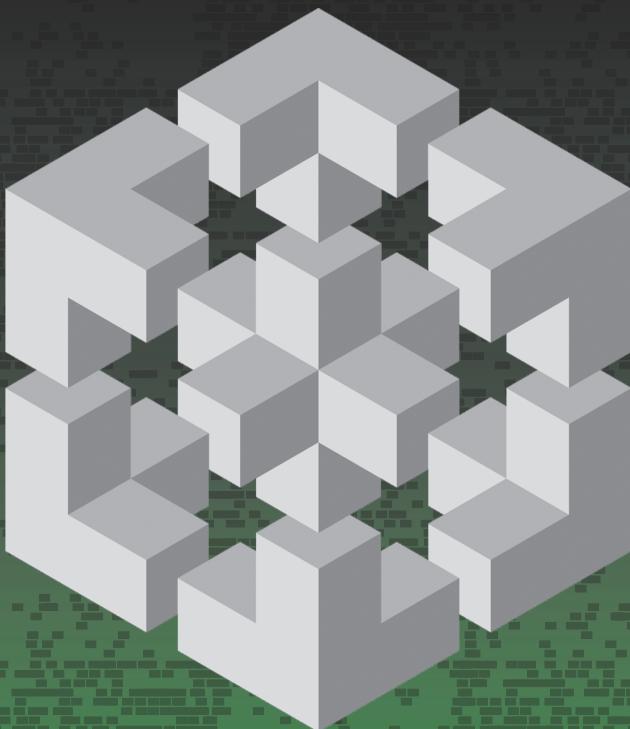
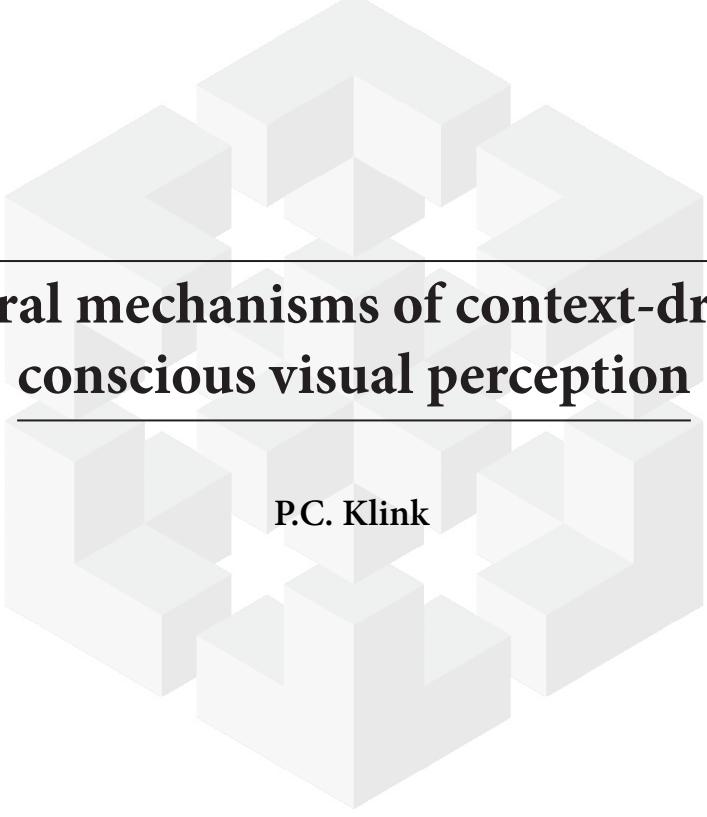


# Neural mechanisms of context-driven conscious visual perception



P. Christiaan Klink





---

# **Neural mechanisms of context-driven conscious visual perception**

---

P.C. Klink

**Reading committee**

Prof. dr. C.J. Erkelens

Utrecht University, the Netherlands

Prof. dr. D.A. Leopold

National Institute of Mental Health, USA

Prof. dr. A.J. van Opstal

Radboud University Nijmegen, the Netherlands

Prof. dr. P.R. Roelfsema

The Netherlands Institute for Neuroscience, the Netherlands

Prof. dr. F.A.J. Verstraten

Utrecht University, the Netherlands



**Universiteit Utrecht**

**Neural mechanisms of context-driven conscious visual perception**

PhD thesis, Functional Neurobiology & Helmholtz Institute

Department of Biology, Faculty of Science, Utrecht University

ISBN/EAN: 978-90-889-1225-2

Cover design: Peter Christiaan Klink

Printed by: Proefschriftmaken.nl || Printyourthesis.com

Published by: Uitgeverij BOXPress, Oisterwijk

Electronically published by: University Library of Utrecht University

© 2011. All rights reserved. No parts of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of the author.

# **Neural mechanisms of context-driven conscious visual perception**

**Neurale mechanismen van contextafhankelijke  
bewuste visuele waarneming**

(met een samenvatting in het Nederlands)

## **Proefschrift**

ter verkrijging van de graad van doctor aan de Universiteit Utrecht  
op gezag van de rector magnificus, prof. dr. J.C. Stoof, ingevolge het  
besluit van het college voor promoties in het openbaar te verdedigen  
op woensdag 12 januari 2011 des middags te 4.15 uur

door

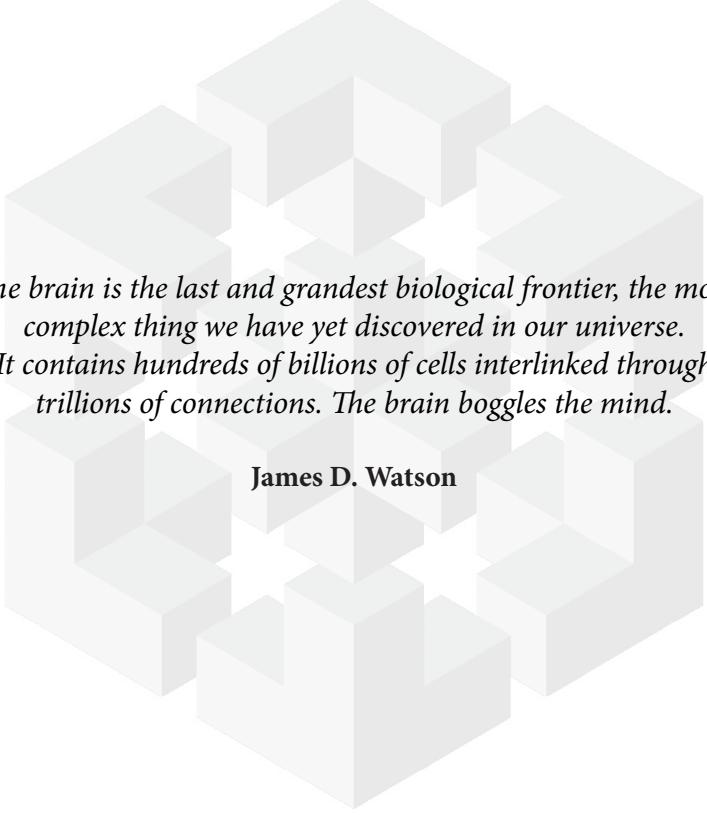
**Peter Christiaan Klink**

geboren op 8 oktober 1980  
te Papendrecht

**Promotoren:** Prof. dr. R.J.A. van Wezel  
Prof. dr. A.V. van den Berg

**Co-promotoren:** Dr. ir. M.J.M. Lankheet  
Dr. R. van Ee

Het in dit proefschrift beschreven onderzoek werd mede mogelijk gemaakt door financiële steun van het High Potential programma van de Universiteit Utrecht.



*The brain is the last and grandest biological frontier, the most complex thing we have yet discovered in our universe. It contains hundreds of billions of cells interlinked through trillions of connections. The brain boggles the mind.*

**James D. Watson**





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	General introduction .....	3
1.2	Neurons, the bits that brains are made of .....	4
1.3	From phrenology to functional specialization .....	5
1.4	Visual cortex: a brief introduction .....	6
1.5	Why study vision?.....	8
1.6	Ambiguous stimuli expose the route from photon to percept .....	9
1.7	Context resolves visual ambiguities .....	12
<b>2</b>	<b>General validity of Levelt's propositions reveals common computational mechanisms for visual rivalry</b>	<b>15</b>
2.1	Abstract .....	17
2.2	Introduction .....	17
2.3	Materials and methods.....	19
2.4	Results .....	20
2.5	Discussion .....	25
2.6	Acknowledgements .....	29
<b>3</b>	<b>Early interactions between neuronal adaptation &amp; voluntary control determine perceptual choices in bistable vision</b>	<b>31</b>
3.1	Abstract .....	33
3.2	Introduction .....	33
3.3	Experiment 1: Percept-choices and stimulus timing.....	35
3.4	Methods .....	36
3.5	Results .....	39
3.6	Experiment 2: Percept-choices and voluntary control .....	41
3.7	Methods .....	42



3.8 Results .....	42
3.9 Eye movements .....	46
3.10 Discussion .....	47
3.11 Acknowledgements .....	50
3.12 Appendix .....	51
<b>4 Intermittent stimulus presentation stabilizes neuronal responses in macaque area MT</b>	<b>57</b>
4.1 Abstract .....	59
4.2 Introduction .....	59
4.3 Results and discussion .....	61
4.4 Conclusions .....	71
4.5 Experimental procedures .....	71
4.6 Acknowledgements .....	74
<b>5 Occlusion-related lateral connections stabilize kinetic depth stimuli through perceptual coupling</b>	<b>75</b>
5.1 Abstract .....	77
5.2 Introduction .....	77
5.3 A neural network model .....	80
5.4 Methods .....	87
5.6 Discussion .....	93
5.7 Acknowledgements .....	97
5.7 Appendix 1: Depth cues in a single kinetic depth cylinder .....	98
5.8 Appendix 2: Perceptual coupling between kinetic depth cylinders .....	99
<b>6 Crossmodal duration perception involves perceptual grouping, temporal ventriloquism &amp; variable internal clock rates</b>	<b>101</b>
6.1 Abstract .....	103
6.2 Introduction .....	103
6.3 General methods .....	106
6.4 Experiment 1: Asymmetric audiovisual distortions in duration perception .....	107
6.5 Experiment 2: The need for crossmodal on- and offset proximity .....	111
6.6 Experiment 3: Auditory distortions of visual duration: A parametric approach .....	112
6.7 Experiment 4: The relative rates of the pacemaker .....	116
6.8 Experiment 5: Intramodal grouping prevents crossmodal duration effects .....	119
6.9 General discussion .....	121
6.10 Acknowledgements .....	126
<b>7 Experience-driven plasticity in binocular vision</b>	<b>127</b>
7.1 Abstract .....	129



---

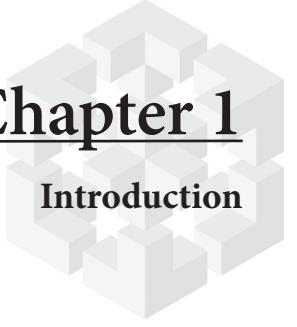
7.2 Results .....	129
7.3 Discussion .....	136
7.4 Experimental procedure .....	138
7.5 Acknowledgements .....	138
7.6 Supplemental experimental procedure .....	139
7.7 Supplemental information 1: Modeling data .....	141
7.8 Supplemental information 2: The 24-hour experiment .....	142
7.9 Supplemental information 3: Superimposition vs. piecemeal .....	143
<b>8 Appendix: Some spikes are more informative than others</b>	<b>147</b>
8.1 Acknowledgements .....	151
<b>9 Summary &amp; General discussion</b>	<b>153</b>
9.1 Summary & General discussion .....	154
9.1.1 <i>Chapter 2: General rivalry mechanisms</i> .....	154
9.1.2 <i>Chapter 3: Temporal context and voluntary control</i> .....	155
9.1.3 <i>Chapter 4: Neurophysiology of temporal context</i> .....	156
9.1.4 <i>Chapter 5: Spatial context</i> .....	156
9.1.5 <i>Chapter 6: Crossmodal context</i> .....	157
9.1.6 <i>Chapter 7: Plasticity</i> .....	157
9.2 Reaching through a window on consciousness .....	158
9.3 Studying inference at the network level .....	161
9.4 Vision in the laboratory .....	162
9.5 Learning the language of neurons .....	163
9.6 In conclusion .....	163
<b>Samenvatting (Summary in Dutch)</b>	<b>165</b>
<b>Bibliography</b>	<b>171</b>
<b>Acknowledgements</b>	<b>195</b>
<b>Publications</b>	<b>201</b>
Journal papers .....	202
Conference abstracts .....	203
Book chapters .....	204
Other publications .....	204
<b>Curriculum Vitae</b>	<b>205</b>



