisible to participants, could be strong evidence of attention binding invisible features in the absence of conscious awareness

As previously described (**Chapter 1.3**), the allocation of attention can be manipulated by requiring participants to perform a task at fixation, or alternatively, to allocate attention over a high contrast mask. Similar to van Boxtel et al., (2010), this

mask also serves the purpose of inducing CFS, such that information presented to the same location of the other eye remains invisible. I propose to mask visual flicker, rather than an afterimage inducer as in van Boxtel et al., (2010). At jittered intervals, the presentation of a crossmodal cue (an auditory tone in the sketch below) will then be presented, in a similar manner to the methods described in **Chapters 3 and 4.** The critical manipulation I now propose would be to vary the congruence of invisible stimulus features over time, unbeknownst to participants. By changing the visual flicker while masked, multi-sensory congruence varies outside of participant awareness. Note that this contrasts the binocular rivalry case, where the feature-content of suppressed visual stimuli is always known, and perhaps not *unconscious* per se. If the visibility of suppressed flicker increases during CFS, and the feature content of this suppressed visual information has been changing outside of participant awareness, then this could provide compelling evidence for attention acting on truly unconscious visual information. To distinguish between a change in criteria vs change in perceptual sensitivity, a forced choice paradigm could also be used, to probe the nature of the suppressed stimulus rather than free report of breakthrough. This too could provide compelling evidence for the dissociation between attention and consciousness, as the critical feature (invisible flicker congruence), changes outside of conscious awareness, yet within the present focus of spatial attention.

7.3.3. Oscillatory behaviour and functional/effective connectivity

In **Chapter 4** I have shown that spatial clusters of increased phase coherence support a change in the contents of consciousness. This relationship between ITPC strength and changes in consciousness is encouraging, in that phase-synchronization is argued to allow for efficient cortical communication, providing a temporal code with which neural activity may become coupled over time, even if spatially distributed (cf. Fries, 2005; Varela, Lachaux, Rodriguez, & Martinerie, 2001). Consequently, the next fruitful avenue to explore will be how inter-regional communication supports a change in subjective experience, encompassing a wider range of frequencies and time-scales.

As mentioned above, dynamic and temporally resolved patterns of neural activity are likely to support different contents of consciousness (Demertzi et al.,

2019; Engel, Fries, König, Brecht, & Singer, 1999). For the study of consciousness, characterising the interaction of spatially distributed neural activity is important, as competing theories and empirical methods debate both the function and relevance of these interactions (Aru et al., 2012; Baars, 2005; Barttfeld, Uhrig, Sitt, Sigman, & Jarraya, 2014; Zeki, 2003). For example, the binding by coherence hypothesis suggests that the resolution of perceptual ambiguity, and indeed emergence of a coherent percept is dependent upon widespread cortical coherence of rhythmic oscillations (Fries, 2005; Varela, 1994; Varela et al., 2001). From a theoretical standpoint, as brain-activity is largely oscillatory and occurring at multiple-temporal scales, activity in one frequency band may occur independently of another, allowing for complex multi-dimensional information processing (Cohen, 2011; Fontolan, Morillon, Liegeois-Chauvel, & Giraud, 2014; Schyns, Thut, & Gross, 2011). It has been suggested that multiple functionally distinct neural networks can be dissociated according to frequency specificity, yet spatially coexist (Wiener & Kanai, 2016) and oscillations relevant for cognitive processes range from delta (~1-4 Hz), to theta (~4-8 Hz) and gamma (~30-100 Hz) band activity (Buzsáki & Draguhn, 2004; M. X. Cohen, 2011; Varela et al., 2001).

Consequently, the data I've collected during crossmodal rivalry could be analysed to quantify dynamic connectivity strength between regions (Bastos & Schoffelen, 2016). The quality and direction of information transfer between sensory regions could be calculated using measures of functional connectivity, such as mutual information (King et al., 2013), transfer entropy (Lindner, Vicente, Priesemann, & Wibral, 2011), or granger causality (Chuang et al., 2012; Marshall, Lackner, Marriott, Santesso, & Segalowitz, 2014). Recent research has endorsed a hierarchical predictive coding model of visual processing, in which spontaneous perception switches can be described by changes in connectivity strength between parietal and visual regions (Megumi, Bahrami, Kanai, & Rees, 2015). Consequently, an early focus of this work would be to see whether the presentation of crossmodal cues alters the connectivity strength between parietal and visual regions prior to a change in consciousness, and the frequency-bands at which this mediation takes place.

7.4 Conclusion

We have come along way, but there is still a long road ahead in the search for the NCCs. The contributions of this thesis lay the grounds for a renewed discussion and understanding of the relationship between attention and consciousness. I have shown how frequency-tagging is a powerful tool to tease apart this relationship, unveiling new interactions between suppressed visual information and the dynamics of attention. It is my modest hope that on this road, both multisensory paradigms and perceptual filling-in will become recognized for the unique opportunities they afford, and contribute toward unveiling the mystery that is our mental life.

REFERENCES

al-Haitham, I. (1989). *The Optics of Ibn Al-Haytham: Books I-III: on Direct Vision*. Warburg Institute, University of London.

Alais, D. (2012). Binocular rivalry: Competition and inhibition in visual perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(1), 87–103. https://doi.org/10.1002/wcs.151

Alais, D., & Blake, R. (2005). *Binocular rivalry*. MIT press.

Alais, D., Blake, R., & Lee, S.-H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience*, *1*(2), 160.

Alais, D., & Melcher, D. (2007). Strength and coherence of binocular rivalry depends on shared stimulus complexity. *Vision Research*, *47*(2), 269–279.

Alais, D., Newell, F. N., & Mamassian, P. (2010). Multisensory Processing in Review: from Physiology to Behaviour. In *Seeing and Perceiving* (Vol. 23). https://doi.org/10.1371/journal.pone.0011283

Alais, D., & Parker, A. (2006). Independent binocular rivalry processes for motion and form. *Neuron*, *52*(5), 911–920.

Alais, D., van Boxtel, J. J. A., Parker, A., & van Ee, R. (2010). Attending to auditory signals slows visual alternations in binocular rivalry. *Vision Research*, *50*(10), 929–935. https://doi.org/10.1016/j.visres.2010.03.010

Ales, J. M., Farzin, F., Rossion, B., & Norcia, A. M. (2012). An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. *Journal of Vision*, *12*(10), 18.

Aller, M., Giani, A., Conrad, V., Watanabe, M., & Noppeney, U. (2015). A spatially collocated sound thrusts a flash into awareness. *Frontiers in Integrative Neuroscience*, *9*(February), 16. https://doi.org/10.3389/fnint.2015.00016

Almonte, F., Jirsa, V. K., Large, E. W., & Tuller, B. (2005). Integration and segregation in auditory streaming. *Physica D: Nonlinear Phenomena*, *212*(1–2), 137–159. https://doi.org/10.1016/j.physd.2005.09.014

Alp, N., Kogo, N., Van Belle, G., Wagemans, J., & Rossion, B. (2016). Frequency tagging yields an objective neural signature of Gestalt formation. *Brain and Cognition*, *104*, 15–24. https://doi.org/10.1016/j.bandc.2016.01.008

Alpers, G., Ruhleder, M., Walz, N., Mühlberger, A., & Pauli, P. (2005). Binocular Rivalry Between Emotional and Neutral Stimuli: a Validation using Fear Conditioning and EEG. *International Journal of Psychophysiology*, *57*, 25–32. https://doi.org/doi:10.1016/j.ijpsycho.2005.01.008

Alsius, A., & Munhall, K. G. (2013). Detection of Audiovisual Speech Correspondences Without Visual Awareness. *Psychological Science*, *24*(4), 423–431. https://doi.org/10.1177/0956797612457378

Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, *15*(9), 839–843. https://doi.org/10.1016/j.cub.2005.03.046

Andersen, S. K., Hillyard, S. A., & Müller, M. M. (2008). Attention Facilitates Multiple Stimulus Features in Parallel in Human Visual Cortex. *Current Biology*, *18*(13), 1006–1009. https://doi.org/10.1016/j.cub.2008.06.030

Anstis, S. (1996). Adaptation to peripheral flicker. *Vision Research*, *36*(21), 3479–3485. https://doi.org/10.1016/0042-6989(96)00016-8

Anstis, S. (2010). Visual filling-in. *Current Biology*, *20*(16), R664–R666.

Anstis, S., & Greenlee, M. W. (2014). Contour Erasure and Filling-in: New

Observations. *I-Perception*, *5*, 79–86. https://doi.org/10.1068/i0624rep
Aru, J., & Bachmann, T. (2015). *The Constitution of Phenomenal Consciousness: Toward a science and theory* (Vol. 92). https://doi.org/10.1075/aicr.92
Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, *36*(2), 737–746. https://doi.org/10.1016/j.neubiorev.2011.12.003
Baars, B. J. (1993). *A cognitive theory of consciousness*. Cambridge University Press.
Baars, B. J. (1997). *In the theater of consciousness: The workspace of the mind*. Oxford University Press, USA.
Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*(1), 47–52. https://doi.org/10.1016/S1364-6613(00)01819-2
Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research*, Vol. 150, pp. 45–53. https://doi.org/10.1016/S0079-6123(05)50004-9
Babington-Smith, B. (1961). An Unexpected Effect of Attention in Peripheral Vision. *Nature*, *189*, 776. Retrieved from http://dx.doi.org/10.1038/189776a0
Bahrami, B., Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2008a). Spatial attention can modulate unconscious orientation processing. *Perception*, *37*(10), 1520–1528. https://doi.org/10.1068/p5999
Bahrami, B., Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2008b). Unconscious orientation processing depends on perceptual load. *Journal of Vision*, *8*(3), 12. https://doi.org/10.1167/8.3.12
Bahrami, B., Lavie, N., & Rees, G. (2007). Attentional Load Modulates Responses of Human Primary Visual Cortex to Invisible Stimuli. *Current Biology*, *17*(6), 509–513. https://doi.org/10.1016/j.cub.2007.01.070
Banks, M. I., Uhlrich, D. J., Smith, P. H., Krause, B. M., & Manning, K. A. (2011). Descending projections from extrastriate visual cortex modulate responses of cells in primary auditory cortex. *Cerebral Cortex*, *21*(11), 2620–2638.
Barttfeld, P., Uhrig, L., Sitt, J. D., Sigman, M., & Jarraya, B. (2014). *Signature of consciousness in the dynamics of resting-state brain activity*. https://doi.org/10.1073/pnas.1418031112
Başar-Eroglu, C., Strüber, D., Kruse, P., Başar, E., & Stadler, M. (1996). Frontal gamma-band enhancement during multistable visual perception. *International Journal of Psychophysiology*, *24*(1–2), 113–125.
Bastos, A. M., & Schoffelen, J.-M. (2016). A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls. *Frontiers in Systems Neuroscience*, *9*, 175. https://doi.org/10.3389/fnsys.2015.00175
Battelli, L., Pascual-Leone, A., & Cavanagh, P. (2007). The "when" pathway of the right parietal lobe. *Trends in Cognitive Sciences*, *11*(5), 204–210. https://doi.org/10.1016/j.tics.2007.03.001
Bayne, T., & Chalmers, D. (2003). What is the Unity of Consciousness. In *The Unity of Consciousness: Binding, Integration, Dissociation* (pp. 23–58). https://doi.org/10.1093/acprof:oso/9780198508571.003.0002
Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, *15*(2), 145–153. https://doi.org/10.1016/j.conb.2005.03.011
Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507. Retrieved from

https://pdfs.semanticscholar.org/7155/80a7be4c1945b2ab608bd43dd4f718587643.pdf
Benjamini, Y., & Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing under Dependency. *The Annals of Statistics*, *29*(4), 1165–1188.
Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, *87*(1), 527–570.
Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*(June), 1–20. https://doi.org/10.3389/fninf.2015.00016
Bizley, J. K., & King, A. J. (2008). Visual–auditory spatial processing in auditory cortical neurons. *Brain Research*, *1242*, 24–36.
Blake, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, *96*(1), 145–167. https://doi.org/10.1037/0033-295X.96.1.145
Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind*, *2*(1), 5–38.
Blake, R., Brascamp, J. W., & Heeger, D. J. (2014). Can binocular rivalry reveal neural correlates of consciousness? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130211. https://doi.org/10.1098/rstb.2013.0211
Blake, R., & Fox, R. (1974). Binocular rivalry suppression: Insensitive to spatial frequency and orientation change. *Vision Research*, *14*(8), 687–692.
Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews. Neuroscience*, *3*(1), 13–21. https://doi.org/10.1038/nrn701
Blake, R., O'Shea, R. P., & Mueller, T. J. (1992). Spatial zones of binocular rivalry in central and peripheral vision. *Visual Neuroscience*, *8*(05), 469–478. https://doi.org/10.1017/S0952523800004971
Blake, R., Tadin, D., Sobel, K. V, Raissian, T. A., & Chong, S. C. (2006). *Strength of early visual adaptation depends on visual awareness*. Retrieved from www.pnas.orgcgidoi10.1073pnas.0509634103
Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, *18*(2), 227–247.
Block, N. (2008). Consciousness and cognitive access. *Proceedings of the Aristotelian Society*, *108*(1_pt_3), 289–317. Oxford University Press Oxford, UK.
Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, *15*(12), 567–575.
Bokil, H., Andrews, P., Kulkarni, J. E., Mehta, S., & Mitra, P. P. (2010). Chronux: A platform for analyzing neural signals. *Journal of Neuroscience Methods*, *192*(1), 146–151. https://doi.org/10.1016/j.jneumeth.2010.06.020
Boly, M., Massimini, M., Tsychiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are The Neural Correlates Of Consciousness In The Front Or In The Back Of The Cerebral Cortex? Clinical And Neuroimaging Evidence. *Journal of Neuroscience*, *37*(40), 9603–9613. Retrieved from http://biorxiv.org/content/early/2017/03/19/118273.abstract
Bonneh, Y. S., Cooperman, A., & Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature*, *411*(6839), 798–801. https://doi.org/10.1038/35081073
Bonneh, Y. S., Donner, T. H., Cooperman, A., Heeger, D. J., & Sagi, D. (2013). *Motion-induced Blindness and Troxler Fading : Common and Different*