---

# Chapter 1

## **Introduction**



*All our knowledge is the offspring of our perceptions.*

**Leonardo da Vinci**



*Several years into my PhD research, I suddenly realized that I had learned quite a lot since I started. The chapters in this thesis represent a rather discrete sampling of my PhD-studentship. Consequently, they merely provide the reader with the end products of a specific subset of ideas that happened to work out quite nicely. There is, however, a lot more to the story of how this thesis was eventually written. To fill in some of the blanks, I will add a little context to each chapter, by means of a short preface, written in italics (similar to this paragraph). I shamelessly stole this idea, with permission, from my former colleague Jan Brascamp, who introduced it in his thesis in 2008. The prefaces to each chapter may contain a reflection on how I think the work fits into the literature, an assessment of the things I may have learned from the experimental process rather than from the data, a simple anecdote on how the idea for the experiment was conceived, a description of the inevitable problems we ran into while collecting the data, or a recollection of one of those sacred moments when a ‘huh?’ finally turned into an ‘ooooh’ and we began to understand the meaning of our results. I hope that these little bits of highly unscientific content can add something extra to the reading experience.*

---



## 1.1 General introduction

Neuroscience has come a long way. When, in approximately 400 BC, Hippocrates first considered the brain to be the organ from which sensations, emotions and intellect arise<sup>1</sup>, he could probably not have imagined that 21<sup>st</sup> century neuroscientists would be able to decode and predict these sensations, emotions, and decisions, merely by looking at patterns of recorded brain activity (Haynes & Rees, 2006; Quian Quiroga & Panzeri, 2009). Unfortunately, the mere fact that one is able to predict the outcome of a process does not necessarily mean that one also understands how this process works. In a sense, current ‘decoding’ approaches are very similar to predicting tomorrow’s weather, based on a thorough analysis of the recent dynamics in a large set of parameters that describe the current and recent weather patterns. While such analysis is very useful in deciding whether to harvest your crops today or tomorrow, it will tell you very little about how weather really works.

Whereas this is a crass oversimplification of the potential of novel decoding techniques, it does illustrate that with all technological progress that has been made over the past 2,400 years, we still know remarkably little about how the brain actually works. It may not be surprising that this is not a simple puzzle to solve, provided that there are roughly 100 billion neurons in a human brain. These neurons communicate with each other via, on average, ten thousand synaptic connections each. And to make things even more complex, these connections between neurons are not static, but constantly change in communicative strength and precise connectivity as a result of the recent activity patterns of the involved neurons.

Such a single neuron is a complex little biological machine of its own, whose intracellular communication employs eloquent molecular mechanisms and whose structure and function are blueprinted in the base pairs of its DNA. Between the order of base pairs in my own DNA and my conscious(?) decision to compare brain activity decoding with weather forecasting in the first paragraph of my thesis, there is more than enough uncovered territory to keep neuroscience busy for at least another several thousand years.

The experiments that are described in this thesis focus on a small, yet versatile, problem of brain functioning: conscious visual perception. The many advantages of studying vision in order to gain knowledge about general brain functions will be explained later in this introduction. First, the lay reader will be provided with some background information about the organization of the brain in general and the visual system in particular. These sections will be very concise, providing merely the most basic information that is necessary for a better understanding of the research chapters in this thesis<sup>2</sup>. It is unlikely that any experienced neuroscientist will find novel information in these sections and it is recommended that they skip forwards to the subheading ‘Why study Vision?’. Here one can find a motivation for studying vision and an explanation of the types of stimuli that are most often used in the research described in this thesis. This introductory chapter ends with a brief overview of the research topics presented in the different chapters of this thesis.

<sup>1</sup> “Men ought to know that from the brain and from the brain only arise our pleasures, joys, laughter, and jests as well as our sorrows, pains, griefs and tears. (...) It is the same thing which makes us mad or delirious, inspires us with dread and fear, whether by night or by day, brings us sleeplessness, inopportune mistakes, aimless anxieties, absent-mindedness and acts that are contrary to habit” - Hippocrates.

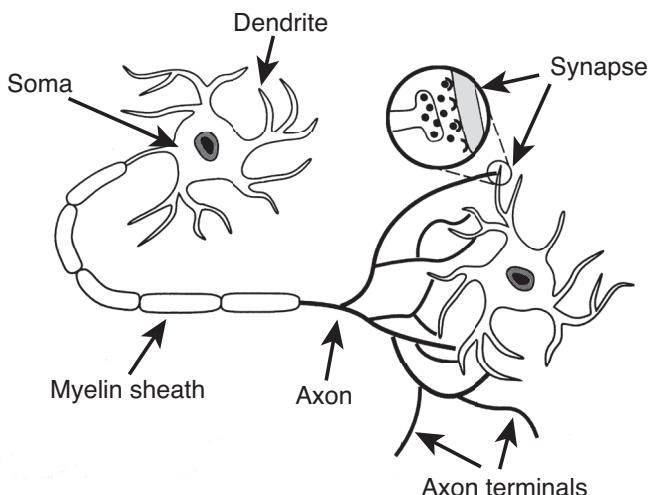
<sup>2</sup> More detailed background information about neuroscience in general can be found in the excellent textbooks *Principles of Neural Science* (Kandel et al.) and *Neuroscience* (Purves et al.). An extensive description of structure and functioning of the visual system can be found in *Vision Science* (Palmer) or *Perception* (Sekuler and Blake).



## 1.2 Neurons, the bits that brains are made of

Neurons are electrically excitable cells that specialize in receiving, integrating, and propagating information, both chemically and electrically. They are the main functional components of the brain and the smallest computational units that will be considered here. A neuron can be functionally subdivided into three parts (Figure 1-1). At the receiving end are the dendrites, thin tree-like structures that collect chemical signals from other cells and translate these chemical signals into graded electric activity. This activity is called the membrane potential since it is based on the electrical difference between the inner and outer side of the cell's membrane. When the membrane potential crosses a certain threshold level, the cell body (or soma) converts this analogue signal into a discrete series of stereotypical impulses termed action potentials, or spikes. These action potentials then propagate away from the soma over the axon, a long, thin projection that forms the emitting part of the neuron. The axons of most neurons are covered with a myelin sheath, an isolation material that ensures a fast conduction of action potentials. At the far end of the axon (the terminal), action potentials are converted back into chemical signals when chemical substances called neurotransmitters are released from the neuron into the extracellular space. Usually, the terminals of one neuron's axon are close to another neuron's dendrite. The small extracellular space between the axon terminal and the next dendrite is called the synapse. Here, information transfers from one neuron to the other.

A single neuron can already perform several major computational operations. By changing the surface structure of its dendritic tree, a neuron can alter its sensitivity to incoming neurotransmitters. By manipulating the precise mechanisms that convert membrane potentials into sequences of action potentials, a whole additional range of computations can be



**FIGURE 1-1.** The basic structure of a neuron. Electrical impulses generated in the soma travel along a myelin-coated axon to the axon terminals, where they are translated in a chemical signal of neurotransmitters that are released into the synapse. At the receiving end of the synapse are the dendrites of another neuron that translate the chemical signal back into an electrical potential that may motivate its soma to generate action potentials. (Figure adapted from Vision Science, by Palmer)

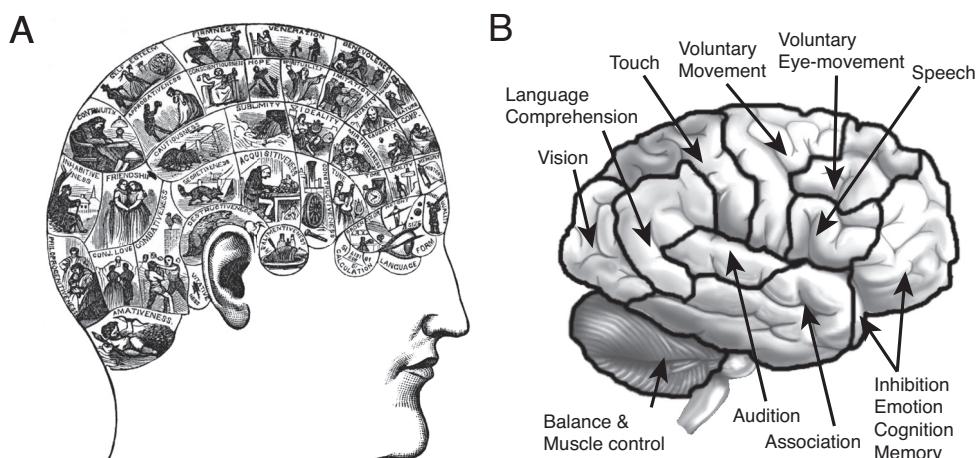


performed (Agüera y Arcas et al., 2003; Koch & Segev, 2000). Finally, by changing the gain at its outgoing synapses, a neuron can determine how information is distributed to the next set of receiving cells. With billions of these computational units and trillions of these synapses, the brain effortlessly performs all the complex computations that allows us to see, feel, move, talk, think, remember, and do all the other things that we usually take for granted.

### 1.3 From phrenology to functional specialization

A brain is more than a large collection of neurons that are held together by a skull and labeled with a face. Its functional organization contains an eloquent structure of highly specialized, interconnected areas, each with their own typical function and cytoarchitecture. The gist of functional specialization was first introduced by nineteenth century phrenologists. They believed that the brain contains a limited set of different functions, and wrongfully hypothesized that one might spatially localize these functions by feeling the bumps on the skull (Figure 1-2A).

While phrenology is nowadays universally considered a pseudoscience, the rapid rise of functional magnetic resonance imaging (fMRI) in neuroscience in the early 1990's sometimes resulted in what has been labeled 'modern phrenology'. An abundance of neuroscientific studies merely 'localizing' the 'neural correlate' of some brain function might have somewhat occluded the true potential of novel brain imaging techniques. In fMRI, a person is placed in a large magnetic scanner that can measure blood flow and blood oxygen levels in the brain. Since active neurons need more oxygen than inactive ones, these fMRI scanners can provide the researcher with an indirect measure of localized brain activity. The spatial resolution of the fMRI signal strongly depends on the strength of the magnetic field, but in a typical scan, a single data-unit will be a 'voxel' of approximately 2x2x2 mm, containing up to



**FIGURE 1-2.** A) A nineteenth century phrenology map of the brain. Phrenologist believed that the brain consisted of different functional areas, whose location could be derived from the bumps on the skull. B) A modern brain map indicating a more accurate interpretation of functional specialization.



a million neurons. Apart from the intrinsic value of cleverly designed fMRI paradigms, such basic functional localization can tell a neurophysiologist where to zoom in and start looking for the actual neuronal mechanisms underlying the functions in which these areas were shown to be involved.

A very coarse functional brain map is depicted in Figure 1-2B. As can be seen, a large portion of the primate cerebral cortex is involved in visual perception. Within this vision area, an even more specialized functional subdivision can be made, with specific areas involved in the perception of, e.g., orientation, color, motion, or depth.

## 1.4 Visual cortex: a brief introduction

The brain areas involved in vision are mostly superficially located in the outer layers of the brain that are together called the cerebral cortex. The functional subdivisions within visual cortex are usually referred to with a number that roughly corresponds to the temporal order in which visual information reaches these areas. V1, or the primary visual cortex, is the first cortical area that receives visual information. V2 gets its information from V1 and reroutes it to V3, etc. However, before visual information reaches the cortex altogether, it must first pass several other, subcortical relay stations (Figure 1-3A).

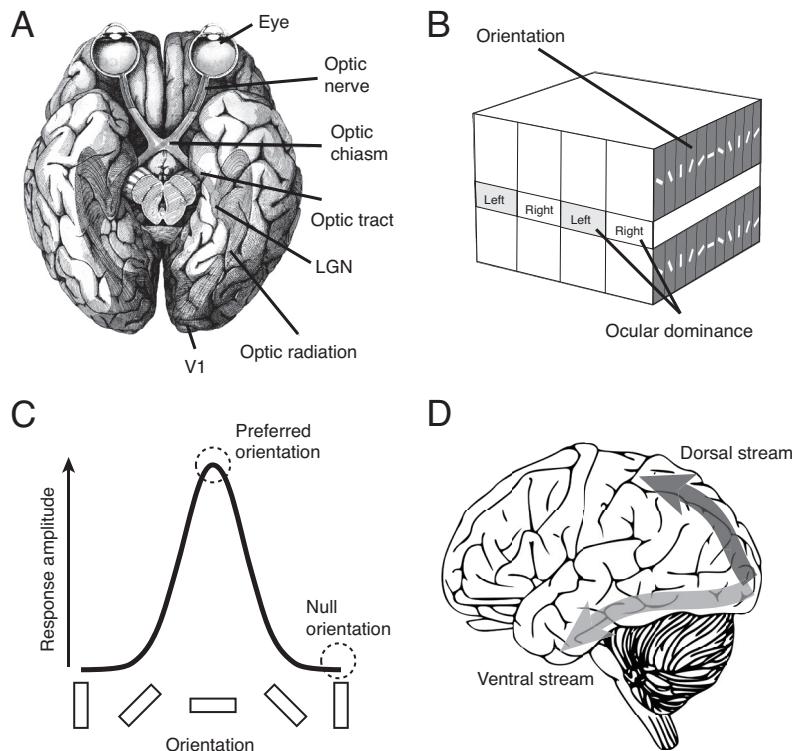
In short, visual information is detected by the photoreceptors that are located in the retinas of the two eyes. Retinal information then initially travels towards the primary visual cortex through the two optic nerves. Since the two eyes are inevitably positioned at two different locations in the head, the two retinal projections of the visual world are not exactly the same, but contain slightly shifted versions of each other. This shift is called binocular disparity and the brain later uses it to calculate stereoscopic depth, allowing three-dimensional vision (Qian, 1997). The two optic nerves meet at the optic chiasm, where the visual information is reordered. Information from the left side of the visual field (but from both eyes) is further processed in the right hemisphere of the brain, whereas the right side of the visual field is processed in the left hemisphere. Both the left and right optic tracts project from the optic chiasm to a subcortical relay structure on either side of the brain called the lateral geniculate nucleus (LGN). In the LGN, incoming information is combined with a large amount of feedback information that is projected down from the cortex to these subcortical structures. The output of the LGN then projects to the primary visual cortex through the optic radiation via a very specific organization of in- and output layers (Hubel & Wiesel, 1972).

The neurons in V1 are retinotopically organized, which means that the neurons that receive information from a specific part of the retinal projection (and thus of the visual field) are grouped together in clusters (Glickstein & Whitteridge, 1987; Wandell et al., 2007). The small area of visual space to which a neuron selectively responds is called the receptive field (Hubel & Wiesel, 1962; Hubel & Wiesel, 1963). Receptive fields are relatively small in primary visual cortex (roughly one degree of visual angle, which corresponds to an area of about two centimeters in diameter viewed from one meter distance) but become increasingly larger in later cortical areas. V1 neurons predominantly respond to stimulation via either one of the two eyes, and neurons with a similar eye-preference are grouped in so-called ocular dominance columns that are oriented perpendicular to the cortical surface (Figure 1-3B) (Hubel & Wiesel, 1968). Within these ocular dominance columns, neurons are again grouped in even



smaller subcolumns that are based on the neurons' orientation selectivity (Hubel et al., 1977; Hubel et al., 1978).

If a simple bar of light is presented in the receptive field of an orientation selective V1 neuron, the magnitude of its response will depend on the orientation of the light bar. Each V1 neuron has its own a preferred orientation to which it will respond with maximal activity. Consequently, each neuron also has its own 'null-orientation' that is usually orthogonal to the preferred orientation and evokes hardly any response, or no response at all (Figure 1-3C). All intermediate orientations will evoke intermediate levels of response magnitude depending on how similar they are to either the preferred or null orientation. Just like V1 cells are tuned to stimulus orientation, other cells in visual cortex can be tuned to other visual features. The motion direction and speed tuning of neurons in area V5 (also called MT) are another prominent example of such feature tuning (Born & Bradley, 2005) that will be exploited in some of the research described in this thesis.



**FIGURE 1-3.** A) The path of visual information before it reaches the cortex. Visual input is captured by the retinas in the eyes and travels through the optic nerve, via the optic chiasm and optic tract to the lateral geniculate nuclei (LGNs). Here it is reorganized and delivered to the cortex via the optic radiation. B) A cortical column demonstrating ocular dominance columns and orientation columns. C) An orientation tuning-curve. A neuron will respond with maximal response amplitude to its 'preferred orientation' and hardly to its 'null orientation'. Similar tuning curves can also be measured for other features to which neurons respond selectively. D) The ventral and dorsal stream of visual processing.



Cortical columns of neurons that are selective to similar visual features but responsive to different locations of visual space are connected via axonal projections called ‘horizontal connections’ (Gilbert & Wiesel, 1983; Gilbert & Wiesel, 1989; Rockland & Lund, 1983). They are called horizontal because they run parallel to the cortical surface and perpendicular to the vertically organized columns. The number of horizontal connections is high for columns that are located near to each other within the cortex and that have receptive fields relatively near to each other in visual space, but they become sparser when the cortical distance and the spatial separation of receptive fields become larger. Whereas ocular dominance largely disappears in cortical areas beyond V1, where cells are predominantly binocular, the columnar feature organization with horizontal connectivity also exists in many later visual cortical areas.

Concerning cortical visual processing beyond area V1, it is common to make a distinction between two functional and anatomical routes of visual processing (Figure 1-3D)(Goodale & Milner, 1992). The ‘ventral stream’ is the pathway that leaves V1 towards the inferior temporal lobe via V2 and V4. It is also called the ‘what-pathway’ or ‘vision-for-perception stream’ since it is predominantly involved in tasks of object recognition and form representation. The other pathway, the ‘dorsal stream’, runs from V1 up to the posterior parietal cortex via V2 and V5/MT. It is involved in spatial awareness and the perception of motion and it is therefore also called the ‘where-pathway’ or ‘vision-for-action stream’.

The concise description of visual cortical organization above is of course a simplification of the real complexity of the visual system. Major omissions include a description of feedback connections from higher cortical areas back to lower ones (e.g. to mediate attention) and a mentioning of the abundance of recurrent connectivity within visual cortex (Lamme et al., 1998). However, this summary hopefully conveys enough of the essential aspects of the neuronal basis of visual processing to promote a better understanding the research in this thesis. In particular, the feature tuning of individual cells, the columnar organization of similarly tuned cells, and the horizontal connectivity between clusters of similarly tuned neurons are aspects of neuronal organization that will be addressed in later chapters.

## 1.5 Why study vision?

Humans, as well as many other mammals, are visually oriented animals. Just imagine someone cycling through traffic in a crowded city. The seeming effortlessness with which the brain processes the large amounts of incoming visual information starkly contrasts with the complexity of the actual operations that are performed. To illustrate, one instantly recognizes the surrounding cars, even if they are partially occluded. One is able to detect potentially threatening elements from the multitude of less relevant cues and accurately avoid them by estimating the upcoming trajectories of multiple moving objects. Moreover, the brain is performing all these calculations while many elements in the visual scene are moving autonomically, the body that the brain is in is moving as well, the head is moving relative to the body, and the eyes relative to the head. Given the flexibility that is required of the visual system to achieve the remarkably stable experience usually associated with seeing, the relatively large proportion of cortex that is involved in vision may not be all that surprising anymore.

The impressive amount of existing knowledge about visual processing mechanisms provides a more practical reason for studying vision. Visual perception has fascinated researchers



in psychology, physiology and philosophy for centuries and this broadly shared fascination has resulted in a considerable amount of theory about visual processing, knowledge about the inner workings of the neurons involved in vision, and computational models that describe and predict the rules of visual perception. At the same time, many aspects of visual processing are still unknown, and a full roadmap of the neural mechanisms that create lively visual experiences of color, shape, depth and motion out of basic patterns of photons that hit the retina is still far away. We are, however, slowly filling in the missing pieces of the puzzle.

Studying visual perception also has the advantage of vision being a sensory system that can easily be manipulated, approached and recorded from with a broad range of experimental techniques. It is relatively simple to create visual stimuli with high spatial and temporal resolution, especially when modern computer hardware can be used. The eyes can be easily individually stimulated to investigate processes of binocular visual processing, while the large amount of cortical surface attributed to vision offers good opportunities for fMRI experiments. Furthermore, the fact that visual processing predominantly occurs in the cortex, near the outside of the brain, allows non-invasive recording of brain activity with electroencephalography (EEG), stimulation with transcranial magnetic stimulation (TMS), and easy access with invasive electrodes with which one can record the activity of single neurons or small neuronal clusters in animal models. A very informative technique to investigate the neural mechanisms underlying visual perception without directly recording brain activity is human psychophysics. With this technique, human observers are presented with cleverly designed visual stimuli and asked to perform a highly specific task. Combined with existing knowledge of the physiological foundations of visual processing, such experiments provide invaluable new insights into the functioning of the visual system as a whole.

A final reason to study the neural mechanisms of visual perception is that many general aspects of brain functioning are present in a rather specialized version in vision. Examples of such mechanisms are visual learning, visual memory, visual plasticity, etc. With all the existing knowledge and practical advantages listed above, it could be very beneficial to first investigate these general brain functions in the specific context of vision. In this thesis, the focus will be on a very general and important aspect of visual processing: the role of context in the mechanisms that convert the physical reality of the outside world into the perceptual reality of conscious experiences. Perception is a global process. In constructing conscious perceptual experiences, information is integrated over space, time, sensory modalities, etc. With ‘context’ we mean all the information that is available to the brain, but that is not directly present in a particular pattern or object. However, from this information, specific additional features of the pattern or object may be inferred. For instance, without looking into an oven we may infer from a particular smell, perhaps in combination with the presence of thick smoke in the proximity of the oven, that a forgotten pizza has changed its color to black. In general, ambiguous visual stimuli are eminently suitable tools to study the underlying neural mechanisms of such processes of contextual inference in conscious visual perception.

## 1.6 Ambiguous stimuli expose the route from photon to percept

In studying visual perception, a lot can be learned from the ‘mistakes’ the visual system appears to be making every now and then. Optical illusions reveal the general processing rules

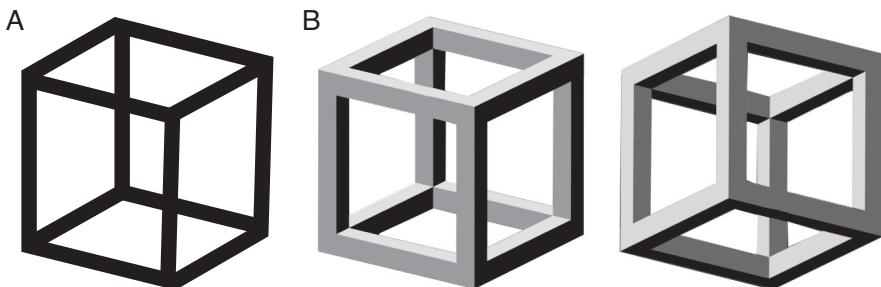


and shortcuts that are implemented by the brain to allow the fast and accurate performance at which it usually operates. In a sense, ambiguous visual stimuli are the ultimate optical illusions (Figure 1-4). Giving rise to multiple, equally valid, yet mutually exclusive perceptual interpretations, visual ambiguities demonstrate the visual systems capacity of creating crisp, stable percepts in the absence of definitive sensory evidence (Hoffman, 2000). Simultaneously, they create a kind of neural breadcrumb trace from stimulus to percept. At the start of this trace, there is the ambiguous stimulus for which different percepts can arise from the same physical input. At the end of the trace, there is the conscious percept that, at any particular moment, appears to be an unrivaled representation of the physical input. This apparent perceptual robustness, however, disappears upon continuous viewing of ambiguous stimuli, when the different possible stimulus interpretations are perceived to switch dominance every few seconds while the stimulus remains constant.

This perceptual switching behavior is the reason that ambiguous stimuli are also sometimes called ‘multistable stimuli’ (or ‘bistable stimuli’, if there are only two possible perceptual interpretations). The entire underlying process involved in these perceptual switches is known as ‘visual rivalry’ or ‘visual competition’ implying that the possible perceptual interpretations somehow compete against each other for access to visual awareness (Blake & Logothetis, 2002). Since the perceptual alterations of the conscious percept occur with a stimulus that remains constant, the neural route from stimulus to percept must crucially change its trajectory somewhere when a perceptual switch occurs. By carefully studying how a broad range of factors affects the endpoints and trajectories of the traces from stimulus to percept, the original mechanisms that underlie conscious visual perception can be incrementally unraveled.