Mechanisms. *9*(3). https://doi.org/10.1371/journal.pone.0092894
Boremanse, A., Norcia, A. M., & Rossion, B. (2013). An objective signature for visual binding of face parts in the human brain. *Journal of Vision*, *13*(11), 6-. https://doi.org/10.1167/13.11.6
Boring, E. G. (1930). A new ambiguous figure. *The American Journal of Psychology*.
Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. https://doi.org/10.1163/156856897X00357
Brascamp, J. W., & Blake, R. (2012). Inattention Abolishes Binocular Rivalry: Perceptual Evidence. *Psychological Science*, *23*(10), 1159–1167. https://doi.org/10.1177/0956797612440100
Brascamp, J. W., Blake, R., & Knapen, T. (2015a). Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nature Neuroscience*, (October). https://doi.org/10.1038/nn.4130
Brascamp, J. W., Blake, R., & Knapen, T. (2015b). Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nature Neuroscience*, *18*(11), 1672–1678. https://doi.org/10.1038/nn.4130
Brascamp, J. W., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable Perception and the Role of Frontoparietal Cortex in Perceptual Inference. *Annual Review of Psychology*, *69*(1), annurev-psych-010417-085944. https://doi.org/10.1146/annurev-psych-010417-085944
Britz, J., Britz, J., Landis, T., Landis, T., Michel, C. M., & Michel, C. M. (2009). Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cerebral Cortex (New York, N.Y. : 1991)*, *19*(1), 55–65. https://doi.org/10.1093/cercor/bhn056
Britz, J., & Pitts, M. A. (2011). Perceptual reversals during binocular rivalry: ERP components and their concomitant source differences. *Psychophysiology*, *48*(11), 1490–1499.
Brooks, J. L., & Palmer, S. E. (2010). *Cue Competition Affects Temporal Dynamics of Edge-assignment in Human Visual Cortex*. Retrieved from https://www-mitpressjournals-org.ezproxy.lib.monash.edu.au/doi/pdf/10.1162/jocn.2010.21433
Brouwer, G. J., & van Ee, R. (2007). Visual Cortex Allows Prediction of Perceptual States during Ambiguous Structure-From-Motion. *Journal of Neuroscience*, *27*(5), 1015–1023. https://doi.org/10.1523/JNEUROSCI.4593-06.2007
Brown, R. J., & Norcia, A. M. (1997). A method for investigating binocular rivalry in real-time with the steady-state VEP. *Vision Research*, *37*(17), 2401–2408. https://doi.org/10.1016/S0042-6989(97)00045-X
Budd, T. W., & Timora, J. R. (2013). Steady state responses to temporally congruent and incongruent auditory and vibrotactile amplitude modulated stimulation. *International Journal of Psychophysiology*, *89*(3), 419–432. https://doi.org/10.1016/j.ijpsycho.2013.06.001
Busch, N. A., Dubois, J., & VanRullen, R. (2009). The Phase of Ongoing EEG Oscillations Predicts Visual Perception. *Journal of Neuroscience*, *29*(24), 7869–7876. https://doi.org/10.1523/JNEUROSCI.0113-09.2009
Busch, N. A., & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, *107*(37), 16048–16053. https://doi.org/10.1073/pnas.1004801107
Bushara, K. O., Hanakawa, T., Immisch, I., Toma, K., Kansaku, K., & Hallett, M. (2003). Neural correlates of cross-modal binding. *Nature Neuroscience*, *6*(2), 190–195. https://doi.org/10.1038/nn993

Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, *304*(5679), 1926–1929.
Calvert, G. A. (2001). Crossmodal Processing in the Human Brain: Insights from Functional Neuroimaging Studies. *Cerebral Cortex*, *11*(12), 1110–1123. https://doi.org/10.1093/cercor/11.12.1110
Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1–3), 191–205. https://doi.org/10.1016/j.jphysparis.2004.03.018
Carmel, D., Walsh, V., Lavie, N., & Rees, G. (2010). Right parietal TMS shortens dominance durations in binocular rivalry. *Current Biology*, *20*(18). https://doi.org/10.1016/j.cub.2010.07.036
Carter, O. L., Konkle, T., Wang, Q., Hayward, V., & Moore, C. (2008). Tactile rivalry demonstrated with an ambiguous apparent-motion quartet. *Current Biology*, *18*(14), 1050–1054.
Carter, O. L., & Pettigrew, J. D. (2003). A common oscillator for perceptual rivalries? *Perception*, *32*(3), 295–305. https://doi.org/10.1068/p3472
Chakravarthi, R., & VanRullen, R. (2012). Conscious updating is a rhythmic process. *Proceedings of the National Academy of Sciences*, *109*(26), 10599–10604. https://doi.org/10.1073/pnas.1121622109
Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*(3), 200–219.
Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford university press.
Chen, A., Wang, A., Wang, T., Tang, X., & Zhang, M. (2017). Behavioral Oscillations in Visual Attention Modulated by Task Difficulty. *Frontiers in Psychology*, *8*, 1630. https://doi.org/10.3389/fpsyg.2017.01630
Chen, Y. C., Yeh, S. L., & Spence, C. (2011). Crossmodal constraints on human perceptual awareness: Auditory semantic modulation of binocular rivalry. *Frontiers in Psychology*, *2*(SEP). https://doi.org/10.3389/fpsyg.2011.00212
Chen, Y., Seth, A. K., Gally, J. A., & Edelman, G. M. (2003). The power of human brain magnetoencephalographic signals can be modulated up or down by changes in an attentive visual task. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(6), 3501–3506. https://doi.org/10.1073/pnas.0337630100
Chica, A. B., & Bartolomeo, P. (2012). Attentional routes to conscious perception. *Frontiers in Psychology*, *3*, 1. https://doi.org/10.3389/fpsyg.2012.00001
Chong, S. C., & Blake, R. (2006). Exogenous attention and endogenous attention influence initial dominance in binocular rivalry. *Vision Research*, *46*(11), 1794–1803. https://doi.org/10.1016/j.visres.2005.10.031
Chong, S. C., Tadin, D., & Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *Journal of Vision*, *5*(11), 6.
Chopin, A., & Mamassian, P. (2010). Task usefulness affects perception of rivalrous images. *Psychological Science*, *21*(12), 1886–1893. https://doi.org/10.1177/0956797610389190
Chow, H. M., Horovitz, S. G., Carr, W. S., Picchioni, D., Coddington, N., Fukunaga, M., … Braun, A. R. (2013). Rhythmic alternating patterns of brain activity distinguish rapid eye movement sleep from other states of consciousness. *Proceedings of the National Academy of Sciences*, *110*(25), 10300–10305.
Chuang, C. H., Huang, C. S., Lin, C. T., Ko, L. W., Chang, J. Y., & Yang, J. M. (2012). Mapping information flow of independent source to predict conscious

level: A granger causality based brain-computer interface. *Proceedings - 2012 International Symposium on Computer, Consumer and Control, IS3C 2012*, 813–816. https://doi.org/10.1109/IS3C.2012.209
Cobb, W. a, Morton, H. B., & Ettlinger, G. (1967). Cerebral potentials evoked by pattern reversal and their suppression in visual rivalry. *Nature*, *216*(5120), 1123–1125. https://doi.org/10.1038/2161123b0
Cohen, M. A., Cavanagh, P., Chun, M. M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences*, *16*(8), 411–417. https://doi.org/10.1016/j.tics.2012.06.013
Cohen, M. A., & Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends in Cognitive Sciences*, *15*(8), 358–364. https://doi.org/10.1016/j.tics.2011.06.008
Cohen, M. X. (2011). It's about Time. *Frontiers in Human Neuroscience*, *5*(January), 2. https://doi.org/10.3389/fnhum.2011.00002
Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT Press.
Cohen, M. X., & Gulbinaite, R. (2017). Rhythmic entrainment source separation: Optimizing analyses of neural responses to rhythmic sensory stimulation. *NeuroImage*, 070862. https://doi.org/10.1101/070862
Colon, E., Legrain, V., & Mouraux, a. (2012). Steady-state evoked potentials to study the processing of tactile and nociceptive somatosensory input in the human brain. *Neurophysiologie Clinique*, *42*(5), 315–323. https://doi.org/10.1016/j.neucli.2012.05.005
Conrad, V., Bartels, A., Kleiner, M., & Noppeney, U. (2010). Audiovisual interactions in binocular rivalry. *Journal of Vision*, *10*(10), 27. https://doi.org/10.1167/10.10.27
Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755
Cosmelli, D., David, O., Lachaux, J.-P., Martinerie, J., Garnero, L., Renault, B., & Varela, F. (2004). Waves of consciousness: ongoing cortical patterns during binocular rivalry. *Neuroimage*, *23*(1), 128–140.
Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/no DOI found
Covic, A., Keitel, C., Porcu, E., Schröger, E., & Müller, M. M. (2017). Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: A frequency-tagging study. *NeuroImage*, *161*(April), 32–42. https://doi.org/10.1016/j.neuroimage.2017.08.022
Crick, F., & Koch, C. (1990a). Some reflections on visual awareness. *Cold Spring Harbor Symposia on Quantitative Biology*, *55*, 953–962. Cold Spring Harbor Laboratory Press.
Crick, F., & Koch, C. (1990b). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275. Retrieved from http://papers.klab.caltech.edu/22/1/148.pdf
Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, *375*(6527), 121–123.
Crick, F., & Koch, C. (2003). A Framework for Consciousness. *Nature Neuroscience*, *6*(2), 119–126. https://doi.org/10.1038/nn0203-119

Davidson, M. J., Alais, D., van Boxtel, J. J. A., & Tsuchiya, N. (2018). Attention periodically samples competing stimuli during binocular rivalry. *BioRxiv*. Retrieved from http://biorxiv.org/content/early/2018/01/25/253740.abstract

Davis, H. (1964). Enhancement of evoked cortical potentials in humans related to a task requiring a decision. *Science*, *145*(3628), 182–183.

De Graaf, T. A., Hsieh, P. J., & Sack, A. T. (2012). The "correlates" in neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, Vol. 36, pp. 191–197. https://doi.org/10.1016/j.neubiorev.2011.05.012

de Jong, R., Toffanin, P., & Harbers, M. (2010). Dynamic crossmodal links revealed by steady-state responses in auditory-visual divided attention. *International Journal of Psychophysiology*, *75*(1), 3–15. https://doi.org/10.1016/j.ijpsycho.2009.09.013

De Meo, R., Murray, M. M., Clarke, S., Matusz, P. J., Soto-Faraco, S., & Wallace, M. T. (2015). Top-down control and early multisensory processes: Chicken vs. egg. *Frontiers in Integrative Neuroscience*, *9*(3), 1–6. https://doi.org/10.3389/fnint.2015.00017

De Weerd, P. (2006). Perceptual filling-in: more than the eye can see. *Progress in Brain Research*, *154*(SUPPL. A), 227–245. https://doi.org/10.1016/S0079-6123(06)54012-9

De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Perceptual filling-in: A parametric study. *Vision Research*, *38*(18), 2721–2734. https://doi.org/10.1016/S0042-6989(97)00432-X

De Weerd, P., Gattass, R., Desimone, R., & Ungerleider, L. G. (1995). Repsonses of cells in monkey visual cortex during Perceptual filling-in of an artifical scotoma. *Nature*, *377*, 731–734.

De Weerd, P., Smith, E., & Greenberg, P. (2006). Effects of selective attention on perceptual filling-in. *Journal of Cognitive Neuroscience*, *18*(3), 335–347. https://doi.org/10.1162/jocn.2006.18.3.335

Degerman, A., Rinne, T., Pekkola, J., Autti, T., Jääskeläinen, I. P., Sams, M., & Alho, K. (2007). Human brain activity associated with audiovisual perception and attention. *NeuroImage*, *34*(4), 1683–1691. https://doi.org/10.1016/j.neuroimage.2006.11.019

Dehaene, S., & Changeux, J.-P. (2005). Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. *PLoS Biology*, *3*(5), e141.

Dehaene, S., & Changeux, J. P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, *70*(2), 200–227. https://doi.org/10.1016/j.neuron.2011.03.018

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211. https://doi.org/10.1016/j.tics.2006.03.007

Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, *95*(24), 14529–14534.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, *79*(1–2), 1–37.

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology*, *5*(10), e260.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *13*, 9–21. https://doi.org/http://dx.doi.org/10.1016/j.jneumeth.2003.10.009

Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., … Voss, H. U. (2019). Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Science Advances*, *5*(2), eaat7603.

Dennett, D. C. (1991). *Consciousness Explained. Boston (Little, Brown and Co) 1991.*

Deroy, O., Chen, Y.-C., & Spence, C. (2014). Multisensory constraints on awareness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1641), 20130207–20130207. https://doi.org/10.1098/rstb.2013.0207

Deroy, O., Spence, C., & Noppeney, U. (2016, October). Metacognition in Multisensory Perception. *Trends in Cognitive Sciences*, Vol. 20, pp. 736–747. https://doi.org/10.1016/j.tics.2016.08.006

Devyatko, D., Appelbaum, L. G., & Mitroff, S. R. (2016). A Common Mechanism for Perceptual Reversals in Motion-Induced Blindness, the Troxler Effect, and Perceptual Filling-In. *Perception*. https://doi.org/10.1177/0301006616672577

Dieter, K. C., Brascamp, J. W., Tadin, D., & Blake, R. (2016). Does visual attention drive the dynamics of bistable perception? *Attention, Perception, & Psychophysics*, *78*(7), 1861–1873. https://doi.org/10.3758/s13414-016-1143-2

Dieter, K. C., Melnick, M. D., & Tadin, D. (2015). When can attention influence binocular rivalry? *Attention, Perception, & Psychophysics*, *77*(6), 1908–1918. https://doi.org/10.3758/s13414-015-0905-6

Dieter, K. C., Melnick, M. D., & Tadin, D. (2016). Perceptual training profoundly alters binocular rivalry through both sensory and attentional enhancements. *Proceedings of the National Academy of Sciences*, *113*(45), 12874–12879. https://doi.org/10.1073/pnas.1602722113

Dieter, K. C., & Tadin, D. (2011). Understanding Attentional Modulation of Binocular Rivalry: A Framework Based on Biased Competition. *Frontiers in Human Neuroscience*, *5*, 155. https://doi.org/10.3389/fnhum.2011.00155

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

Doner, J., Lappin, J. S., & Perfetto, G. (1984). Detection of three-dimensional structure in moving optical patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 1.

Donner, T. H., Sagi, D., Bonneh, Y. S., & Heeger, D. J. (2008). Opposite neural signatures of motion-induced blindness in human dorsal and ventral visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(41), 10298–10310. https://doi.org/10.1523/JNEUROSCI.2371-08.2008

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, *3*(3), 277–283. https://doi.org/10.1038/72991

Driver, J., & Noesselt, T. (2008). Multisensory Interplay Reveals Crossmodal Influences on "Sensory-Specific" Brain Regions, Neural Responses, and Judgments. *Neuron*, *57*(1), 11–23. https://doi.org/10.1016/j.neuron.2007.12.013

Dugué, L., McLelland, D., Lajous, M., & VanRullen, R. (2015). Attention searches nonuniformly in space and in time. *Proceedings of the National Academy of Sciences*, *112*(49), 15214–15219. https://doi.org/10.1073/pnas.1511331112

Dugué, L., Merriam, E. P., Heeger, D. J., & Carrasco, M. (2017). Specific Visual

Subregions of TPJ Mediate Reorienting of Spatial Attention. *Cerebral Cortex*, 1–16. https://doi.org/10.1093/cercor/bhx140
Dugué, L., Roberts, M., & Carrasco, M. (2016). Attention Reorients Periodically. *Current Biology*, *26*(12), 1595–1601. https://doi.org/10.1016/j.cub.2016.04.046
Dugué, L., & VanRullen, R. (2014). The dynamics of attentional sampling during visual search revealed by Fourier analysis of periodic noise interference. *Journal of Vision*, *14*, 1–15. https://doi.org/10.1167/14.2.11.doi
Dugué, L., & VanRullen, R. (2017). Transcranial magnetic stimulation reveals intrinsic perceptual and attentional rhythms. *Frontiers in Neuroscience*, *11*, 154. https://doi.org/10.3389/fnins.2017.00154
Dugué, L., Xue, A. M., & Carrasco, M. (2017). Distinct perceptual rhythms for feature and conjunction searches. *Journal of Vision*, *17*(3), 22. https://doi.org/10.1167/17.3.22
Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, *39*(2), 647–660. https://doi.org/10.1016/j.neuroimage.2007.09.034
Durgin, F. H., Srimant, T. P., & Levi, D. M. (1995). On the filling in of the visual blind spot: some rules of thumb. *Perception*, *24*(7), 827–840. https://doi.org/10.1068/p240827
Dykstra, A. R., Cariani, P. A., & Gutschalk, A. (2017). A roadmap for the study of conscious audition and its neural basis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1714). https://doi.org/10.1098/rstb.2016.0103
Edelman, G. M. (1987). *Neural Darwinism: The theory of neuronal group selection.* Basic books.
Edelman, G. M. (1989). *The remembered present*. Basic Books, New York.
Edelman, G. M. (1993). Neural Darwinism: selection and reentrant signaling in higher brain function. *Neuron*, *10*(2), 115–125.
Edelman, G. M., Gally, J. A., & Baars, B. J. (2011). Biology of consciousness. *Frontiers in Psychology*, *2*, 4.
Edelman, G. M., & Mountcastle, V. B. (1978). *The mindful brain: cortical organization and the group-selective theory of higher brain function.* Massachusetts Inst of Technology Pr.
Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness: How matter becomes imagination*. Basic books.
Engel, a K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Consciousness and Cognition*, *8*(2), 128–151. https://doi.org/10.1006/ccog.1999.0389
Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, *9*(4), 292–303. https://doi.org/10.1038/nrn2258
Faivre, N., Arzi, A., Lunghi, C., & Salomon, R. (2017). Consciousness is more than meets the eye: a call for a multisensory study of subjective experience†. *Neuroscience of Consciousness*, *3*(1), 1–8. https://doi.org/10.1093/nc/nix003
Falchier, A., Schroeder, C. E., Hackett, T. A., Lakatos, P., Nascimento-Silva, S., Ulbert, I., … Smiley, J. F. (2009). Projection from visual areas V2 and prostriata to caudal auditory cortex in the monkey. *Cerebral Cortex*, *20*(7), 1529–1538.
Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience*, *8*(10), 1380.
Farahani, E. D., Goossens, T., Wouters, J., & van Wieringen, A. (2017). Spatiotemporal reconstruction of auditory steady-state responses to acoustic

amplitude modulations: Potential sources beyond the auditory pathway. *NeuroImage*, *148*(December 2016), 240–253. https://doi.org/10.1016/j.neuroimage.2017.01.032
Felleman, D. J., & Van, D. C. E. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, NY: 1991)*, *1*(1), 1–47.
Fiebelkorn, I. C., & Kastner, S. (2019). A rhythmic theory of attention. *Trends in Cognitive Sciences*, *xx*, 1–36. https://doi.org/10.1016/j.tics.2018.11.009
Fiebelkorn, I. C., Pinsk, M. A., & Kastner, S. (2018). A Dynamic Interplay within the Frontoparietal Network Underlies Rhythmic Spatial Attention. *Neuron*, *99*(4), 842-853.e8. https://doi.org/10.1016/J.NEURON.2018.07.038
Fiebelkorn, I. C., Saalmann, Y. B., & Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. *Current Biology*, *23*(24), 2553–2558. https://doi.org/10.1016/j.cub.2013.10.063
Fiorani Jr., M., Rosa, M. G. P., Gattass, R., & Rocha-Miranda, C. E. (1992). Dynamic surrounds of receptive fields in primate striate cortex: A physiological basis for perceptual completion. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(18), 8547–8551. https://doi.org/10.1073/pnas.89.18.8547
Fontolan, L., Morillon, B., Liegeois-Chauvel, C., & Giraud, A.-L. (2014). *The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex*. https://doi.org/10.1038/ncomms5694
Fort, A., Delpuech, C., Pernier, J., & Giard, M.-H. (2002). Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. *Cerebral Cortex*, *12*(10), 1031–1039.
Fox, R., & Check, R. (1968). Detection of motion during binocular rivalry suppression. *Journal of Experimental Psychology*, *78*(3p1), 388.
Fox, R., Todd, S., & Bettinger, L. A. (1975). Optokinetic nystagmus as an objective indicator of binocular rivalry. *Vision Research*, *15*(7), 849–853.
Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognitive Brain Research*, *10*(1–2), 77–83. https://doi.org/10.1016/S0926-6410(00)00024-0
Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., … Murray, M. M. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *Journal of Neurophysiology*, *88*(1), 540–543.
Frassle, S., Sommer, J., Jansen, A., Naber, M., & Einhauser, W. (2014). Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. *Journal of Neuroscience*, *34*(5), 1738–1747. https://doi.org/10.1523/JNEUROSCI.4403-13.2014
Frey, J. N., Ruhnau, P., & Weisz, N. (2015). Not so different after all: The same oscillatory processes support different types of attention. *Brain Research*, *1626*, 183–197. https://doi.org/10.1016/j.brainres.2015.02.017
Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. https://doi.org/10.1016/j.tics.2005.08.011
Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, *88*(1), 220–235. https://doi.org/10.1016/j.neuron.2015.09.034
Fries, P., Neuenschwander, S., Engel, A. K., Goebel, R., & Singer, W. (2001). Rapid feature selective neuronal synchronization through correlated latency shifting.

Nature Neuroscience, *4*(2), 194.
Fries, P., Roelfsema, P. R., Engel, A. K., König, P., & Singer, W. (1997). Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proceedings of the National Academy of Sciences*, *94*(23), 12699–12704.
Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Experimental Brain Research*, *166*(3–4), 455–464. https://doi.org/10.1007/s00221-005-2385-8
Fujisaki, W., & Nishida, S. (2010). A common perceptual temporal limit of binding synchronous inputs across different sensory attributes and modalities. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1692), 2281–2290. https://doi.org/10.1098/rspb.2010.0243
Fujiwara, M., Ding, C., Kaunitz, L., Stout, J. C., Thyagarajan, D., & Tsuchiya, N. (2017). Optokinetic nystagmus reflects perceptual directions in the onset binocular rivalry in Parkinson's disease. *PLoS ONE*, *12*(3), 1–22. https://doi.org/10.1371/journal.pone.0173707
Gallotto, S., Sack, A. T., Schuhmann, T., & de Graaf, T. A. (2017). Oscillatory Correlates of Visual Consciousness. *Frontiers in Psychology*, *8*(July), 1147. https://doi.org/10.3389/fpsyg.2017.01147
Gassel, M. M., & Williams, D. (1963). Visual function in patients with homonymous hemianopia. *Brain*, *86*(2), 229–260.
Geng, H., Song, Q., Li, Y., Xu, S., & Zhu, Y. (2007). Attentional modulation of motion-induced blindness. *Chinese Science Bulletin*, *52*(8), 1063–1070. https://doi.org/10.1007/s11434-007-0178-0
Gerrits, H. J., & Timmerman, G. J. (1969). The filling-in process in patients with retinal scotomata. *Vision Research*, *9*(3), 439–442.
Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, Vol. 10, pp. 278–285. https://doi.org/10.1016/j.tics.2006.04.008
Giabbiconi, C. M., Dancer, C., Zopf, R., Gruber, T., & Müller, M. M. (2004). Selective spatial attention to left or right hand flutter sensation modulates the steady-state somatosensory evoked potential. *Cognitive Brain Research*, *20*(1), 58–66. https://doi.org/10.1016/j.cogbrainres.2004.01.004
Giabbiconi, C. M., Trujillo-Barreto, N. J., Gruber, T., & Müller, M. M. (2007). Sustained spatial attention to vibration is mediated in primary somatosensory cortex. *NeuroImage*, *35*(1), 255–262. https://doi.org/10.1016/j.neuroimage.2006.11.022
Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *J Cogn Neurosci*, *11*(5), 473–490. https://doi.org/10.1162/089892999563544
Glover, S., & Dixon, P. (2004). Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review*, *11*(5), 791–806.
Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.
Gordon, N., Koenig-Robert, R., Tsuchiya, N., Van Boxtel, J. J. A., & Hohwy, J. (2017). Neural markers of predictive coding under perceptual uncertainty revealed with hierarchical frequency tagging. *ELife*, *6*, 1–17. https://doi.org/10.7554/eLife.22749
Gratton, G., Cooper, P., Fabiani, M., Carter, C., & Karayanidis, F. (2017). *Cognitive control: theoretical bases, paradigms, and a view for the future*. (September), 1–