Decades of research with ambiguous stimuli have resulted in a rough division between two types of visual ambiguities that are largely studied in parallel: binocular rivalry stimuli and perceptual rivalry stimuli. Binocular rivalry (Alais & Blake, 2005) relies on the fact that visual information first enters our brain via two isolated retinas, yet our perception contains only one single visual interpretation of the world. This means that somewhere along the route from stimulus to percept, the input from the two retinas must be combined into a single perceptual representation. Usually, the difference between the two retinal projections is only minor and the brain uses this binocular disparity to calculate the three-dimensional structure of the world. However, when the two eyes are presented with radically different images, the brain does not simply combine these two images into a fused single percept. Instead, perception often alternates between the two monocular images. Binocular rivalry has been studied for centuries resulting in a serious amount of knowledge about its temporal and spatial dynamics (Alais & Blake, 2005). The phenomenon initially gained lots of interest after the introduction of the mirror stereoscope (a mirror construction allowing straightforward presentation of dissimilar images to the individual eyes) by Sir Charles Wheatstone in the early nineteenth century, but its more recent revival in visual perception research is driven both by technical advances in neuroscience and by the recognition of binocular rivalry as a promising experimental window on the neural mechanisms of conscious visual perception and, ultimately, consciousness itself (Crick & Koch, 2003).

Perceptual rivalry occurs when the brain cannot unequivocally infer a physical trait from the limited set of visual features that are present in a stimulus (Leopold & Logothetis, 1999). The Necker cube is a classical example of a perceptual rivalry stimulus. Here, the third dimension of the perceived cube is inferred from two-dimensional perspective cues, causing a



**FIGURE 1-4.** A) A Necker cube. Lacking explicit depth information, the Necker cube can be perceived in two distinct orientations. B) If additional depth cues are added the cube stabilizes into a single perceived orientation.

set of flat, two-dimensional lines to be perceived as a three-dimensional cube (Figure 1-4A). Since there are no additional cues that specify the depth order of the inferred surfaces, the orientation of the cube is ambiguous and perception will alternate between the two possible interpretations (Figure 1-4B). Another perceptual rivalry stimulus that is often used in visual neuroscience relies on the brain's capacity to infer structure from visual motion (Andersen & Bradley, 1998; Kourtzi et al., 2008). A set of randomly positioned dots move around on a two-dimensional computer screen as if they were painted on the otherwise transparent surface of an orthogonally viewed rotating cylinder. This stimulus configuration yields a vivid three-dimensional percept of such a rotating cylinder, but in the absence of additional depth cues, the stimulus lacks the depth order information that would define its front and backside and the rotation direction becomes ambiguous.

The two hallmark features that characterize ambiguous visual stimuli are 1) *The mutual exclusivity* of the possible perceptual interpretations (however, see Chapter 7 of this thesis for an example of incomplete exclusivity), and 2) *The perceptual alternations* that occur between these interpretations upon prolonged viewing (Long & Toppino, 2004). Traditionally, the perceptual dynamics of visual rivalry are studied by focusing on these alternations of perception during prolonged exposure to ambiguous stimuli. Such an approach essentially studies the *instability* of the visual system under ambiguous input. In order to study the *stability* of perception instead, one can focus on the mutual exclusivity, which is most clearly probed when the brain initially chooses one percept over the other at the onset of an ambiguous stimulus (Noest et al., 2007).

By temporarily removing the stimulus from view before the instability sets in and causes a perceptual switch, one can repeatedly probe the percept-choice mechanism and obtain a sensitive measure of the underlying neural dynamics. This technique of intermittent presentation of ambiguous stimuli has become common practice over the last few years (Brascamp et al., 2009; Brascamp et al., 2010; Brascamp et al., 2008; Klink et al., 2008a; Kornmeier et al., 2007; Leopold et al., 2002; Maier et al., 2003; Noest et al., 2007; Pastukhov & Braun, 2008; Pearson & Brascamp, 2008) and has already proven its applicability in revealing mechanisms of visual processing that cannot be demonstrated with continuous stimulus presentation. Some of the chapters in this thesis provide examples of this percept-choice approach.



The debate about whether binocular rivalry and perceptual rivalry are driven by common neural machinery (Blake & Logothetis, 2002; Leopold & Logothetis, 1999) is not likely to be resolved in the near future (but see Chapter 2 of this thesis). However, from an experimental point of view, the two types of visual ambiguities are at least equivalent in their dissociation of stimulus and percept. This dissociation makes both types of rivalry highly suitable as tools to investigate how the brain creates stable and coherent conscious visual percepts.

## 1.7 Context resolves visual ambiguities

The brain states associated with the perception of ambiguous stimuli can be compared to a student that is forced to answer the exam question ‘What color does an apple have?’, by choosing between the two correct alternative answers ‘red’ and ‘green’. We can easily see how the impossible choice between the two equally valid answers might be guided by external influences. In an attempt to reach an acceptable answer, the student may look for spatial help by peeking at the answer of the student that is sitting next to him or her. Alternatively, the temporal context of the apple he or she had at lunchtime may enforce an answer driven by his or her most recent experience with the concept ‘apple’. There are many more external influences conceivable that may facilitate an eventual answer to a question that cannot be misanswered, but the message is clear: if the information provided within a problem is insufficient to reach a definitive solution, one looks for additional cues. In this respect, the visual system is no different than our hypothetical student.

When sensory systems are confronted with ambiguous input, context is a natural disambiguator. While this is true on a macro-scale for the perceptual interpretation of ambiguous patterns, it is similarly accurate for all visual processing on a lower micro-level. Since each neuron in (primary) visual cortex has a limited spatiotemporal receptive field, each neuron will consequently only ‘see’ a limited fraction of the full picture. Within these small neuronal spotlights, ambiguities from a lack of information arise naturally, and coherent perception can only be obtained when information is integrated over space and time. Similarly so, at the macro-level of ambiguous visual pattern perception, the influences of context (either in space or time) can be profound and clever experimental manipulations of the sensory context that accompanies an ambiguity can provide valuable insights in the generic neural mechanisms of perception.

This thesis will demonstrate how a range of different contexts can influence visual perception. By combining experimental data, both with existing knowledge of the brain’s functional anatomy, and with computational modeling approaches, the potential of contextual influences to teach us about the mechanisms of conscious visual perception will be discussed. Furthermore, some hypotheses will be postulated about the precise neuronal organization that might underlie our findings. In Chapter 2 we will first demonstrate how binocular rivalry and perceptual rivalry are functionally very similar. Chapter 3 will investigate the influences of temporal context and attention on the perceptual dynamics of both binocular and perceptual rivalry. Chapter 4 will expand the findings concerning temporal context to the level of the single neuron and local cortical network through neurophysiological recordings in the awake behaving non-human primate. Chapter 5 will demonstrate an effect of spatial context that may directly be related to the visual cortical cytoarchitecture. In Chapter 6, we do not



use ambiguous stimuli, but instead we provide an example of crossmodal influences on the visual perception of time. Returning to binocular rivalry in Chapter 7, we show how human psychophysics combined with basic knowledge of neuronal processing can result in valuable new insights about adaptive brain functioning. We reveal a form of plasticity in binocular vision that is highly consistent with previously proposed synaptic learning mechanisms. Finally, the last chapter of this thesis will summarize the work and suggest some future research directions that seem most promising in their potential to answer some of the novel questions and hypotheses raised by the findings presented in this thesis.



# Chapter 2

**General validity of Levelt's propositions reveals  
common computational mechanisms for visual rivalry**

*All generalizations are dangerous, even this one.*

Alexandre Dumas

Published as

Klink, P.C., van Ee, R., & van Wezel, R.J.A. (2008). General validity of Levelt's propositions reveals common computational mechanisms for visual rivalry. *PLoS ONE*, 3(10): e3473



*The chapters of this thesis are not in chronological order. The experiments that are described in this chapter were in fact performed after those in Chapter 3. We had recently concluded that the influences of voluntary control on the perceptual choices of intermittently presented ambiguous stimuli closely resembled a contrast gain mechanism. This notion had inspired a short student's project to find out whether we could evoke effects similar to those of attention by selectively manipulating the dot luminance of one of the two motion directions in an ambiguously rotating structure-from-motion sphere. It turned out that we could, and I was supposed to present these results at one of the Friday afternoon labmeetings. As was often the case, I did not get around to actually putting together this presentation until late on the Thursday evening before the lab-meeting. In these experiments, we had used a paradigm of intermittent stimulus presentation to allow close comparison with the results that were obtained with voluntary control earlier, but while I was composing the slides for the presentation, I started wondering how dot luminance manipulations would influence the perceptual statistics of continuously viewed rival stimuli. Eventually, my curiosity got the best of me and, alone at the lab at this nightly hour, I programmed the experiment to run myself as a subject. The resulting graph reminded me of a figure in a recent paper by Jan Brascamp (Brascamp et al., 2006). In this paper, a reversal of Levelt's classic second proposition is demonstrated to occur with binocular rivalry stimuli at low stimulus contrasts. The similarity with our novel data inspired us to attempt to rephrase the full set of Levelt's propositions in a way that would make them applicable to perceptual rivalry as well. After we had achieved this, we moved on to test the validity of these rephrased propositions outside the binocular rivalry paradigm and ended up with the experiments that are described in this chapter.*

---



## 2.1 Abstract

The mechanisms underlying conscious visual perception are often studied with either binocular rivalry or perceptual rivalry stimuli. Despite existing research into both types of rivalry, it remains unclear to what extent their underlying mechanisms involve common computational rules. Computational models of binocular rivalry mechanisms are generally tested against Levelt's four propositions, describing the psychophysical relation between stimulus strength and alternation dynamics in binocular rivalry. Here we use a bistable rotating structure-from-motion sphere, a generally studied form of perceptual rivalry, to demonstrate that Levelt's propositions also apply to the alternation dynamics of perceptual rivalry. Importantly, these findings suggest that bistability in structure-from-motion results from active cross-inhibition between neural populations with computational principles similar to those present in binocular rivalry. Thus, although the neural input to the computational mechanism of rivalry may stem from different cortical neurons and different cognitive levels the computational principles just prior to the production of visual awareness appear to be common to the two types of rivalry.

## 2.2 Introduction

In a world that provides an abundance of visual information our brain seemingly effortlessly decides which information reaches awareness. In the lab, this process can be studied using stimuli that cause perception to alternate between competing interpretations while staying constant on the retina (Blake & Logothetis, 2002). Two categories of such stimuli can be distinguished. In binocular rivalry the two eyes are independently presented with different visual stimuli (e.g. dissimilarly oriented gratings), causing either eye's image to be perceived in turn. In perceptual rivalry visual information is the same for both eyes but rivalry arises due to the existence of multiple mutually exclusive perceptual interpretations of the stimulus. An example is the well-known Necker cube, which causes perception to alternate between two spatial organizations of a flat line drawing. Binocular and perceptual rivalry are both manifestations of how the visual system handles inconclusive sensory evidence, but it remains unclear whether they include common computational mechanisms.

A comparison of the two types of rivalry tells us that their phenomenological appearance (Leopold & Logothetis, 1999) and temporal dynamics (Brascamp et al., 2005; van Ee, 2005) are similar during continuous viewing. Both types of rivalry can, in a qualitatively similar -yet quantitatively different- manner, be influenced by attentional efforts to hold one of the two alternative percepts dominant (Meng & Tong, 2004; van Ee et al., 2005). Furthermore, when presented with intermittent blank periods, they exhibit qualitatively identical effects of stimulus timing (Klink et al., 2008a; Leopold et al., 2002; Noest et al., 2007), and voluntary control (Klink et al., 2008a). Together, this suggests that even though binocular and perceptual rivalry may arise at different cortical levels (causing quantitative differences), the computational rules to produce perceptual output may be common (causing qualitative similarities).

Several computational models are available that provide an explicit theory of the computational mechanisms that underlie binocular rivalry (Kalarickal & Marshall, 2000; Lehky, 1988; Mueller, 1990; Noest et al., 2007; Wilson, 2007). An important set of constraints for



binocular rivalry models are based on the observations by Levelt (Levelt, 1966) regarding the relation between the strength (contrast) of the eyes' images and the time course of perceptual alternations. Levelt described in four propositions how perceptual dominance durations are affected by changes in the contrasts in either or both of the images engaged in rivalry. For instance, if the contrast of one image is increased this provides a competitive advantage to the associated neural representation, leading to a greater predominance of the corresponding percept (Levelt's rule I).