29. https://doi.org/10.1111/psyp.13016
Gross, J. (2014). Analytical methods and experimental approaches for electrophysiological studies of brain oscillations. *Journal of Neuroscience Methods*, *228*, 57–66.
Guggisberg, A. G., Dalal, S. S., Schnider, A., & Nagarajan, S. S. (2011). The neural basis of event-time introspection. *Consciousness and Cognition*, *20*(4), 1899–1915. https://doi.org/10.1016/j.concog.2011.03.008
Gulbinaite, R., Roozendaal, D. H. M., & Vanrullen, R. (2019). *Attention differentially modulates the amplitude of resonance frequencies in the visual cortex*. 1–40.
Gundlach, C., & Müller, M. M. (2013). Perception of illusory contours forms intermodulation responses of steady state visual evoked potentials as a neural signature of spatial integration. *Biological Psychology*, *94*(1), 55–60. https://doi.org/10.1016/j.biopsycho.2013.04.014
Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology*, *22*(5), 383–388. https://doi.org/10.1016/j.cub.2012.01.004
Haas, L. F. (2003). Hans berger (1873–1941), richard caton (1842–1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(1), 9.
Hanslmayr, S., Gross, J., Klimesch, W., & Shapiro, K. L. (2011). The role of alpha oscillations in temporal attention. *Brain Research Reviews*, *67*(1–2), 331–343. https://doi.org/10.1016/j.brainresrev.2011.04.002
Hanslmayr, S., Volberg, G., Wimber, M., Dalal, S. S., & Greenlee, M. W. (2013). Prestimulus oscillatory phase at 7 Hz gates cortical information flow and visual perception. *Current Biology*, *23*(22), 2273–2278. https://doi.org/10.1016/j.cub.2013.09.020
Harris, K. D., & Thiele, A. (2011). Cortical state and attention. *Nature Reviews Neuroscience*, *12*(9), 509–523. https://doi.org/10.1038/nrn3084
Hartcher-O'Brien, J., Talsma, D., Adam, R., Vercillo, T., Macaluso, E., & Noppeney, U. (2016). The Curious Incident of Attention in Multisensory Integration: Bottom-up vs. Top-down. *Multisensory Research*, *29*(6–7), 557–583. https://doi.org/10.1163/22134808-00002528
Haynes, J.-D., Driver, J., & Rees, G. (2005). Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. *Neuron*, *46*(5), 811–821.
Helfrich, R. F. (2018). The rhythmic nature of visual perception. *Journal of Neurophysiology*, *119*(4), 1251–1253. https://doi.org/10.1152/jn.00810.2017
Henschke, J. U., Noesselt, T., Scheich, H., & Budinger, E. (2015). Possible anatomical pathways for short-latency multisensory integration processes in primary sensory cortices. *Brain Structure and Function*, *220*(2), 955–977.
Herrmann, C. S., Strüber, D., Helfrich, R. F., & Engel, A. K. (2015). EEG oscillations: From correlation to causality. *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2015.02.003
Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science*, *172*(3990), 1357–1360.
Ho, H. T., Leung, J., Burr, D. C., Alais, D., & Morrone, M. C. (2017a). Auditory Sensitivity and Decision Criteria Oscillate at Different Frequencies Separately for the Two Ears. *Current Biology*, *27*(23), 3643-3649.e3. https://doi.org/10.1016/j.cub.2017.10.017
Ho, H. T., Leung, J., Burr, D. C., Alais, D., & Morrone, M. C. (2017b). Auditory

Sensitivity and Decision Criteria Oscillate at Different Frequencies Separately for the Two Ears. *Current Biology*, *27*(23), 3643-3649.e3. https://doi.org/10.1016/j.cub.2017.10.017
Hogendoorn, H. (2016). Voluntary Saccadic Eye Movements Ride the Attentional Rhythm. *Journal of Cognitive Neuroscience*, *28*(10), 1625–1635. https://doi.org/10.1162/jocn
Hohwy, J. (2009). The neural correlates of consciousness: New experimental approaches needed? *Consciousness and Cognition*, *18*(2), 428–438. https://doi.org/10.1016/j.concog.2009.02.006
Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Frontiers in Psychology*, *3*(APR), 1–14. https://doi.org/10.3389/fpsyg.2012.00096
Hohwy, J., Roepstorff, A., & Friston, K. (2008). Predictive coding explains binocular rivalry: An epistemological review. *Cognition*, *108*(3), 687–701.
Holcombe, A. O., & Chen, W.-Y. (2013). Splitting attention reduces temporal resolution from 7 Hz for tracking one object to <3 Hz when tracking three. *Journal of Vision*, *13*(1), 1–19. https://doi.org/10.1167/13.1.12
Hollins, M. (1980). The effect of contrast on the completeness of binocular rivalry suppression. *Perception & Psychophysics*, *27*(6), 550–556.
Hsu, L.-C., Yeh, S.-L., & Kramer, P. (2004). Linking motion-induced blindness to perceptual filling-in. *Vision Research*, *44*(24), 2857–2866.
Hsu, L.-C., Yeh, S. L., & Kramer, P. (2006). A common mechanism for perceptual filling-in and motion-induced blindness. *Vision Research*, *46*(12), 1973–1981. https://doi.org/10.1016/j.visres.2005.11.004
Huang, Y., Chen, L., & Luo, H. (2015). Behavioral Oscillation in Priming: Competing Perceptual Predictions Conveyed in Alternating Theta-Band Rhythms. *Journal of Neuroscience*, *35*(6), 2830–2837. https://doi.org/10.1523/JNEUROSCI.4294-14.2015
Hugrass, L., & Crewther, D. (2012). Willpower and conscious percept: volitional switching in binocular rivalry. *PloS One*, *7*(4), e35963.
Iemi, L., & Busch, N. A. (2018). Moment-to-Moment Fluctuations in Neuronal Excitability Bias Subjective Perception Rather than Strategic Decision-Making. *Eneuro*, *ENEURO.043*. https://doi.org/10.1523/ENEURO.0430-17.2018
Iemi, L., Chaumon, M., Crouzet, S. M., & Busch, N. A. (2017). Spontaneous Neural Oscillations Bias Perception by Modulating Baseline Excitability. *The Journal of Neuroscience*, *37*(4), 807–819. https://doi.org/10.1523/JNEUROSCI.1432-16.2017
Intaitė, M., Koivisto, M., Rukšėnas, O., & Revonsuo, A. (2010). Reversal negativity and bistable stimuli: Attention, awareness, or something else? *Brain and Cognition*, *74*(1), 24–34.
Isoglu-Alkaç, Ü., Basar-Eroglu, C., Ademoglu, A., Demiralp, T., Miener, M., & Stadler, M. (2000). Alpha activity decreases during the perception of Necker cube reversals: an application of wavelet transform. *Biological Cybernetics*, *82*(4), 313–320.
Itthipuripat, S., Garcia, J. O., & Serences, J. T. (2013). Temporal dynamics of divided spatial attention. *Journal of Neurophysiology*, *109*(9), 2364–2373. https://doi.org/10.1152/jn.01051.2012
Jack, B. N., & Hacker, G. (2014). Predictive Coding Explains Auditory and Tactile Influences on Vision during Binocular Rivalry. *Journal of Neuroscience*, *34*(19), 6423–6424. https://doi.org/10.1523/JNEUROSCI.1040-14.2014

Jamison, K. W., Roy, A. V, He, S., Engel, S. A., & He, B. (2015a). SSVEP signatures of binocular rivalry during simultaneous EEG and fMRI. *Journal of Neuroscience Methods*, *243*, 53–62.
Jamison, K. W., Roy, A. V, He, S., Engel, S. a, & He, B. (2015b). SSVEP signatures of binocular rivalry during simultaneous EEG and fMRI. *Journal of Neuroscience Methods*, *243*, 1–10. https://doi.org/10.1016/j.jneumeth.2015.01.024
Kamphuisen, A., Bauer, M., & van Ee, R. (2008). No evidence for widespread synchronized networks in binocular rivalry: MEG frequency tagging entrains primarily early visual cortex. *Journal of Vision*, *8*(5), 1–8. https://doi.org/10.1167/8.5.4
Kanai, R., Carmel, D., Bahrami, B., & Rees, G. (2011). Structural and functional fractionation of right superior parietal cortex in bistable perception. *Current Biology*, *21*(3). https://doi.org/10.1016/j.cub.2010.12.009
Kanai, R., & Tsuchiya, N. (2012). Qualia. *Current Biology*, *22*(10), R392–R396.
Kanai, R., Tsuchiya, N., & Verstraten, F. a J. (2006). The Scope and Limits of Top-Down Attention in Unconscious Visual Processing. *Current Biology*, *16*(23), 2332–2336. https://doi.org/10.1016/j.cub.2006.10.001
Kang, M.-S., & Blake, R. (2005). Perceptual synergy between seeing and hearing revealed during binocular rivalry. *Psichologija*, *32*, 7–15.
Kaplan, I. T., & Metlay, W. (1964). Light intensity and binocular rivalry. *Journal of Experimental Psychology*, *67*(1), 22.
Kaspar, K., Hassler, U., Martens, U., Trujillo-Barreto, N., & Gruber, T. (2010). Steady-state visually evoked potential correlates of object recognition. *Brain Research*, *1343*(August 2015), 112–121. https://doi.org/10.1016/j.brainres.2010.04.072
Katyal, S., Engel, S. A., He, B., & He, S. (2016). Neurons that detect interocular conflict during binocular rivalry revealed with EEG. *Journal of Vision*, *16*(3), 18. https://doi.org/10.1167/16.3.18
Kayser, C., & Logothetis, N. K. (2007). Do early sensory cortices integrate cross-modal information? *Brain Structure and Function*, *212*(2), 121–132. https://doi.org/10.1007/s00429-007-0154-0
Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, *18*(7), 1560–1574. https://doi.org/10.1093/cercor/bhm187
Keil, J., & Senkowski, D. (2018). Neural Oscillations Orchestrate Multisensory Processing. *Neuroscientist*. https://doi.org/10.1177/1073858418755352
Keitel, C., & Müller, M. M. (2015). Audio-visual synchrony and feature-selective attention co-amplify early visual processing. *Experimental Brain Research*, 1–11.
Kennett, S., Eimer, M., Spence, C., & Driver, J. (2001). Tactile-visual links in exogenous spatial attention under different postures: convergent evidence from psychophysics and ERPs. *Journal of Cognitive Neuroscience*, *13*(4), 462–478.
Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible.' *Trends in Cognitive Sciences*, *9*(8), 381–388. https://doi.org/10.1016/j.tics.2005.06.012
Kim, Y.-J., Grabowecky, M., Paller, K. a, Muthu, K., & Suzuki, S. (2007). Attention induces synchronization-based response gain in steady-state visual evoked potentials. *Nature Neuroscience*, *10*(1), 117–125. https://doi.org/10.1038/nn1821
Kim, Y.-J., Grabowecky, M., Paller, K. a, & Suzuki, S. (2011). Differential roles of

frequency-following and frequency-doubling visual responses revealed by evoked neural harmonics. *Journal of Cognitive Neuroscience*, *23*(8), 1875–1886. https://doi.org/10.1162/jocn.2010.21536
Kim, Y.-J., Tsai, J. J., Ojemann, J., & Verghese, P. (2017). Attention to multiple objects facilitates their integration in prefrontal and parietal cortex. *J. Neurosci*, *10*(19), 2370–16. https://doi.org/10.1523/JNEUROSCI.2370-16.2017
Kim, Y.-J., & Verghese, P. (2012). The Selectivity of Task-Dependent Attention Varies with Surrounding Context. *Journal of Neuroscience*, *32*(35), 12180–12191. https://doi.org/10.1523/JNEUROSCI.5992-11.2012
King, J.-R., Sitt, J. D., Faugeras, F., Rohaut, B., El Karoui, I., Cohen, L., … Dehaene, S. (2013). Information sharing in the brain indexes consciousness in noncommunicative patients. *Current Biology*, *23*(19), 1914–1919.
Kingdom, F., & Moulden, B. (1988). Border effects on brightness: A review of findings, models and issues. *Spatial Vision*, *3*(4), 225–262.
Kleinschmidt, A., Buchel, C., Zeki, S., & Frackowiak, R. S. (1998). Human brain activity during spontaneous reversing perception of ambiguous figures. *Proceedings of the Royal Society of London B: Biological Sciences*, *265*(1413), 2427–2433.
Kleinschmidt, A., Sterzer, P., & Rees, G. (2012). *Variability of perceptual multistability: from brain state to individual trait*. 988–1000. https://doi.org/10.1098/rstb.2011.0367
Klink, P. C., Self, M. W., Lamme, V. A. F., & Roelfsema, P. R. (2015). *Theories and methods in the scientific study of consciousness*. https://doi.org/10.1075/aicr.92.02kli
Klink, P. C., van Ee, R., Nijs, M. M., Brouwer, G. J., Noest, A. J., & Van Wezel, R. J. A. (2008). Early interactions between neuronal adaptation and voluntary control determine perceptual choices in bistable vision. *Journal of Vision*, *8*(5), 16.
Klink, P. C., van Wezel, R. J. a., & van Ee, R. (2012). United we sense, divided we fail: context-driven perception of ambiguous visual stimuli. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1591), 932–941. https://doi.org/10.1098/rstb.2011.0358
Kloosterman, N. A., Meindertsma, T., van Loon, A. M., Lamme, V. A. F., Bonneh, Y. S., & Donner, T. H. (2015). Pupil size tracks perceptual content and surprise. *European Journal of Neuroscience*, *41*(8), 1068–1078. https://doi.org/10.1111/ejn.12859
Knapen, T., Brascamp, J. W., Pearson, J., van Ee, R., & Blake, R. (2011). The role of frontal and parietal brain areas in bistable perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(28), 10293–10301. https://doi.org/10.1523/JNEUROSCI.1727-11.2011
Koch, C. (2004). *The quest for consciousness*. New York.
Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). The neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. https://doi.org/10.1038/nrn.2016.22
Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences*, *11*(1), 16–22. https://doi.org/10.1016/j.tics.2006.10.012
Koenig-Robert, R., & VanRullen, R. (2013). SWIFT: A novel method to track the neural correlates of recognition. *NeuroImage*, *81*, 273–282. https://doi.org/10.1016/j.neuroimage.2013.04.116
Komatsu, H. (2006). The neural mechanisms of perceptual filling-in. *Nature Reviews.*