Levelt's complete set of binocular rivalry propositions state that 1) *Increasing the stimulus strength in one eye will increase the predominance of the stimulus*; 2) *Increasing the stimulus strength in one eye will not affect the average duration of dominance in that eye*; 3) *Increasing the stimulus strength in one eye will increase the rivalry alternation rate*; 4) *Increasing the stimulus strength in both eyes will increase the rivalry alternation rate* (Levelt, 1966). More recent observations, dictate a critical re-evaluation of the second proposition (Bossink et al., 1993; Brascamp et al., 2006; Mueller & Blake, 1989). Levelt's second proposition appears to be valid for high-contrast binocular rivalry stimuli, but to reverse for low-contrast stimuli (Brascamp et al., 2006). This means that based on this existing literature the second rule can no longer be regarded as valid and should be rephrased as 'manipulations of stimulus strength mainly influence the dominance durations of the percept from the eye with the strongest stimulus'. We will refer to this new rule as 'the revised second proposition'.

It is currently unclear to what extent theories that have been developed for binocular rivalry can be applied to other forms of rivalry. In this study we investigate whether Levelt's psychophysical observations that lie at the basis of virtually all binocular rivalry models can be generalized to perceptual rivalry. We use a structure-from-motion stimulus (for a review see Andersen & Bradley, 1998) for which a two-dimensional projection of a transparent sphere revolving around a vertical axis gives rise to perceptual rivalry between two depth organizations. In the absence of explicit depth information the sphere is perceived to alternately rotate in either of two directions: with the leftward moving surface in front and the rightward moving surface in the back, or vice versa. We investigate how the time course of perceptual alternations between these rotation directions is affected by changes in the luminance of the dots that define either of the two surfaces. Analogous to binocular rivalry, where an increase in the contrast of one of the conflicting images alters the neural competition process in favor of the corresponding neural representation, a luminance increment of the dots that comprise one of the surfaces in structure-from-motion rivalry alters the competition process such that the brighter surface is perceived in front a larger fraction of the time (Freeman & Driver, 2006; Schwartz & Sperling, 1983). However, it is an open question whether these manipulations – image contrast in binocular rivalry and dot luminance in structure-from-motion rivalry – affect the competition process in similar ways, or whether they are different.

Our results demonstrate that all four propositions regarding contrast and perceptual dynamics in binocular rivalry can without any serious modification be applied to dot luminance in bistable structure-from-motion. An important implication of this finding is that models of binocular rivalry that were inspired by Levelt's propositions can be applied to structure-from-motion rivalry as well. Moreover, given the highly dissimilar nature of the ambiguity in these two forms of rivalry, our results suggest that the neural computations that produce dominance in visual rivalry share common features for a broad range of rivalry stimuli.



## 2.3 Materials and methods

### 2.3.1 Observers

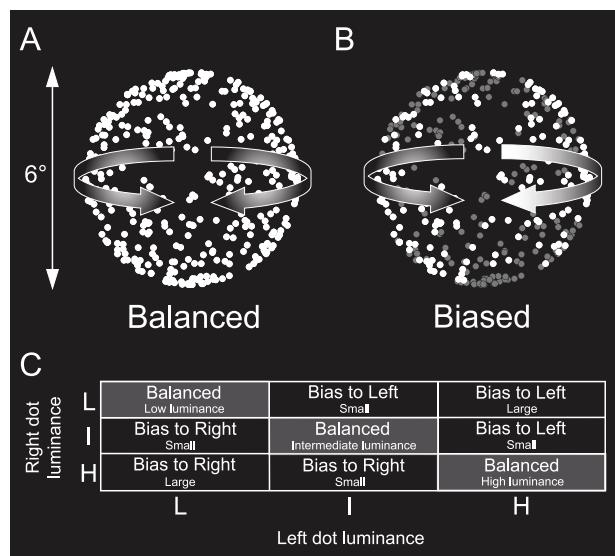
Five observers with normal or corrected to normal vision, ranging in age between 21 and 28 years, participated in our experiment. One of the observers was an author (CK), but the other four (students) were completely naïve with respect to the aims of the study.

### 2.3.2 Apparatus

Visual stimuli were generated on a Macintosh computer in MATLAB (MathWorks, Natick, MA) using the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997) and presented on a 22 inch CRT monitor with a resolution of 1280x1024 pixels and a refresh rate of 100 Hz. Observers used a head- and chinrest and viewed the stimuli from a distance of 120 cm.

### 2.3.3 Stimulus and procedure

The stimulus was a bistable rotating sphere, composed of two transparent layers of 450 random white dots with a sinusoidal speed profile on a black background ( $0.0 \text{ cd/m}^2$ ). The sphere size was 6 degrees, the dot size 0.05 degrees and the rotation speed was 80 degrees per second.



**FIGURE 2-1.** Schematic representation of the stimulus. White dots on a black background moving with a sinusoidal speed profile create the impression of a three-dimensional sphere rotating in depth around the vertical axis. If all dots have the same luminance (A, Balanced) both rotation directions are equally likely and the sphere is perceived to switch rotation direction every few seconds. If the dots moving in opposite directions have a different luminance (B) the sphere is biased towards the perceptual interpretation with the brightest dots in the foreground. C) In our experiment we used three different luminance levels (Low = L, Intermediate = I, High = H) for the two surfaces resulting in nine different sphere stimuli.



The number of dots and rotation velocity were chosen to maximize the number of reversals (Brouwer & van Ee, 2006). The luminance of the dots of both surfaces was manipulated between low, intermediate and high white intensities (corresponding to 25.3, 41.3 and 61.2 cd/m<sup>2</sup> respectively). This resulted in nine (3x3 dot luminance values) sphere stimuli configurations (Figure 2-1). The dots at the two different luminance values were drawn on the screen in random order to avoid a true depth ordering (due to overlapping dots) of the two layers. Stimuli were pseudo randomly chosen from the nine possible configurations and presented for 300 seconds while observers reported the perceived rotation direction of the sphere by pressing one of two buttons on a keyboard. Observers were explicitly instructed to report the direction of the perceived front surface to minimize the role of mixed percepts (Hol et al., 2003). Transition periods between two percepts were not recorded but subjects indicated that they were very short if present at all.

### 2.3.4 Data analysis

From the reported perceptual episodes we calculated the average dominance duration, reversal rates and predominance (percentage time spent in one percept) for all experimental conditions. As an extra test, percept durations were fit to a cumulative gamma-rate distribution function (Brascamp et al., 2005) using a bootstrap routine (1,000 repeats) to more reliably estimate the mean dominance durations. A Kolmogorov-Smirnov analysis demonstrated that more than 92 percent of our fits were significant at  $p = 0.05$ . Because all statistical analyses on the data yielded similar results for the directly calculated and fitted mean dominance durations we only report the results for the directly calculated percept durations. Group data were normalized to an observer's mean percept duration during the intermediate balanced luminance condition or mean reversal rate over all conditions. Statistical differences between conditions were tested with one-way analyses of variance (ANOVA).

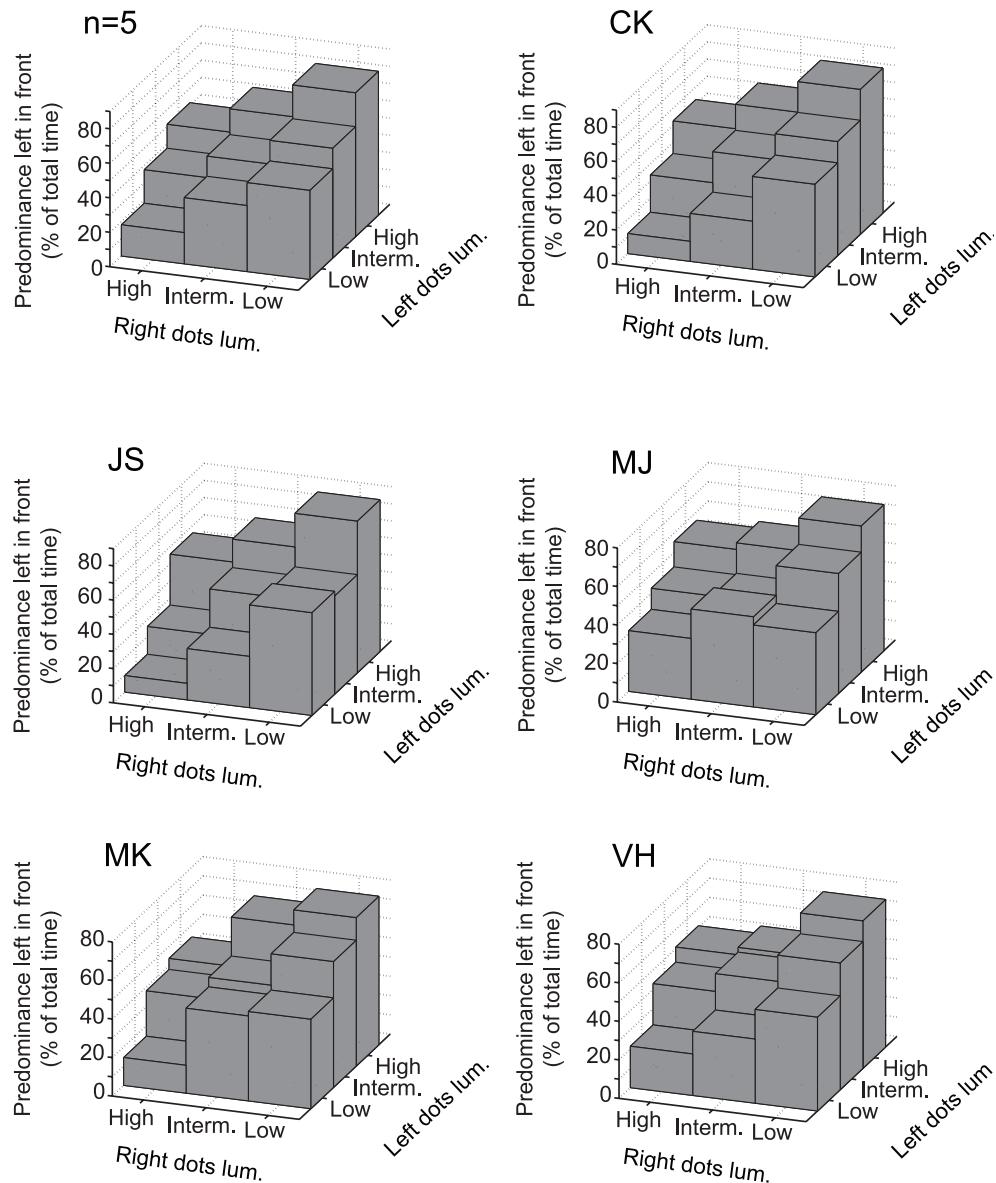
## 2.4 Results

To systematically evaluate the validity of the four propositions for perceptual rivalry we need to make a small -merely semantic- change to Levelt's original propositions. Stimulus strength and perceptual interpretation are tightly coupled in perceptual rivalry, but unlike in binocular rivalry they are not exactly similar. We updated the propositions accordingly and the results of our experiments will be presented following the original order of the propositions. Importantly, in the balanced stimulus conditions none of our observers demonstrated a significant bias for either of the two rotation directions of the bistable sphere (ANOVA,  $p > 0.27$ ).

### 2.4.1 Proposition 1

*Increasing the stimulus strength of one perceptual interpretation of a bistable stimulus increases the predominance of this perceptual interpretation*

Observers are more likely to perceive the surface with the brighter dots in the foreground and this effect is more prominent for larger dot luminance differences. Statistical analysis of



**FIGURE 2-2.** Predominance as a function of individual surface dot luminance. For both the group data ( $n = 5$ ) and individual observers a balanced increase of stimulus luminance (the diagonal) does not affect the predominance. If the luminance of only one layer of dots is manipulated the predominance shifts towards the perceptual interpretation with the brightest dots in the foreground.



the predominance data revealed significant increases of predominance (percentage of the total time that a percept is dominant) with increasing dot luminance for both left and rightwards moving dots (Left:  $F(2,12) = 36.56$ ,  $p < 0.001$ ; Right:  $F(2,12) = 20.28$ ,  $p < 0.001$ ). Likewise, decreasing dot luminance led to significant decreases in predominance of the corresponding percept (Left:  $F(2,12) = 31.73$ ,  $p < 0.0001$ ; Right:  $F(2,12) = 28.35$ ,  $p < 0.0001$ ). Balanced dot luminance manipulations did not have any significant effect on predominance ( $F(2,12) = 0.19$ ,  $p = 0.83$ ). Figure 2-2 shows the predominance of leftward rotating spheres (left in front) for all combinations of dot luminance. A similar pattern is present for all observers and the average group data (top left panel). These findings confirm that the perception of a bistable sphere is consistent with the first proposition.