Neuroscience, *7*(3), 220–231. https://doi.org/10.1038/nrn1869
Komatsu, H., Kinoshita, M., & Murakami, I. (2000). Neural responses in the retinotopic representation of the blind spot in the macaque V1 to stimuli for perceptual filling-in. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *20*(24), 9310–9319. https://doi.org/20/24/9310 [pii]
Kondo, H. M., Kitagawa, N., Kitamura, M. S., Koizumi, A., Nomura, M., & Kashino, M. (2012). Separability and commonality of auditory and visual bistable perception. *Cerebral Cortex*, *22*(8), 1915–1922. https://doi.org/10.1093/cercor/bhr266
Kornmeier, J., & Bach, M. (2004). Early neural activity in Necker-cube reversal: Evidence for low-level processing of a gestalt phenomenon. *Psychophysiology*, *41*(1), 1–8. https://doi.org/10.1046/j.1469-8986.2003.00126.x
Kornmeier, J., & Bach, M. (2005). The Necker cube—an ambiguous figure disambiguated in early visual processing. *Vision Research*, *45*(8), 955–960.
Kornmeier, J., & Bach, M. (2006). Bistable perception—along the processing chain from ambiguous visual input to a stable percept. *International Journal of Psychophysiology*, *62*(2), 345–349.
Kornmeier, J., & Bach, M. (2012). Ambiguous figures - what happens in the brain when perception changes but not the stimulus. *Frontiers in Human Neuroscience*, *6*(March), 51. https://doi.org/10.3389/fnhum.2012.00051
Kornmeier, J., Ehm, W., Bigalke, H., & Bach, M. (2007). Discontinuous presentation of ambiguous figures: How interstimulus-interval durations affect reversal dynamics and ERPs. *Psychophysiology*, *44*(4), 552–560.
Kornmeier, J., Pfäffle, M., & Bach, M. (2011). Necker cube: stimulus-related (low-level) and percept-related (high-level) EEG signatures early in occipital cortex. *Journal of Vision*, *11*(9), 12.
Kovacs, I., Papathomas, T. V, Yang, M., & Fehér, Á. (1996). When the brain changes its mind: Interocular grouping during binocular rivalry. *Proceedings of the National Academy of Sciences*, *93*(26), 15508–15511.
Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540. https://doi.org/10.1038/nn.2303
Lack, L. C. (1978). *Selective attention and the control of binocular rivalry* (Vol. 11). Mouton De Gruyter.
Laing, C. R., & Chow, C. C. (2002). A spiking neuron model for binocular rivalry. *J Comput Neurosci*, *12*(1), 39–53. https://doi.org/10.1023/A:1014942129705
Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The Leading Sense: Supramodal Control of Neurophysiological Context by Attention. *Neuron*, *64*(3), 419–430. https://doi.org/10.1016/j.neuron.2009.10.014
Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences*, *7*(1), 12–18.
Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, *10*(11), 494–501. https://doi.org/10.1016/j.tics.2006.09.001
Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience*, *1*(3), 204–220. https://doi.org/10.1080/17588921003731586
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.

Landau, A. N. (2018). Neuroscience: A Mechanism for Rhythmic Sampling in Vision. *Current Biology*, Vol. 28, pp. R830–R832. https://doi.org/10.1016/j.cub.2018.05.081
Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, *22*(11), 1000–1004. https://doi.org/10.1016/j.cub.2012.03.054
Landau, A. N., Schreyer, H. M., Van Pelt, S., & Fries, P. (2015). Distributed Attention Is Implemented through Theta-Rhythmic Gamma Modulation. *Current Biology*, *25*(17), 2332–2337. https://doi.org/10.1016/j.cub.2015.07.048
Lansing, R. W. (1964). Electroencephalographic Correlates of Binocular Rivalry in Man. *Science (New York, N.Y.)*, *146*(August), 1325–1327. https://doi.org/10.1126/science.146.3649.1325
Laureys, S. (2005). The neural correlate of (un) awareness: lessons from the vegetative state. *Trends in Cognitive Sciences*, *9*(12), 556–559.
Lawwill, T., & Biersdorf, W. R. (1968). Binocular Rivalry and Visual Evoked Responses. *Investigative Opthalmology*, *7*(4), 378–385.
Lee, S.-H., & Blake, R. (1999). Rival ideas about binocular rivalry. In *Vision Research* (Vol. 39). Retrieved from www.psy.vanderbilt.edu/faculty/blake
Lee, S.-H., & Blake, R. (2004). A fresh look at interocular grouping during binocular rivalry. *Vision Research*, *44*(10), 983–991.
Leopold, D. a., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254–264. https://doi.org/10.1016/S1364-6613(99)01332-7
Levelt, W. J. M. (1965). *On binocular rivalry*. Van Gorcum Assen.
Li, H.-H., Carrasco, M., & Heeger, D. J. (2015). Deconstructing Interocular Suppression: Attention and Divisive Normalization. *PLOS Computational Biology*, *11*(10), e1004510. https://doi.org/10.1371/journal.pcbi.1004510
Li, H.-H., Rankin, J., Rinzel, J., Carrasco, M., & Heeger, D. J. (2017). Attention model of binocular rivalry. *Proceedings of the National Academy of Sciences*, 201620475. https://doi.org/10.1073/pnas.1620475114
Limbach, K., & Corballis, P. M. (2016). Prestimulus alpha power influences response criterion in a detection task. *Psychophysiology*, *53*(8), 1154–1164. https://doi.org/10.1111/psyp.12666
Lin, Z., & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, *87*(4), 195–211. https://doi.org/10.1016/j.pneurobio.2008.09.002
Lindner, M., Vicente, R., Priesemann, V., & Wibral, M. (2011). TRENTOOL: A Matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neuroscience*, *12*(1), 119.
Ling, S., & Carrasco, M. (2006). When sustained attention impairs perception. *Nature Neuroscience*, *9*(10), 1243–1245. https://doi.org/10.1038/nn1761
Llinás, R. R., & Paré, D. (1991). Of dreaming and wakefulness. *Neuroscience*, *44*(3), 521–535.
Llinás, R. R., Ribary, U., Contreras, D., & Pedroarena, C. (1998). The neuronal basis for consciousness. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*(1377), 1841–1849. https://doi.org/10.1098/rstb.1998.0336
Logothetis, N. K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *353*(1377), 1801–1818. https://doi.org/10.1098/rstb.1998.0333
Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*,

453(7197), 869–878.
Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature*, *380*(6575), 621.
Loorits, K. (2014). Structural qualia: A solution to the hard problem of consciousness. *Frontiers in Psychology*, *5*(MAR), 1–9. https://doi.org/10.3389/fpsyg.2014.00237
Lou, L. (1999). Selective peripheral fading: Evidence for inhibitory sensory effect of attention. *Perception*, *28*(4), 519–526. https://doi.org/10.1068/p2816
Luck, S. J., & Kappenman, E. S. (2011). *The Oxford handbook of event-related potential components*. Oxford university press.
Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural Correlates of Perceptual Rivalry in the Human Brain. *Science*, *280*(5371), 1930–1934. https://doi.org/10.1126/science.280.5371.1930
Lumer, E. D., & Rees, G. (1999). Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proceedings of the National Academy of Sciences*, *96*(4), 1669–1673.
Lunghi, C., & Alais, D. (2013). Touch Interacts with Vision during Binocular Rivalry with a Tight Orientation Tuning. *PLoS ONE*, *8*(3), 1–8. https://doi.org/10.1371/journal.pone.0058754
Lunghi, C., & Alais, D. (2015). Congruent tactile stimulation reduces the strength of visual suppression during binocular rivalry. *Scientific Reports*, *5*. https://doi.org/10.1038/srep09413
Lunghi, C., Binda, P., & Morrone, M. C. (2010). Touch disambiguates rivalrous perception at early stages of visual analysis. *Current Biology*, *20*(4), R143-4. https://doi.org/10.1016/j.cub.2009.12.015
Lunghi, C., Morrone, M. C., & Alais, D. (2014). Auditory and tactile signals combine to influence vision during binocular rivalry. *The Journal of Neuroscience*, *34*(3), 784–792. https://doi.org/10.1523/JNEUROSCI.2732-13.2014
Macaluso, E, Frith, C. D., & Driver, J. (2002). Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. *Neuron*, *34*(4), 647–658.
Macaluso, Emiliano, Frith, C. D., & Driver, J. (2000). Modulation of human visual cortex by crossmodal spatial attention. *Science*, *289*(5482), 1206–1208.
Maier, A., Panagiotaropoulos, T. I., Tsuchiya, N., & Keliris, G. A. (2012). Binocular rivalry: a gateway to consciousness. In *Frontiers Research Foundation*. https://doi.org/10.3389/978-2-88919-069-0
Maier, A., Wilke, M., Aura, C., Zhu, C., Ye, F. Q., & Leopold, D. A. (2008). Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. *Nature Neuroscience*, *11*(10), 1193–1200. https://doi.org/10.1038/nn.2173
Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024
Marshall, W. J., Lackner, C. L., Marriott, P., Santesso, D. L., & Segalowitz, S. J. (2014). Using phase shift granger causality to measure directed connectivity in EEG recordings. *Brain Connectivity*, *4*(10), 826–841. https://doi.org/10.1089/brain.2014.0241
Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Dyar, T. a. (2006). Microsaccades counteract visual fading during fixation. *Neuron*, *49*(2), 297–305. https://doi.org/10.1016/j.neuron.2005.11.033

Mathes, B., Strüber, D., Stadler, M. A., & Basar-Eroglu, C. (2006). Voluntary control of Necker cube reversals modulates the EEG delta-and gamma-band response. *Neuroscience Letters*, *402*(1–2), 145–149.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To See or Not to See: Prestimulus α Phase Predicts Visual Awareness. *Journal of Neuroscience*, *29*(9), 2725–2732. https://doi.org/10.1523/JNEUROSCI.3963-08.2009

McDonald, J. J., Teder-Sälejärvi, W. A., Russo, F. Di, & Hillyard, S. A. (2003). Neural Substrates of Perceptual Enhancement by Cross-Modal Spatial Attention. *Journal of Cognitive Neuroscience*, *15*(1), 10–19. https://doi.org/10.1162/089892903321107783

McEwen, C., Paton, B., Tsuchiya, N., & van Boxtel, J. J. A. (2018). *Motion-induced Blindness as a Tool to Measure Attentional Biases and the Link to Attention-deficit/hyperactivity Traits*.

Megumi, F., Bahrami, B., Kanai, R., & Rees, G. (2015a). Brain activity dynamics in human parietal regions during spontaneous switches in bistable perception. *NeuroImage*, *107*, 190–197. https://doi.org/10.1016/j.neuroimage.2014.12.018

Megumi, F., Bahrami, B., Kanai, R., & Rees, G. (2015b). Brain activity dynamics in human parietal regions during spontaneous switches in bistable perception. *NeuroImage*, *107*, 190–197.

Meng, M., Remus, D. A., & Tong, F. (2005). Filling-in of visual phantoms in the human brain. *Nature Neuroscience*, *8*(9), 1248–1254. https://doi.org/10.1038/nn1518

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, *4*(7), 539–551. https://doi.org/10.1167/4.7.2

Mercier, M. R., Foxe, J. J., Fiebelkorn, I. C., Butler, J. S., Schwartz, T. H., & Molholm, S. (2013). Auditory-driven phase reset in visual cortex: Human electrocorticography reveals mechanisms of early multisensory integration. *NeuroImage*, *79*, 19–29. https://doi.org/10.1016/j.neuroimage.2013.04.060

Meredith, M. A., & Stein, B. E. (1983). Interactions among converging sensory inputs in the superior colliculus. *Science*, *221*(4608), 389–391.

Merikle, P. M., & Joordens, S. (1997). Parallels between perception without attention and perception without awareness. *Consciousness and Cognition*, *6*(2–3), 219–236. https://doi.org/10.1006/ccog.1997.0310

Miller, S. M. (2007). On the correlation/constitution distinction problem (and other hard problems) in the scientific study of consciousness. *Acta Neuropsychiatrica*, *19*(3), 159–176. https://doi.org/10.1111/j.1601-5215.2007.00207.x

Miller, S. M. (2013). *The constitution of visual consciousness: Lessons from binocular rivalry* (Vol. 90). John Benjamins Publishing.

Miller, S. M. (2014). Closing in on the constitution of consciousness. *Frontiers in Psychology*, *5*(November), 1–18. https://doi.org/10.3389/fpsyg.2014.01293

Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience*, *27*(15), 4120–4131.

Mitchell, J. F., Stoner, G. R., & Reynolds, J. H. (2004). Object-based attention determines dominance in binocular rivalry. *Nature*, *429*(6990), 410–413. https://doi.org/10.1038/nature02584

Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory

processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, *14*(1), 115–128. https://doi.org/10.1016/S0926-6410(02)00066-6
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
Morillon, B., Schroeder, C. E., & Wyart, V. (2015). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, *5*, 1–9. https://doi.org/10.1038/ncomms6255
Morrill, R. J., & Hasenstaub, A. R. (2018). *Systems/Circuits Visual Information Present in Infragranular Layers of Mouse Auditory Cortex*. https://doi.org/10.1523/JNEUROSCI.3102-17.2018
Morrone, M. C., & Lunghi, C. (2013). Early interaction between vision and touch during binocular rivalry. *Multisensory Research*, *26*(3), 291–306. https://doi.org/10.1163/22134808-00002411
Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., & Laurienti, P. J. (2008). Modality-specific selective attention attenuates multisensory integration. *Experimental Brain Research*, *184*(1), 39–52. https://doi.org/10.1007/s00221-007-1080-3
Mozolic, J. L., Joyner, D., Hugenschmidt, C. E., Peiffer, A. M., Kraft, R. A., Maldjian, J. A., & Laurienti, P. J. (2008). Cross-modal deactivations during modality-specific selective attention. *BMC Neurology*, *8*, 1–11. https://doi.org/10.1186/1471-2377-8-35
Mudrik, L., Faivre, N., & Koch, C. (2014). Information integration without awareness. *Trends in Cognitive Sciences*, 1–9. https://doi.org/10.1016/j.tics.2014.04.009
Müller, M. M., Andersen, S., Trujillo, N. J., Valdes-Sosa, P., Malinowski, P., & Hillyard, S. A. (2006). Feature-selective attention enhances color signals in early visual areas of the human brain. *Proceedings of the National Academy of Sciences*, *103*(38), 14250–14254.
Müller, M. M., & Hubner, R. (2002). Can the spotlight of attention be shaped like a doughnut? Evidence from steady-state visual evoked potentials. *Psychological Science*, *13*(2), 119–124. https://doi.org/10.1111/1467-9280.00422
Müller, M. M., Malinowski, P., Gruber, T., & Hillyard, S. (2003). Sustained Division of the attentional spotlight. *Nature*, *424*(6946), 309–312. https://doi.org/10.1038/nature01744
Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20-28 Hz range. *Cognitive Brain Research*, *6*(4), 249–261. https://doi.org/10.1016/S0926-6410(97)00036-0
Müller, M. M., Teder-Salejarvi, W. A., & Hillyard, S. (1998). The time course of cortical facilitatin during cued shifts of spatial attention. *Nature Neuroscience*, *1*(7), 631–634.
Murray, M. M., Molholm, S., Michel, C. M., Heslenfeld, D. J., Ritter, W., Javitt, D. C., … Foxe, J. J. (2005). Grabbing your ear: Rapid auditory-somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebral Cortex*, *15*(7), 963–974. https://doi.org/10.1093/cercor/bhh197
Naber, M., Frässle, S., & Einhäuser, W. (2011). Perceptual Rivalry: Reflexes Reveal the Gradual Nature of Visual Awareness. *PLoS ONE*, *6*(6), e20910.

https://doi.org/10.1371/journal.pone.0020910
Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450.
Naue, N., Rach, S., Strüber, D., Huster, R. J., Zaehle, T., Körner, U., & Herrmann, C. S. (2011). Auditory event-related response in visual cortex modulates subsequent visual responses in humans. *Journal of Neuroscience*, *31*(21), 7729–7736.
Necker, L. A. (1832). LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *1*(5), 329–337.
New, J. J., & Scholl, B. J. (2008). "Perceptual scotomas": A functional account of motion-induced blindness: Research article. *Psychological Science*, *19*(7), 653–659. https://doi.org/10.1111/j.1467-9280.2008.02139.x
Noppeney, U., Faivre, N., Lunghi, C., Deroy, O., Spence, C., & Aller, M. (2016). The Complex Interplay Between Multisensory Integration and Perceptual Awareness. *Multisensory Research*, *29*(6–7), 585–606. https://doi.org/10.1163/22134808-00002529
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), 1–46. https://doi.org/10.1167/15.6.4.doi
Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Steady-state evoked potentials as an index of multisensory temporal binding. *NeuroImage*, *60*(1), 21–28. https://doi.org/10.1016/j.neuroimage.2011.11.065
O' Regan, J. K. (1992). Solving the Real Mysteries of Visual-Perception - the World as an Outside Memory. *Canadian Journal of Psychology-Revue Canadienne De Psychologie*, *46*(3), 461–488.
O'Donnell, B. F., Hendler, T., & Squires, N. K. (1988). Visual evoked potentials to illusory reversals of the Necker cube. *Psychophysiology*, *25*(2), 137–143.
O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(05), 939–973. https://doi.org/10.1017/S0140525X01000115
O'Shea, R. P., & Blake, R. (1986). Dichoptic temporal frequency differences do not lead to binocular rivalry. *Perception & Psychophysics*, *39*(1), 59–63. https://doi.org/10.3758/BF03207584
Odegaard, B., Knight, R. T., & Lau, H. (2017). Should A Few Null Findings Falsify Prefrontal Theories Of Conscious Perception? *The Journal of Neuroscience*, *37*(40), 9593–9602. Retrieved from http://biorxiv.org/content/early/2017/03/30/122267.abstract
Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology*, *10*(5). https://doi.org/10.1371/journal.pcbi.1003588
Ooi, T. L., & He, Z. J. (1999). Binocular rivalry and visual awareness: The role of attention. *Perception*, *28*(5), 551–574.
Ooi, T. L., & He, Z. J. (2003). A distributed intercortical processing of binocular rivalry: psychophysical evidence. *Perception*, *32*(2), 155–166.
Orbach, J., Zucker, E., & Olson, R. (1966). Reversibility of the Necker cube: VII. Reversal rate as a function of figure-on and figure-off durations. *Perceptual and Motor Skills*, *22*(2), 615–618.
Owen, A. M. (2008). Disorders of consciousness. *Annals of the New York Academy of Sciences*, *1124*, 225–238. https://doi.org/10.1196/annals.1440.013

Paffen, C. L. E., & Alais, D. (2011). Attentional Modulation of Binocular Rivalry. *Frontiers in Human Neuroscience*, *5*, 105. https://doi.org/10.3389/fnhum.2011.00105

Paffen, C. L. E., Alais, D., & Verstraten, F. A. J. (2006). Attention speeds binocular rivalry. *Psychological Science*, *17*(9), 752–756. https://doi.org/10.1111/j.1467-9280.2006.01777.x

Paffen, C. L. E., & Van der Stigchel, S. (2010). Shifting spatial attention makes you flip: Exogenous visual attention triggers perceptual alternations during binocular rivalry. *Attention, Perception, & Psychophysics*, *72*(5), 1237–1243.

Palmer, T. D., & Ramsey, A. K. (2012). The function of consciousness in multisensory integration. *Cognition*, *125*(3), 353–364. https://doi.org/10.1016/j.cognition.2012.08.003

Panagiotaropoulos, T. I., Kapoor, V., & Logothetis, N. K. (2014). Subjective visual perception: from local processing to emergent phenomena of brain activity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1641), 20130534. https://doi.org/10.1098/rstb.2013.0534

Parkkonen, L., Andersson, J., Hämäläinen, M., & Hari, R. (2008). Early visual brain areas reflect the percept of an ambiguous scene. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20500–20504. https://doi.org/10.1073/pnas.0810966105

Pastor, M. A., Artieda, J., Arbizu, J., Marti-Climent, J. M., Peñuelas, I., & Masdeu, J. C. (2002). Activation of human cerebral and cerebellar cortex by auditory stimulation at 40 Hz. In *The Journal of neuroscience : the official journal of the Society for Neuroscience* (Vol. 22). https://doi.org/22/23/10501 [pii]

Pei, F., Pettet, M. W., & Norcia, A. M. (2018). *Neural correlates of object-based attention*. (2002), 588–596.

Perlstein, W. M., Cole, M. A., Larson, M., Kelly, K., Seignourel, P., & Keil, A. (2003). Steady-state visual evoked potentials reveal frontally-mediated working memory activity in humans. *Neuroscience Letters*, *342*(3), 191–195. https://doi.org/10.1016/s0304-3940(03)00226-x

Pessoa, L, Thompson, E., & Noë, A. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *The Behavioral and Brain Sciences*, *21*(6), 723–748; discussion 748-802. https://doi.org/10.1017/S0140525X98001757

Pessoa, Luiz, Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: from modulation of sensory processing to top-down control. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(10), 3990–3998. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12764083

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2014). Linear and nonlinear mixed effects models. *R Package Version*, *3*.

Pitts, M. A., Gavin, W. J., & Nerger, J. L. (2008). Early top-down influences on bistable perception revealed by event-related potentials. *Brain and Cognition*, *67*(1), 11–24.

Pitts, M. A., Lutsyshyna, L. A., & Hillyard, S. A. (2018). The relationship between attention and consciousness: an expanded taxonomy and implications for "no-report" paradigms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(20170348). https://doi.org/10.1098/rstb.2017.0348

Plourde, G. (2006). *Auditory Evoked Potentials* (pp. 129–139). pp. 129–139. Elsevier Inc.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148.
Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience*, *3*(11), 1153.
Porcu, E., Keitel, C., & Müller, M. M. (2013). Concurrent visual and tactile steady-state evoked potentials index allocation of inter-modal attention: A frequency-tagging study. *Neuroscience Letters*, *556*, 113–117. https://doi.org/10.1016/j.neulet.2013.09.068
Porcu, E., Keitel, C., & Müller, M. M. (2014). Visual, auditory and tactile stimuli compete for early sensory processing capacities within but not between senses. *NeuroImage*, *97*(APRIL 2014), 224–235. https://doi.org/10.1016/j.neuroimage.2014.04.024
Posner, M. I. (1994). Attention: the mechanisms of consciousness. *Proceedings of the National Academy of Sciences*, *91*(16), 7398–7403. https://doi.org/10.1073/pnas.91.16.7398
Posner, M. I. (2012). Attentional networks and consciousness. *Frontiers in Psychology*, *3*, 64. https://doi.org/10.3389/fpsyg.2012.00064
Proskovec, A. L., Heinrichs-Graham, E., Wiesman, A. I., McDermott, T. J., & Wilson, T. W. (2018). Oscillatory dynamics in the dorsal and ventral attention networks during the reorienting of attention. *Human Brain Mapping*, *39*(5), 2177–2190. https://doi.org/10.1002/hbm.23997
Raij, T., Ahveninen, J., Lin, F., Witzel, T., Jääskeläinen, I. P., Letham, B., … Stufflebeam, S. (2010). Onset timing of cross-sensory activations and multisensory interactions in auditory and visual sensory cortices. *European Journal of Neuroscience*, *31*(10), 1772–1782.
Ramachandran, V. S. (1992). Filling in gaps in perception: Part I. *Current Directions in Psychological Science*, *1*(6), 199–205.
Ramachandran, V. S. (1993). Filling in gaps in perception: Part II. Scotomas and phantom limbs. *Current Directions in Psychological Science*, *2*(2), 56–65.
Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, *350*(6320), 699–702.
Ramachandran, V. S., Gregory, R. L., & Aiken, W. (1993). Perceptual fading of visual texture borders. *Vision Research*, *33*(5–6), 717–721. https://doi.org/10.1016/0042-6989(93)90191-X
Regan, D. (1977). Steady-state evoked potentials. *JOSA*, *67*(11), 1475–1489.
Reynolds, J. H., & Pasternak, T. (2000). Attention Increases Sensitivity of V4 Neurons increases the magnitude of the neuronal response elic. *Neuron*, *26*, 703–714. https://doi.org/10.1016/S0896-6273(00)81206-4
Rogers-Ramachandran, D. C., & Ramachandran, V. S. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research*, *38*(1), 71–77. https://doi.org/10.1016/S0042-6989(97)00131-4
Romei, V., Gross, J., & Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. *Current Biology*, *22*(9), 807–813. https://doi.org/10.1016/j.cub.2012.03.025
Ronconi, L., Oosterhof, N. N., Bonmassar, C., & Melcher, D. (2017). Multiple oscillatory rhythms determine the temporal organization of perception. *Proceedings of the National Academy of Sciences*, *114*(51), 1345–13440. https://doi.org/10.1073/pnas.1714522114
Ross, B., Picton, T. W., Herdman, A. T., Hillyard, S. A., & Pantev, C. (2004). The

effect of attention on the auditory steady-state response. *Neurology and Clinical Neurophysiology*, *2004*. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-29144497317&partnerID=40&md5=7327ba18e4cca38ccd3038db913da509

Roy, A. V., Jamison, K. W., He, S., Engel, S. A., & He, B. (2017). Deactivation in the posterior mid-cingulate cortex reflects perceptual transitions during binocular rivalry: Evidence from simultaneous EEG-fMRI. *NeuroImage*, *152*(October 2016), 1–11. https://doi.org/10.1016/j.neuroimage.2017.02.041

Rubin, E. (1980). *Visuell wahrgenommene figuren [visually perceived forms]*. Gyldenhal: Copenhagen, Denmark.