#### 2.4.2 (Revised) proposition 2

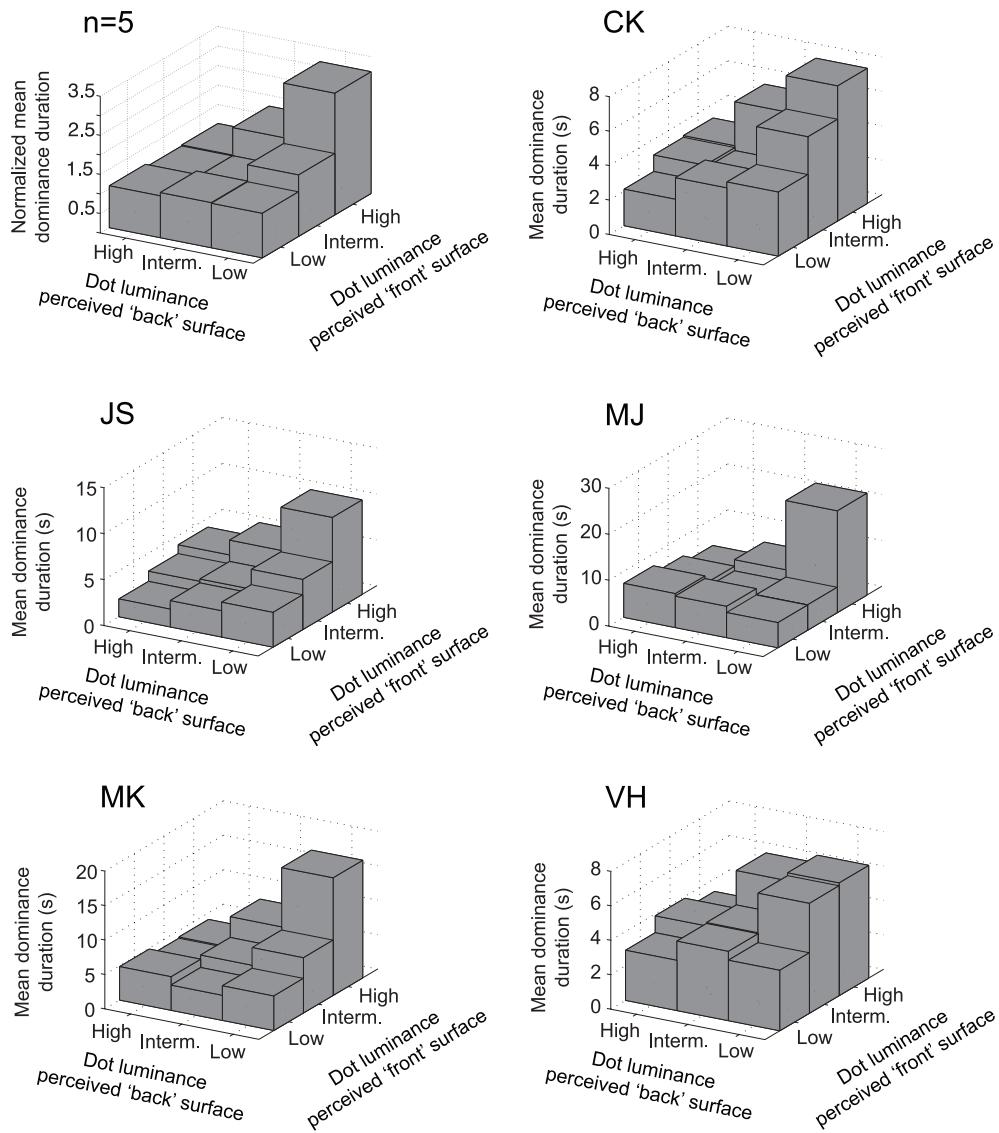
*Manipulations of stimulus strength of one perceptual interpretation of a bistable stimulus will mainly influence the average dominance duration of the perceptual interpretation corresponding to the strongest stimulus*

Increased predominance of a percept can be the result of longer dominance durations of the percept, shorter dominance durations of the opposite percept or both. Recently it has been shown that changes in predominance in binocular rivalry mainly results from changes in the average dominance duration of the strongest stimulus (Brascamp et al., 2006). Our bistable rotating sphere demonstrates similar results for perceptual rivalry confirming our revised second proposition for perceptual rivalry.

Figure 2-3 demonstrates that starting with a high luminance stimulus, a decrease in dot luminance of one of the two dot surfaces only affects the mean percept duration of the 3-D percept with the alternative surface (consisting of the brighter dots) in front. For example the percept of a sphere with high luminance dots in both the front and back has approximately the same mean dominance duration as the percept of a sphere with low luminance dots in the front and high luminance dots in the back, but significantly shorter average percept durations than the percept with high luminance dots in the front and low luminance dots in the back.

In other words, decreasing the stimulus strength of one perceptual interpretation does not influence the average dominance duration of this percept but it does influence average dominance durations of the opposing percept. A statistical analysis of the effect confirms that decreasing the stimulus strength of a perceptual interpretation has no significant effect on the dominance duration of the same percept ( $F(2,33) = 0.16$ ,  $p = 0.8525$ ) but does have a significant effect on the dominance duration of the opposite percept ( $F(2,33) = 17.41$ ,  $p < 0.0001$ ). As in binocular rivalry, the opposite holds true for low luminance stimuli. Here an increase in stimulus strength does increase the mean dominance duration of the same percept ( $F(2,33) = 16.92$ ,  $p < 0.0001$ ) while leaving the dominance durations of the opposite percepts unaffected ( $F(2,33) = 0.62$ ,  $p = 0.54$ ). This pattern of effects is present for all individual observers as well as the group data (Figure 2-3).

Figure 2-4 plots the effects of a manipulated dot luminance on mean dominance duration in a different way for the group data and individual observers. Starting with a fixed dot luminance (arrow), the dot luminance of one surface is varied (solid line) while that of the other is fixed (dashed line). Figure 2-4 clearly demonstrates that changes in mean dominance



**FIGURE 2-3.** Percept dominance durations as a function of differential dot luminance. Dominance durations are plotted for percepts with defined luminance for the perceived 'front' and 'back' surface of the sphere. For both the group data ( $n = 5$ , data for each observer is normalized to the mean percept duration at intermediate contrast) and the individual observers, dominance durations are longest when observers perceive the brightest possible dots in the foreground and de dimmest possible dots in the background. A decrease of the dot luminance of one of the layers of a balanced high luminance stimulus does not decrease the durations of the episodes when this layer is perceived in the foreground. Instead, it increases the durations of the opposite perceptual interpretation. If however the dot luminance of a balanced low luminance stimulus is increased, this only influences the duration of the episodes when the varied dot luminance is perceived as the foreground (see also Figure 2-4).



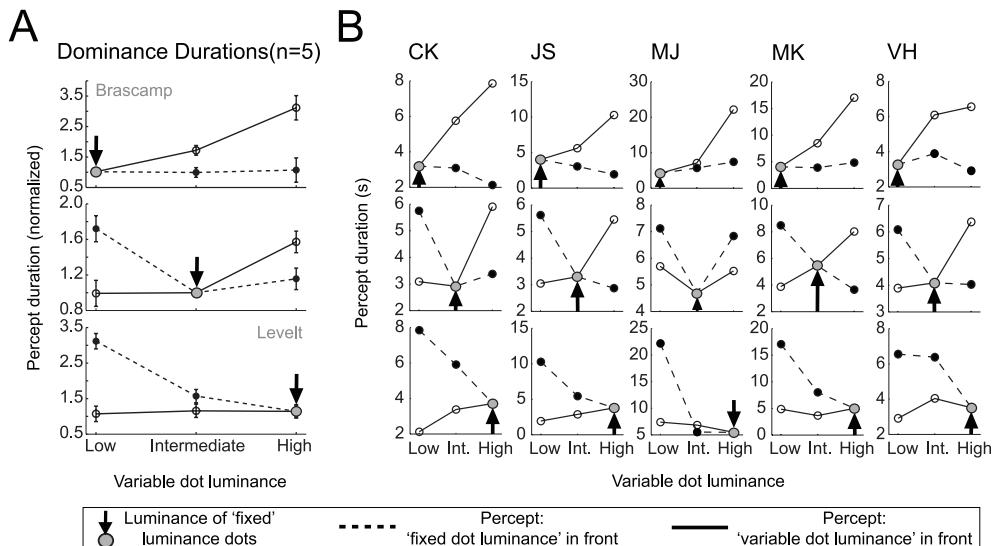
duration predominantly occur for the percept with the brightest dots in front. This effect is independent of which dot luminance is manipulated and consistent with our revised second proposition.

### 2.4.3 Proposition 3

*Manipulating the stimulus strength of one perceptual interpretation of a bistable stimulus will influence the average rivalry reversal rate*

Levelt's third proposition directly followed from his first and second propositions. It states that increasing stimulus strength increases the predominance of the corresponding stimulus by reducing the mean dominance duration of the other stimulus rather than increasing its own mean dominance duration. This automatically results in higher reversal rates when the strength of one of the two stimuli is increased.

Following the same line of reasoning our revised second proposition predicts that decreasing the dot luminance of one of the surfaces in a high luminance sphere would result in an increase of the dominance durations of the percept with the other (brighter) surface in front, leading to a lower reversal rate. Figure 2-5 demonstrates that this is indeed the case (group data is normalized to the overall mean reversal rate for an observer). The results for individual observers in Figure 2-5 are a little noisier but the same pattern is clearly pres-



**FIGURE 2-4.** Starting with a balanced stimulus of low, intermediate or high luminance (indicated with a grey dot and an arrow) the dot luminance of one of the two layers is manipulated while that of the other remains fixed. Mean percept durations are plotted for episodes when the sphere is perceived with the fixed dot luminance surface in the foreground (dotted lines) or with the variable dot luminance in the foreground (solid lines). For both the average group data (A) and the individual observers (B) manipulations of dot luminance mainly affect the mean dominance durations of the percept with the brightest dots in the foreground. Error bars represent the standard error of the mean.



ent. Statistical analysis of this effect revealed that both decreasing the dot luminance of the rightward moving dots ( $F(2,12) = 18.42, p < 0.0001$ ) and the leftward moving dots ( $F(2,12) = 24.49, p < 0.0001$ ) significantly decreased the reversal rate.

Our revised second proposition also predicts that an increase of the dot luminance of one of the surfaces in a low luminance sphere causes longer mean dominance durations of the percept with these dots in the front while hardly influencing the dominance durations of the alternative percept. Naturally this would also result in a decrease of the reversal rate. Figure 2-5 demonstrates that this is indeed the case and a statistical analysis demonstrated that for both dot surfaces this effect was significant (Left:  $F(2,12) = 11.56, p < 0.001$ ; Right:  $F(2,12) = 21.34, p < 0.002$ ).

Note that reversal rates can also increase as a result of increasing the stimulus strength. If we start off with a stimulus consisting of one high- and one low-luminance dot-surface and we increase the dot luminance of the low-luminance surface, our second proposition predicts that the average dominance duration of the manipulated percept remains unchanged whereas that of the fixed percept decreases resulting in an increase of reversal rates. Figure 2-5 demonstrates that this is indeed what happens.