Safran, A. B., & Landis, T. (1996). Plasticity in the adult visual cortex: implications for the diagnosis of visual field defects and visual rehabilitation. *Current Opinion in Ophthalmology*, *7*(6), 53–64.

Sakaguchi, Y. (2001). Target/surround asymmetry in perceptual filling-in. *Vision Research*, *41*(16), 2065–2077.

Sakaguchi, Y. (2006). Contrast dependency in perceptual filling-in. *Vision Research*, *46*(20), 3304–3312. https://doi.org/10.1016/j.visres.2006.05.015

Sasaki, Y. (2007). Processing local signals into global patterns. *Current Opinion in Neurobiology*, *17*(2), 132–139. https://doi.org/10.1016/j.conb.2007.03.003

Schieting, S., & Spillman, L. (1987). Flicker adaptation in the peripheral retina. *Vision Research*, *27*(2), 277–284.

Schölvinck, M. L., & Rees, G. (2009). Attentional influences on the dynamics of motion-induced blindness. *Journal of Vision*, *9*(1), 38.1-9. https://doi.org/10.1167/9.1.38

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, "unisensory" processing. *Current Opinion in Neurobiology*, *15*(4), 454–458. https://doi.org/10.1016/j.conb.2005.06.008

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9–18. https://doi.org/10.1016/j.tins.2008.09.012

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. *Current Opinion in Neurobiology*, *20*(2), 172–176. https://doi.org/10.1016/j.conb.2010.02.010

Schurger, A., Cowey, A., Cohen, J. D., Treisman, A., & Tallon-Baudry, C. (2008). Distinct and independent correlates of attention and awareness in a hemianopic patient. *Neuropsychologia*, *46*(8), 2189–2197.

Schyns, P. G., Thut, G., & Gross, J. (2011). Cracking the code of oscillatory activity. *PLoS Biology*, *9*(5). https://doi.org/10.1371/journal.pbio.1001064

Sekar, K., Findley, W. M., Poeppel, D., & Llinás, R. R. (2013). Cortical response tracking the conscious experience of threshold duration visual stimuli indicates visual perception is all or none. *Proceedings of the National Academy of Sciences*, *110*(14), 5642–5647.

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*(10), 1391.

Sergent, C., & Naccache, L. (2012). Imaging neural signatures of consciousness:'What','When','Where'and 'How'does it work? *Archives Italiennes de Biologie*, *150*(2/3), 91–106.

Seth, A. K. (2007). Models of consciousness. *Scholarpedia*, *2*(1), 1328. https://doi.org/10.4249/scholarpedia.1328

Seth, A. K., Izhikevich, E., Reeke, G. N., & Edelman, G. M. (2006). Theories and measures of consciousness: an extended framework. *Proceedings of the National Academy of Sciences*, *103*(28), 10799–10804.
Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182–186. https://doi.org/10.1016/j.tics.2008.02.003
Silberstein, R. B., Ciorciari, J., & Pipingas, A. (1995). Steady-state visually evoked potential topography during the Wisconsin card sorting test. *Electroencephalography and Clinical Neurophysiology*, *96*(1), 24–35. https://doi.org/10.1016/0013-4694(94)00189-R
Silberstein, R. B., Harris, P. G., Nield, G. A., & Pipingas, A. (2000). Frontal steady-state potential changes predict long-term recognition memory performance. *International Journal of Psychophysiology*, *39*(1), 79–85. https://doi.org/10.1016/s0167-8760(00)00118-5
Silberstein, R. B., Schier, M. A., Pipingas, A., Ciorciari, J., Wood, S. R., & Simpson, D. G. (1990a). Steady-state visually evoked potential topography associated with a visual vigilance task. *Brain Topography*, *3*(2), 337–347. https://doi.org/10.1007/BF01135443
Silberstein, R. B., Schier, M. A., Pipingas, A., Ciorciari, J., Wood, S. R., & Simpson, D. G. (1990b). Steady-state visually evoked potential topography associated with a visual vigilance task. *Brain Topography*, *3*(2), 337–347.
Smout, C. A., & Mattingley, J. B. (2018). Spatial attention enhances the neural representation of invisible signals embedded in noise. *Journal of Cognitive Neuroscience*, *30*(8), 1119–1129. https://doi.org/doi:10.1162/jocn_a_01283
Snyder, a. Z. (1992). Steady-state vibration evoked potentials: Description of technique and characterization of responses. *Electroencephalography and Clinical Neurophysiology - Evoked Potentials*, *84*(3), 257–268. https://doi.org/10.1016/0168-5597(92)90007-X
Sobel, K. V, & Blake, R. (2002). How context influences predominance during binocular rivalry. *Perception*, *31*(7), 813–824.
Song, K., Meng, M., Chen, L., Zhou, K., & Luo, H. (2014). Behavioral Oscillations in Attention: Rhythmic α Pulses Mediated through θ Band. *Journal of Neuroscience*, *34*(14), 4837–4844. https://doi.org/10.1523/JNEUROSCI.4856-13.2014
Spaak, E., de Lange, F. P., & Jensen, O. (2014). Local Entrainment of Alpha Oscillations by Visual Stimuli Causes Cyclic Modulation of Perception. *Journal of Neuroscience*, *34*(10), 3536–3544. https://doi.org/10.1523/JNEUROSCI.4385-13.2014
Spillmann, L., & De Weerd, P. (2003). Mechanisms of surface completion: Perceptual filling-in of texture. *Filling-in: From Perceptual Completion to Cortical Reorganization*, 81–105.
Spillmann, L., Otte, T., Hamburger, K., & Magnussen, S. (2006). Perceptual filling-in from the edge of the blind spot. *Vision Research*, *46*(25), 4252–4257. https://doi.org/10.1016/j.visres.2006.08.033
Spitzer, H., Desimone, R., & Moran, J. (1988). *Increased Attention Enhances Both Behavioral and Neuronal Performance*. *240*(4850), 338–340.
Spong, P., Haider, M., & Lindsley, D. B. (1965). Selective attentiveness and cortical evoked responses to visual and auditory stimuli. *Science*, *148*(3668), 395–397.
Sporns, O., Gally, J. A., Reeke, G. N., & Edelman, G. M. (1989). Reentrant signaling among simulated neuronal groups leads to coherency in their oscillatory activity. *Proceedings of the National Academy of Sciences*, *86*(18), 7265–7269.

Squires, K. C., Hillyard, S. A., & Lindsay, P. H. (1973a). Cortical potentials evoked by confirming and disconfirming feedback following an auditory discrimination. *Perception & Psychophysics*, *13*(1), 25–31.
Squires, K. C., Hillyard, S. A., & Lindsay, P. H. (1973b). Vertex potentials evoked during auditory signal detection: Relation to decision criteria. *Perception & Psychophysics*, *14*(2), 265–272.
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
Srinivasan, R., Fornari, E., Knyazeva, M. G., Meuli, R., & Maeder, P. (2007). fMRI responses in medial frontal cortex that depend on the temporal frequency of visual input. *Experimental Brain Research*, *180*(4), 677–691.
Srinivasan, R., & Petrovic, S. (2006). MEG Phase Follows Conscious Perception during Binocular Rivalry Induced by Visual Stream Segregation. *Cerebral Cortex*, *16*(5), 597–608. https://doi.org/10.1093/cercor/bhj016
Srinivasan, R., Russell, D. P., Edelman, G. M., & Tononi, G. (1999). Increased Synchronization of Neuromagnetic Responses during Conscious Perception. *The Journal of Neuroscience*, *19*(13), 5435–5448. https://doi.org/10.1523/JNEUROSCI.19-13-05435.1999
Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255–266.
Sterzer, P., & Kleinschmidt, A. (2007). A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences*, *104*(1), 323–328. https://doi.org/10.1073/pnas.0609006104
Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, *13*(7), 310–318. https://doi.org/10.1016/j.tics.2009.04.006
Sterzer, P., Stein, T., Ludwig, K., Rothkirch, M., & Hesselmann, G. (2014). Neural processing of visual information under interocular suppression: A critical review. *Frontiers in Psychology*, *5*(MAY), 1–12. https://doi.org/10.3389/fpsyg.2014.00453
Stevenson, R. J., & Mahmut, M. K. (2013). Detecting olfactory rivalry. *Consciousness and Cognition*, *22*(2), 504–516.
Storm, J. F., Boly, M., Casali, A. G., Massimini, M., Olcese, U., Pennartz, C. M. A., & Wilke, M. (2017). Consciousness Regained: Disentangling Mechanisms, Brain Systems, and Behavioral Responses. *The Journal of Neuroscience*, *37*(45), 10882–10893. https://doi.org/10.1523/JNEUROSCI.1838-17.2017
Strüber, D., Baar-Eroglu, C., Miener, M., & Stadler, M. (2001). EEG gamma-band response during the perception of Necker cube reversals. *Visual Cognition*, *8*(3–5), 609–621.
Strüber, D., & Herrmann, C. S. (2002). MEG alpha activity decrease reflects destabilization of multistable percepts. *Cognitive Brain Research*, *14*(3), 370–382.
Stürzel, F., & Spillmann, L. (2001). Texture fading correlates with stimulus salience. *Vision Research*, *41*(23), 2969–2977. https://doi.org/10.1016/S0042-6989(01)00172-9
Sutoyo, D., & Srinivasan, R. (2009). Nonlinear SSVEP responses are sensitive to the perceptual binding of visual hemifields during conventional “eye” rivalry and

interocular “percept” rivalry. *Brain Research*, *1251*, 245–255. https://doi.org/10.1016/j.brainres.2008.09.086
Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188.
Tagliazucchi, E., von Wegner, F., Morzelewski, A., Brodbeck, V., Jahnke, K., & Laufs, H. (2013). Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep. *Proceedings of the National Academy of Sciences*, *110*(38), 15419–15424.
Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *Front Biosci*, *14*, 321–332.
Tallon-Baudry, C. (2012). On the neural mechanisms subserving consciousness and attention. *Frontiers in Psychology*, *3*(JAN), 1–11. https://doi.org/10.3389/fpsyg.2011.00397
Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences*, *3*(4), 151–162.
Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, *14*(9), 400–410. https://doi.org/10.1016/j.tics.2010.06.008
Tanaka, K., Kuriki, S., Nemoto, I., & Uchikawa, Y. (2013). Auditory Steady-State Responses in Magnetoencephalogram and Electroencephalogram: Phenomena, Mechanisms, and Applications. *Advanced Biomedical Engineering*, *2*, 55–62. https://doi.org/10.1016/0269-7491(91)90075-8
ten Oever, S., Romei, V., van Atteveldt, N., Soto-Faraco, S., Murray, M. M., & Matusz, P. J. (2016). The COGs (context, object, and goals) in multisensory processing. *Experimental Brain Research*, *234*(5), 1307–1323. https://doi.org/10.1007/s00221-016-4590-z
Thomas, V., Davidson, M., Zakavi, P., Tsuchiya, N., & van Boxtel, J. J. A. (2017). Simulated forward and backward self motion, based on realistic parameters, causes motion induced blindness. *Scientific Reports*, *7*(1), 1–14. https://doi.org/10.1038/s41598-017-09424-6
Thorne, J. D., De Vos, M., Viola, F. C., & Debener, S. (2011). Cross-Modal Phase Reset Predicts Auditory Task Performance in Humans. *Journal of Neuroscience*, *31*(10), 3853–3861. https://doi.org/10.1523/JNEUROSCI.6176-10.2011
Thorne, J. D., & Debener, S. (2014). Look now and hear what’s coming: On the functional role of cross-modal phase reset. *Hearing Research*, *307*, 144–152. https://doi.org/10.1016/j.heares.2013.07.002
Tiitinen, H. T., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Näätänen, R. (1993). Selective attention enhances the auditory 40-Hz transient response in humans. *Nature*, *364*(6432), 59.
Tlumak, A. I., Durrant, J. D., Delgado, R. E., & Boston, J. R. (2012). Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: profile in adults. *International Journal of Audiology*, *50*(7), 448–458. https://doi.org/10.3109/14992027.2011.560903
Tlumak, A. I., Rubinstein, E., & Durrant, J. D. (2007). Meta-analysis of variables that affect accuracy of threshold estimation via measurement of the auditory steady-state response (ASSR). *International Journal of Audiology*, *46*(11), 692–710. https://doi.org/10.1080/14992020701482480
Tomassini, A., Ambrogioni, L., Medendorp, W. P., & Maris, E. (2017). Theta oscillations locked to intended actions rhythmically modulate perception. *ELife*, *6*(e25618). https://doi.org/10.7554/eLife.25618