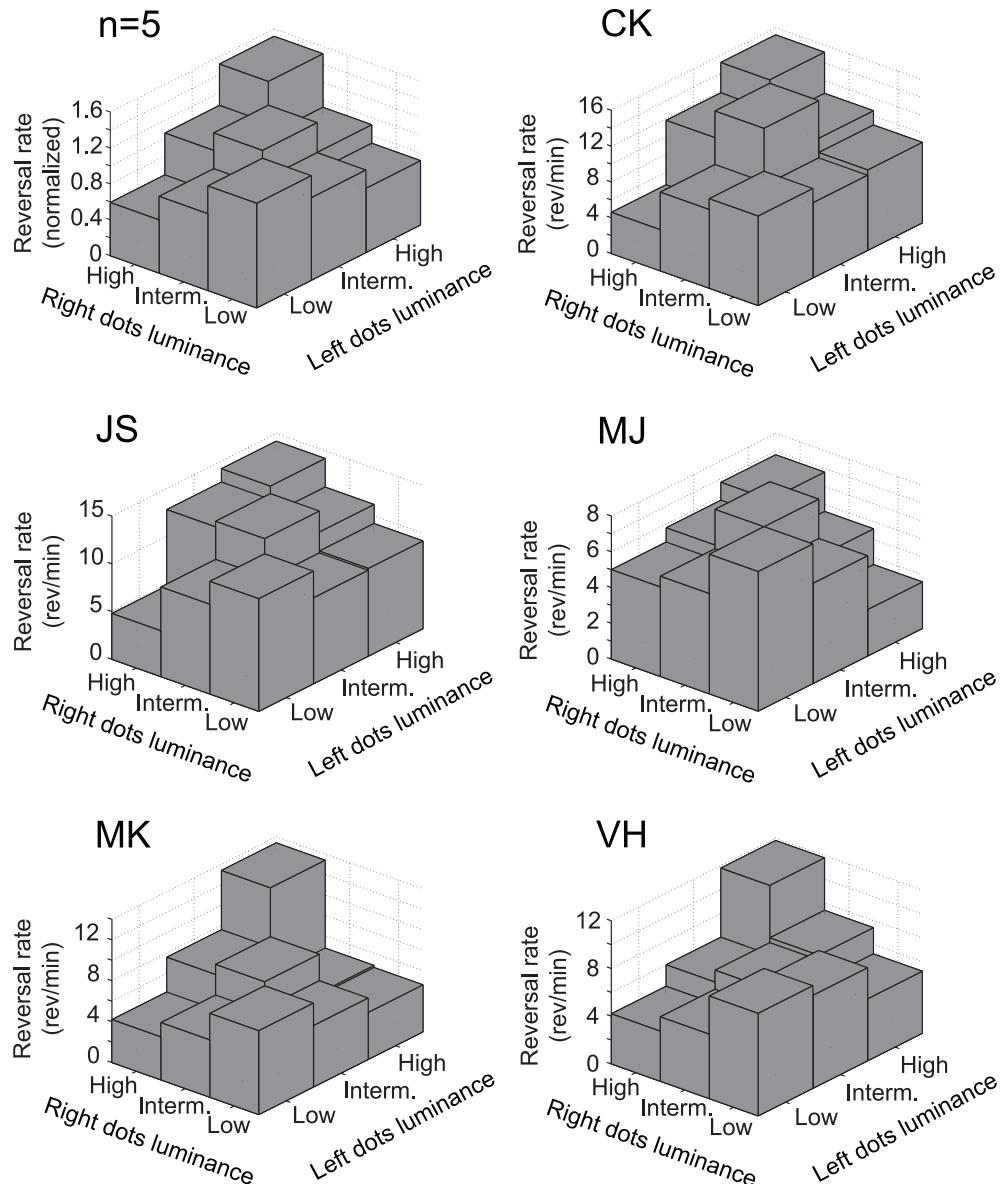
#### 2.4.4 Proposition 4

*Increasing the general stimulus strength of a bistable stimulus will increase the average rivalry reversal rate*

Until this point we have focused on changing the dot luminance in one of the two layers to manipulate the strength of the perceptual interpretation with these dots as the front surface. A test of proposition 4 however requires increasing the general strength of the stimulus. Changing the dot luminance of both surfaces with the same amount might be used to accomplish this manipulation. Figure 2-5 demonstrates the effect of stimulus strength on reversal rate with the bars on the diagonal of the plot. Increasing the stimulus strength does indeed increase the rivalry reversal rate (group data:  $F(2,12) = 4.24, p < 0.05$ ). Whereas balanced manipulations of dot luminance are unlikely to change the strength of ‘sphereness’ of the stimulus, it will influence the neural dynamics (e.g. adaptation speed) at earlier neural levels where the individual dots or surfaces are processed. The differences between subjects present in Figure 2-5 probably reflect differences in their individual neural dynamics at these non-rivalry stages.

## 2.5 Discussion

Binocular rivalry and perceptual rivalry provide unique windows on visual consciousness. Since perception alternates vividly in the absence of stimulus changes, the alternations can only result from the internal mechanism that shapes subjective experiences (Koch, 2004). However, it remains unclear how similar these internal mechanisms are for binocular rivalry and perceptual rivalry. We have shown that crucial constraints for binocular rivalry models inspired by Levelt’s four propositions can just as well be applied to the perceptual rivalry of a bistable rotating structure-from-motion sphere. Predominance shifts towards the strongest



**FIGURE 2-5.** Reversal rates as a function of dot luminance of the two motion-defined surfaces. For the average group data ( $n = 5$ ) the reversal rates of individual observers were normalized to the mean reversal rate over all conditions. Logically following from the demonstrated effect of dot luminance induced percept probability on the mean dominance durations (Figures 2-3 & 2-4) these plots demonstrate that for both the group data and the individual observers' reversal rates decrease when the luminance difference between the two surfaces increases. Furthermore, reversal rates for balanced stimuli increase when the dot luminance, and thus stimulus strength, increases.

---



perceptual interpretation (*I*), only the mean dominance of the strongest perceptual interpretation is influenced by dot luminance-based changes in percept probability (*II*), reversal rates change consistent with dominance duration (*III*) and the reversal rates increase if dot luminance of all the dots is increased (*IV*). The validity of the revised second proposition in particular implies that visual competition in perceptual rivalry involves an active process of cross-inhibiting neural populations with computational principles much like we find in binocular rivalry. Thus, although the neural input to the computational mechanism of rivalry may stem from different cortical neurons and different cognitive levels the computational principles just prior to the production of visual awareness appear to be common to the two types of rivalry.

There is considerable evidence supporting the idea that perceptual outputs in binocular and perceptual rivalry are at least partially based on a common computational mechanism. Percept durations under continuous viewing conditions are distributed similarly (Brascamp et al., 2005) and their drift and serial correlation are also comparable (van Ee, 2005). Quantitatively, observers attempting to hold one percept as long as possible through voluntary control affect percept durations in perceptual rivalry more than in binocular rivalry, but the qualitative dynamic aspects are similar (Meng & Tong, 2004; van Ee et al., 2005), even in terms of the individual fit parameters of percept duration distributions (van Ee et al., 2006). It has also been reported that observers with slow perceptual switches in one bistability paradigm are also slow switchers in another paradigm (Carter & Pettigrew, 2003; Pettigrew, 2001). If stimuli are presented with intermittent blank periods, binocular and perceptual rivalry exhibit similar qualitative effects of stimulus timing on the percept sequences (Klink et al., 2008a; Leopold et al., 2002; Noest et al., 2007) and they are comparably influenced by voluntary control (Klink et al., 2008a). Eye movements affect the two types of rivalry in a qualitatively different way. They play a greater causal role in producing perceptual alternations in binocular rivalry than in perceptual rivalry (van Dam & van Ee, 2006b). This qualitatively different effect of eye movements and the quantitative differences with voluntary control and stimulus timing are consistent with the idea that binocular rivalry is a more low-level type of rivalry than perceptual rivalry (Klink et al., 2008a; Meng & Tong, 2004; van Dam & van Ee, 2006b; van Ee et al., 2005).

Our current findings, together with the studies mentioned in this paragraph, suggest that although binocular and perceptual rivalry may arise at different cortical levels, which causes quantitative differences (Blake & Logothetis, 2002), the computational rules that eventually produce perceptual output may be common (causing qualitative similarities). Note that we talk about common computational principles, not common neural machinery. Indeed, multiple bistable attributes of single binocular or perceptual rivalry stimuli undergo independent switching dynamics, suggesting that attribute-specific rivalry occurs in parallel at different levels of visual processing (Grossmann & Dobbins, 2006). Our view is in line with a recently developed physiologically and mechanistically plausible model for visual rivalry (Noest et al., 2007), which is developed in terms of minimal neural activity. In this minimal model, even a single neural stage -distinguishing this model from other existing models- of rapidly inhibiting but slowly adapting percept representations can qualitatively explain all experimental findings in perceptual and binocular rivalry to date. Quantitative differences between types of rivalry can be explained in this model with different gain factors resulting from various pre-rivalry processing stages (Klink et al., 2008a).



Any computational model of visual rivalry needs to be tested against experimentally established characteristics. Levelt's four propositions (Levelt, 1966) are probably the best-known critical tests for models of binocular rivalry (recent examples are (Laing & Chow, 2002; Lankheet, 2006; Moreno-Bote et al., 2007; Shpiro et al., 2007; Stollenwerk & Bode, 2003; Wilson, 2007)). It would be very useful to know if theories that have been developed to understand binocular rivalry could also be applied to other forms of rivalry. Levelt's four propositions make an excellent starting point to resolve this issue with respect to the alternation process in visual rivalry. For plaid rivalry, a manipulation of the stimulus strength of only one perceptual interpretation has already been claimed to result in behavior consistent with Levelt's second proposition (Hupe & Rubin, 2003; Rubin & Hupe, 2005), but a detailed and systematic analysis was never reported. Our study offers the first complete and systematic test of perceptual rivalry against Levelt's four propositions revealing that all rules regarding contrast in binocular rivalry (with inclusion of the revised second proposition) can -without any fundamental modification- be applied to dot luminance in bistable structure-from-motion.

The independent manipulation of stimulus strengths mentioned in Levelt's original propositions has long hindered a systematic application of the propositions to perceptual rivalry where we have only one stimulus. It is however questionable whether this independent manipulation of stimulus strength is essential. The ongoing debate about what is rivaling during binocular rivalry primarily focuses on competition between information from the two individual eyes, the two stimulus patterns or a combination of the two (Blake, 2001; Logothetis et al., 1996; Tong et al., 2006). Regardless of this debate, the competition clearly takes place between neural representations rather than between stimuli and the conflict leading to visual rivalry first presents itself when populations of neurons start coding for mutually exclusive perceptual interpretations. Without putting any claims on the exact content or location of this conflict, it seems likely that the most active neural population will 'win' the competition and eventually shape conscious perception. Increasing the stimulus strength of one of the two stimuli in binocular rivalry will increase the activity of the neural population coding for the corresponding percept, thereby increasing its chances to win the competition. Since this automatically decreases the probability that the opposing neural population wins the competition, it illustrates that even a unilateral manipulation of stimulus strength in binocular rivalry is still a relative manipulation at the relevant level of competing neural representations.

Our current findings are in line with previous studies that suggest that structure-from-motion is constructed through surface representations (Li & Kingdom, 1999; Nawrot & Blake, 1991b; Treue et al., 1995) and that the rivalry in a bistable structure-from-motion sphere takes place between the two surfaces competing for the 'front-location' in their depth ordering (Brouwer & van Ee, 2006). Our dot luminance manipulations bias the sphere stimulus towards the interpretation with the brightest surface in the front. The exact mechanism that establishes the bias is not crucial to our findings. Possible explanations could be that brighter objects are perceived to be closer to the observer (Schwartz & Sperling, 1983) or that lower contrast dots are perceived to move slower (Krekelberg et al., 2006). The surface-based interpretation of structure-from-motion rivalry is consistent with the recent finding of surface based attentional modulation of neuronal activity in the area MT of the monkey brain (Wan-nig et al., 2007) and the existence of a motion after effect specific for surface depth order (Sohn & Seiffert, 2006).



In conclusion, we have shown that perceptual rivalry in bistable structure-from-motion stimuli complies with all four of the Levelt-inspired propositions, much like binocular rivalry does. Our findings do not indicate that all relevant processes underlying binocular and bistable structure-from-motion take place at the same neural level. However, the strong similarities between the two do suggest that their output is produced by -at least partially- similar computational mechanisms, justifying a generalization of computational models of visual competition over binocular and perceptual rivalry.

## 2.6 Acknowledgements

We thank Pim Levelt, Jan Brascamp and Randolph Blake for providing valuable comments and support on earlier versions of this manuscript.