Tomassini, A., Spinelli, D., Jacono, M., Sandini, G., & Morrone, M. C. (2015). Rhythmic Oscillations of Visual Contrast Sensitivity Synchronized with Action. *Journal of Neuroscience*, *35*(18), 7019–7029. https://doi.org/10.1523/JNEUROSCI.4568-14.2015
Tomko, G. J., & Crapper, D. R. (1974). Neuronal variability: non-stationary responses to identical visual stimuli. *Brain Research*, *79*(3), 405–418.
Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends in Cognitive Sciences*, *10*(11), 502–511. https://doi.org/10.1016/j.tics.2006.09.003
Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, *5*(1), 42.
Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, *215*(3), 216–242. https://doi.org/215/3/216 [pii]
Tononi, G. (2012). Integrated Information Theory of Consciousness: An Updated Account. *Arch Ital Biol*, *150*, 56–90. https://doi.org/10.4449/aib.v149i5.1388
Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated Information Theory: from consciousness to its physical substrates. *Nature Reviews Neuroscience*, *in press*. https://doi.org/10.1038/nrn.2016.44
Tononi, G., & Edelman, G. M. (1998). Consciousness and complexity. *Science*, *282*(5395), 1846–1851.
Tononi, G., Edelman, G. M., & Sporns, O. (1998). Complexity and coherency: integrating information in the brain. *Trends in Cognitive Sciences*, *2*(12), 474–484.
Tononi, G., & Koch, C. (2015). Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 1–18. https://doi.org/10.1038/425107b
Tononi, G., & Sporns, O. (2003). Measuring information integration. *BMC Neuroscience*, *4*(1), 31.
Tononi, G., Srinivasan, R., Russell, D. P., & Edelman, G. M. (1998a). Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *Proceedings of the National Academy of Sciences*, *95*(6), 3198–3203. https://doi.org/10.1073/pnas.95.6.3198
Tononi, G., Srinivasan, R., Russell, D. P., & Edelman, G. M. (1998b). Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *Proceedings of the National Academy of Sciences*, *95*(6), 3198–3203. https://doi.org/10.1073/pnas.95.6.3198
Troxler, D. (1804). Ueber das Verschwinden gegebener Gegenstande innerhalb unseres Gesichtskreises. *Ophthalmologische Bibliothek*, *2*, 1–119.
Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*(8), 1096–1101. https://doi.org/10.1167/4.8.61
Tsuchiya, N., Wilke, M., Frassle, S., & Lamme, V. A. F. (2015). No-report paradigms: Extracting the true neural correlates of consciousness. *TICS*. https://doi.org/10.1017/CBO9781107415324.004
Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends in Cognitive Sciences*, *19*(12), 757–770. https://doi.org/10.1016/j.tics.2015.10.002
Van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: Flexible use of general operations. *Neuron*, *81*(6), 1240–1253. https://doi.org/10.1016/j.neuron.2014.02.044
van Boxtel, J. J. A., Alais, D., Erkelens, C. J., & van Ee, R. (2008). The Role of

Temporally Coarse Form Processing during Binocular Rivalry. *PLoS ONE*, *3*(1), e1429. https://doi.org/10.1371/journal.pone.0001429
van Boxtel, J. J. A., Knapen, T., Erkelens, C. J., & van Ee, R. (2008). Removal of monocular interactions equates rivalry behavior for monocular, binocular, and stimulus rivalries. *Journal of Vision*, *8*(15), 13–13. https://doi.org/10.1167/8.15.13
van Boxtel, J. J. A., & Tsuchiya, N. (2014). *De-confounding the neural constitution of phenomenal consciousness from attention, report and memory*.
van Boxtel, J. J. A., Tsuchiya, N., & Koch, C. (2010a). Consciousness and Attention: On Sufficiency and Necessity. *Frontiers in Psychology*, *1*(DEC), 1–13. https://doi.org/10.3389/fpsyg.2010.00217
van Boxtel, J. J. A., Tsuchiya, N., & Koch, C. (2010b). Opposing effects of attention and consciousness on afterimages. *Proceedings of the National Academy of Sciences*, *107*(19), 8883–8888. https://doi.org/10.1073/pnas.0913292107
van Boxtel, J. J. A., van Ee, R., & Erkelens, C. J. (2007). Dichoptic masking and binocular rivalry share common perceptual dynamics. *Journal of Vision*, *7*(14), 3. https://doi.org/10.1167/7.14.3
van Driel, J., Knapen, T., van Es, D. M., & Cohen, M. X. (2014). Interregional alpha-band synchrony supports temporal cross-modal integration. *NeuroImage*, *101*, 404–415. https://doi.org/10.1016/j.neuroimage.2014.07.022
van Ee, R., van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *29*(37), 11641–11649. https://doi.org/10.1523/JNEUROSCI.0873-09.2009
VanRullen, R. (2013). Visual attention: A rhythmic process? *Current Biology*, *23*(24), R1110–R1112. https://doi.org/10.1016/j.cub.2013.11.006
VanRullen, R. (2016a). Perceptual cycles. *Trends in Cognitive Sciences*, *20*(10), 723–735. https://doi.org/10.1167/15.12.1401
VanRullen, R. (2016b). Perceptual Rhythms. *Stevens' Handbook of Experimental Psychology*, 1–51.
VanRullen, R. (2018). Attention Cycles. *Neuron*, *99*(4), 632–634. https://doi.org/10.1016/j.neuron.2018.08.006
VanRullen, R., Carlson, T., & Cavanagh, P. (2007). The blinking spotlight of attention. *Proceedings of the National Academy of Sciences*, *104*(49), 19204–19209. https://doi.org/10.1073/pnas.0707316104
VanRullen, R., & Dubois, J. (2011). The psychophysics of brain rhythms. *Frontiers in Psychology*, *2*(AUG), 1–10. https://doi.org/10.3389/fpsyg.2011.00203
Varela, F. (1994). Resonant cell assemblies: a new approach to cognitive functions and neuronal synchrony. *Biological Research*, *28*(1), 81–95.
Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, *2*(4), 229–239.
Vialatte, F. B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives. *Progress in Neurobiology*, *90*(4), 418–438. https://doi.org/10.1016/j.pneurobio.2009.11.005
Von Helmholtz, H. (1867). *Handbuch der physiologischen Optik* (Vol. 9). Leipzig: Leopold Voss.
Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony. *Attention, Perception & Psychophysics*, *72*(4), 871–884. https://doi.org/10.3758/APP

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. https://doi.org/10.1037/a0029333

Wales, R., & Fox, R. (1970). Increment detection thresholds during binocular rivalry suppression. *Perception & Psychophysics*, *8*(2), 90–94.

Walter, S., Quigley, C., Andersen, S. K., & Mueller, M. M. (2012). Effects of overt and covert attention on the steady-state visual evoked potential. *Neuroscience Letters*, *519*(1), 37–41. https://doi.org/10.1016/j.neulet.2012.05.011

Wang, J., Clementz, B. A., & Keil, A. (2007). The neural correlates of feature-based selective attention when viewing spatially and temporally overlapping images. *Neuropsychologia*, *45*(7), 1393–1399. https://doi.org/10.1016/j.neuropsychologia.2006.10.019

Watanabe, M., Cheng, K., Murayama, Y., Ueno, K., Asamizuya, T., Tanaka, K., & Logothetis, N. K. (2011). Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. *Science*, *334*(6057), 829–831.

Weil, R. S., Kilner, J. M., Haynes, J. D., & Rees, G. (2007). Neural correlates of perceptual filling-in of an artificial scotoma in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 5211–5216. https://doi.org/10.1073/pnas.0609294104

Weil, R. S., & Rees, G. (2011). A new taxonomy for perceptual filling-in. *Brain Research Reviews*, *67*(1–2), 40–55. https://doi.org/10.1016/j.brainresrev.2010.10.004

Weil, R. S., Wykes, V., Carmel, D., & Rees, G. (2012). Opposite effects of perceptual and working memory load on perceptual filling-in of an artificial scotoma. *Cognitive Neuroscience*, *3*(1), 36–44. https://doi.org/10.1080/17588928.2011.603829

Weilnhammer, V. A., Ludwig, K., Hesselmann, G., & Sterzer, P. (2013). Frontoparietal cortex mediates perceptual transitions in bistable perception. *The Journal of Neuroscience*, *33*(40), 16009–16015.

Wheatstone, C. (1838). XVIII. Contributions to the physiology of vision.- Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, *128*, 371–394.

Whittle, P. (1965). Binocular rivalry and the contrast at contours. *Quarterly Journal of Experimental Psychology*, *17*(3), 217–226.

Wiener, M., & Kanai, R. (2016). Frequency tuning for temporal perception and prediction. *Current Opinion in Behavioral Sciences*, *8*, 1–6.

Wilson, H. R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(24), 14499–14503. https://doi.org/10.1073/pnas.2333622100

Wilson, H. R., Blake, R., & Lee, S.-H. (2001). Dynamics of travelling waves in visual perception. *Nature*, *412*(6850), 907.

Winter, B. (2013). *A very basic tutorial for performing linear mixed effects analyses*. (Tutorial 2), 1–22.

Wu, Z., & Yao, D. (2007). The Influence of Cognitive Tasks on Different Frequencies Steady-state Visual Evoked Potentials. *Brain Topography*, *20*(2), 97–104. https://doi.org/http://dx.doi.org/10.1007/s10548-007-0035-0

Wyart, V., Dehaene, S., & Tallon-Baudry, C. (2012). Early dissociation between

neural signatures of endogenous spatial attention and perceptual awareness during visual masking. *Frontiers in Human Neuroscience*, *6*(February), 1–14. https://doi.org/10.3389/fnhum.2012.00016
Wyart, V., & Tallon-Baudry, C. (2008). Neural dissociation between visual awareness and spatial attention. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(10), 2667–2679. https://doi.org/10.1523/JNEUROSCI.4748-07.2008
Wyart, V., & Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *Journal of Neuroscience*, *29*(27), 8715–8725.
Yoshor, D., Bosking, W. H., Ghose, G. M., & Maunsell, J. H. R. (2007). Receptive fields in human visual cortex mapped with surface electrodes. *Cerebral Cortex*, *17*(10), 2293–2302. https://doi.org/10.1093/cercor/bhl138
Yuval-Greenberg, S., & Deouell, L. Y. (2009). The broadband-transient induced gamma-band response in scalp EEG reflects the execution of saccades. *Brain Topography*, *22*(1), 3–6.
Yuval-Greenberg, S., & Heeger, D. J. (2013). Continuous flash suppression modulates cortical activity in early visual cortex. *Journal of Neuroscience*, *33*(23), 9635–9643.
Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., & Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron*, *58*(3), 429–441.
Zaretskaya, N., Thielscher, A., Logothetis, N. K., & Bartels, A. (2010). Disrupting parietal function prolongs dominance durations in binocular rivalry. *Current Biology*, *20*(23), 2106–2111. https://doi.org/10.1016/j.cub.2010.10.046
Zeki, S. (2003). The disunity of consciousness. *Trends in Cognitive Sciences*, *7*(5), 214–218.
Zemon, V., & Ratliff, F. (1984). Intermodulation components of the visual evoked potential: responses to lateral and superimposed stimuli. *Biological Cybernetics*, *50*(6), 401–408.
Zhang, D., Hong, B., Gao, X., Gao, S., & Röder, B. (2011). Exploring steady-state visual evoked potentials as an index for intermodal and crossmodal spatial attention. *Psychophysiology*, *48*(5), 665–675. https://doi.org/10.1111/j.1469-8986.2010.01132.x
Zhang, H., Morrone, M. C., & Alais, D. (2019). Behavioural oscillations in visual orientation discrimination reveal distinct modulation rates for both sensitivity and response bias. *Scientific Reports*, *9*(1), 1115.
Zhang, P., Jamison, K., Engel, S., He, B., & He, S. (2011). Binocular Rivalry Requires Visual Attention. *Neuron*, *71*(2), 362–369. https://doi.org/10.1016/j.neuron.2011.05.035
Zhou, W., & Chen, D. (2009). Binaral rivalry between the nostrils and in the cortex. *Current Biology*, *19*(18), 1561–1565.
Zhou, W., Jiang, Y., He, S., & Chen, D. (2010). Olfaction modulates visual perception in binocular rivalry. *Current Biology*, *20*(15), 1356–1358.
Zoefel, B., & VanRullen, R. (2017). Oscillatory mechanisms of stimulus processing and selection in the visual and auditory systems: State-of-the-art, speculations and suggestions. *Frontiers in Neuroscience*, *11*, 296. https://doi.org/10.3389/fnins.2017.00296
Zou, J., He, S., & Zhang, P. (2016). *Binocular rivalry from invisible patterns*. https://doi.org/10.1073/pnas.1604816113

Appendix 1: Simulated forward and backward self-motion, based on realistic parameters, causes motion induced blindness

A.1. Article Introduction

The following article was completed during candidature, requiring time to design the software for stimulus presentation, provide supervision for data analysis, and preparation of the manuscript for publication. This article has previously been submitted for consideration at Monash University for the award of Honours, by first author V. Thomas, and thus has not been included as an empirical chapter in this dissertation.

Although not included in the narrative of this thesis, this motion-induced-blindness investigation helped to form the ideas that were central to the second empirical axis of this thesis. Specifically, we here examine peripheral disappearances over a moving stimulus background, testing targets in each of the four quadrants independently. After the success of this paradigm to induce peripheral disappearances, my investigations into the combination of PFI and frequency-tagging began in earnest.

A.1.1. Article access

In the interest of space and paper waste, this article has not been reproduced here. The article is accessible online:
https://www.ncbi.nlm.nih.gov/pubmed/28851914

For any assistance in accessing this or other articles, please don't hesitate to contact the author.