# Chapter 3

---

**Early interactions between neuronal adaptation & voluntary control determine perceptual choices in bistable vision**

*“You’re not paying Attention,” said the Hatter.  
“If you don’t pay him, you know, he won’t perform.”*

**Francis Crick,**  
**(quoted as after Lewis Carroll)**

**Published as**  
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), 16:1-18



---

*This chapter describes the first project I worked on when I started as a PhD student. As such, it encapsulates many 'first experiences'. The first time I worked with a computational model, the first time I programmed an experiment (or even programmed at all), the first time I had to concisely write everything down in a manuscript, the first time I had to juggle co-author suggestions, the first time I submitted a paper, the first time I had to deal with rejections of a manuscript, the first time I had to rework a paper based on referee comments, and the first time I published my work in a peer-reviewed journal. As one can imagine, it was one big learning experience and the people involved in this project have had a substantial influence on my further development as a researcher. Whereas I still really like the content and message of the paper, I am a bit more critical about the actual writing. Concise and clear writing is something most PhD students have to learn, and at the time of writing this manuscript this was very much true for me. The help that I received from my co-authors, both during data analysis and during the many rewrites of this manuscript has helped me to find my way around a dataset and improved my writing skills significantly.*

---



### 3.1 Abstract

At the onset of bistable stimuli the brain needs to choose which of the competing perceptual interpretations will first reach awareness. Stimulus manipulations and cognitive control both influence this choice process, but the underlying mechanisms and interactions remain poorly understood. Using intermittent presentation of bistable visual stimuli we demonstrate that short interruptions cause perceptual reversals upon the next presentation, whereas longer inter-stimulus intervals stabilize the percept. Top-down voluntary control biases this process but does not override the timing dependencies. Extending a recently introduced low-level neural model, we demonstrate that percept choice dynamics in bistable vision can be fully understood with interactions in early neural processing stages. Our model includes adaptive neural processing preceding a rivalry resolution stage with cross-inhibition, adaptation and an interaction of the adaptation levels with a neural baseline. Most importantly, our findings suggest that top-down attentional control over bistable stimuli interacts with low-level mechanisms at early levels of sensory processing before perceptual conflicts are resolved and perceptual choices about bistable stimuli are made.

### 3.2 Introduction

How does one choose between alternatives that are completely equal in every possible aspect? This classic problem is known in philosophy as *Buridan's ass* and tells the story of an ass that starves to death because it is incapable of choosing between two equally distant stacks of hay that are of exactly the same size and quality. A neurophysiological counterpart of this 14<sup>th</sup> century paradox can be found in the visual perception of bistable stimuli, containing equal evidence for two mutually exclusive percepts. When confronted with such a stimulus, the brain quickly chooses -in a non-random fashion- which interpretation will reach awareness (Leopold et al., 2002; Maier et al., 2003; Noest et al., 2007; Orbach et al., 1963). Classic decision theory, the accumulation of sensory signal towards a winner-takes-all decision, accounts for decisions based on unequal stimulus evidence (Gold & Shadlen, 2007) and thus predicts random percept-choices for bistable stimuli. The fact that the visual system makes non-random choices between interpretations with equivalent evidence implies that this choice-process must rely on dynamically evolving internal states. Indeed, it has been found that percept-choices at the onset of bistable vision depend on stimulus timing (Kanai & Verstraten, 2005; Noest et al., 2007; Orbach et al., 1966), the preceding perceptual history (Brascamp et al., 2007; Leopold et al., 2002; Maier et al., 2003) and both bottom-up (e.g. Chong & Blake, 2006; Mitchell et al., 2004) and top-down (Chong & Blake, 2006) attentional states.

A recent study (Noest et al., 2007) focused on finding the minimal neural mechanism that explains how the percept choice process can yield perceptual repetitions when stimuli are presented with relatively long intermittent blanks (Kanai et al., 2007a; Leopold et al., 2002; Maier et al., 2003; Noest et al., 2007; Orbach et al., 1966; Pearson & Clifford, 2004) and perceptual alternations when the blanks are short (Noest et al., 2007; Orbach et al., 1966). The resulting low-level model is based on a dynamic balance between the adaptation levels of cross-inhibiting, percept-related neural pools and an interaction of this adaptation with a near-threshold neural baseline (for details see Noest et al., 2007 and the Appendix of this chapter). This



interaction creates a small ‘head start’ for the more adapted neural population at the next stimulus onset that may overcome the effects of adaptation causing the more adapted neural population to become dominant and a perceptual repetition to occur. The model predicts and demonstrates that percept choices depend on the perceptual history and stimulus timing, both important determinants of the neuronal adaptation state. Importantly, the model only describes the dynamics of the single, core neural stage that resolves the rivalry. In reality, it is obvious that rivalry resolution does not occur at the first stage of neural processing and there will be stages of neural processing that precede rivalry resolution. These pre-rivalry stages exhibit general neural features like adaptation that will not change the nature of the percept-choice process itself but will nevertheless have clear effects. Reality thus forces us to consider these stages when comparing psychophysical data with model-driven predictions. In the current study, we extend the single-stage model of Noest et al. (2007) with the implementation of such adaptive pre-rivalry neural processing stages in their simplest possible form. They are assumed to have the same neural dynamics as the rivalry resolving stages, but lack cross-inhibition. The interaction of adaptation with the fixed neural baseline is functionally irrelevant for neural pools that do not interact (Noest et al., 2007) and is therefore not implemented in these stages. This model extension provides novel predictions that are consistent with the findings of our psychophysical experiments and help to understand the neural mechanism underlying percept choices at the onset of bistable stimuli.

An interesting feature of bistable stimuli under continuous viewing conditions is that subjects can, to a certain extent, voluntarily control their perception (Meng & Tong, 2004; van Ee et al., 2005). It is however not known whether similar voluntarily control can be exerted over percept-choices at stimulus onset. In this study we investigate the possible effect and underlying mechanisms of voluntary control over percept choices at the onset of bistable visual stimuli. The minimal neural mechanism that can generate top-down choice biases can directly be derived from existing, well-established literature. Advances in the field of visual attention (e.g. Alais & Blake, 1999; Chong & Blake, 2006; Ling & Carrasco, 2006; Ooi & He, 1999; Treue & Maunsell, 1999; Verstraten & Ashida, 2005; Womelsdorf et al., 2006) in addition to the finding of common neural substrates for attention and voluntary control (Slotnick & Yantis, 2005), and the demonstration of stimulus-feature dependency of voluntary control (Brouwer & van Ee, 2006; Suzuki & Peterson, 2000) suggest that voluntary control influences perception by attention-driven gain modulations at early stages of visual processing. As noted before, these stages must exist preceding the previously modeled rivalry-resolution stage. Such early gain modulations cause an imbalance in the input signal of the rivalry-resolving stage, a situation that has previously been used to explain the existence of classic after-effects that may override the influence of stimulus timing and perceptual history on percept-choices (Noest et al., 2007). Our extended model predicts that a difference between the input signals to the rivalry-resolving stage of only a few percent causes a substantial bias towards choosing the high-gain percept. Furthermore, the non-linearities in the relation between adaptation and percept-choices predict pre-stage gain imbalances to interact with stimulus timing and perceptual history, a notion that is confirmed by our second experiment in which we vary both stimulus timing and voluntary control instructions.

As a first step towards understanding the effects of voluntary control we investigate how neural stages that precede the rivalry resolution influence the percept-choices. Theoretically, adaptation that occurs prior to rivalry resolution will -to some extent- normalize the input to



the rivalry-resolving stage. Longer stimulus ON-time durations will cause more adaptation in these pre-rivalry stages, thereby decreasing the amplitude of their output signal (which is the rivalry stage's input) while keeping the 'stimulus energy' (duration  $\times$  amplitude) more or less constant (comparable to classic contrast normalization). Our model predicts that this pre-rivalry normalization process makes the build-up of adaptation in the rivalry stage independent of the stimulus ON-time duration. This effect emerges from the exact same minimal neural structure we will introduce to explain voluntary control, but yields more simple predictions since it involves only two parameters (presentation duration and interstimulus interval), rather than three (voluntary control added to the mix). In our first experiment two different types of bistable stimuli are used that may be resolved at different levels in the cortical hierarchy. It has been suggested that the visual competition for a set of binocular rivalry gratings takes place at a lower level of neural processing than that for perceptual rivalry such as a structure-from-motion defined sphere with an ambiguous rotation direction (Meng & Tong, 2004; van Dam & van Ee, 2006b; van Ee et al., 2005). We demonstrate that this difference between stimuli is reflected in the amount of pre-stage adaptation that reveals itself in the relationship between stimulus timing and percept-choices.

Our second, and main experiment directly probes the effect of voluntary control on percept-choices. The results demonstrate the existence of voluntary control over percept choice at the onset of bistable vision and are consistent with model predictions implementing small stimulus biases. They also confirm our prediction that voluntary control interacts with stimulus presentation dynamics and supports the proposal that top-down voluntary control indeed modulates neural gains at very early levels of visual processing.

### 3.3 Experiment 1: Percept-choices and stimulus timing

It has previously been shown that the stimulus ON and OFF durations determine whether the intermittent presentation of bistable visual stimuli result in perceptual repetitions or alternations (Noest et al., 2007). This effect depends crucially on the build-up of adaptation during stimulus ON-time ( $T_{on}$ ) and decay of adaptation during stimulus OFF-time ( $T_{off}$ ) (Figure 3-1C). These predictions are based on an input signal to the rivalry-resolving stage that remains constant in amplitude during the stimulus presentation duration. Since visual rivalry is not resolved at the level of the retina, there will be stages of neural processing that precede the neural stage where the rivalry resolution takes place. These stages -like any neural system- will be subject to adaptation, causing the amplitude of the input to the rivalry-resolving stage to decrease for increasing ON-time. If the amount of adaptation is different for the two individual percept-related neural processes in pre-rivalry stages it might result in unbalanced input to the rivalry stage leading to perceptual biases. But more in general, this pre-rivalry adaptation effectively functions as a normalization process that -if there is enough adaptation- will result in percept-choices that are independent of the stimulus ON-time (Figure 3-1D). In this first experiment we test this assumption using two bistable stimuli that are thought to rival at different processing levels. If our assumptions are correct we expect to find that percept-choices are relatively independent of stimulus presentation duration. Furthermore, if any effect of ON-time will still be present, it is more likely to occur with the binocular rivalry stimulus, which is generally regarded to be a more low-level kind of bistable stimulus.



## 3.4 Methods

### 3.4.1 Visual stimuli

Two different types of bistable stimuli were used: an ambiguously rotating, structure from motion (SFM) sphere and two orthogonally oriented gratings presented binocularly. The sphere was composed of two transparent layers of random white dot patterns on a black background moving in opposite directions with a sinusoidal speed profile (Figure 3-1A). Due to structure-from-motion effects these moving dots created the vivid impression of a three-dimensional rotating sphere (for a review, see Andersen & Bradley, 1998; Bradley et al., 1998; Dodd et al., 2001). As no unambiguous depth information was present in this stimulus the perceived rotation direction was bistable. The sphere was presented in the center of a computer monitor ( $1024 \times 768$ , 85 Hz) at a distance of 122 centimeters, with a yellow fixation square ( $4.2 \times 4.2$  arcmin) in its center. It was 2 degrees in diameter, while the dots were 2.8 arcmin and moved with a sinusoidal speed profile with a peak angular speed of 60 degrees per second. The luminance of the white dots was  $21.7 \text{ cd/m}^2$  and background luminance was  $0.13 \text{ cd/m}^2$ ; the dot density was 40 dots per squared visual degree. The dot lifetime was infinite, but at the start of each stimulus presentation the dots were randomly positioned to prevent tracking individual dots over stimulus presentations.

The binocular rivalry stimulus (Figure 3-1A) consisted of a dichoptically presented pair of sine wave gratings (spatial frequency = 1.75 cycles/degree) at orthogonal orientations using a mirror stereoscope. Each grating was multiplied by a two-dimensional Gaussian envelope ( $\sigma = 0.5$  degrees), resulting in an effective stimulus size of about 2.4 degrees in diameter. At the peak of the Gaussian function the luminance measured  $61 \text{ cd/m}^2$ ; the lowest luminance was  $\sim 0 \text{ cd/m}^2$ . The gratings were kept in anti-phase throughout the experiment, while the orientations remained the same for each eye. The gratings were presented on a grey background with a luminance of  $15 \text{ cd/m}^2$  and were accompanied by four dark grey lines (1.95 degrees by 0.12 degrees) presented to both eyes to support correct binocular fusion of the images. These lines had a luminance of  $3.7 \text{ cd/m}^2$  and were positioned 2.9 degrees from the gratings. The binocular rivalry stimulus was presented on a computer monitor ( $1280 \times 1024$ , 85 Hz) in the centre of the screen, 47 centimeters from the subject.

### 3.4.2 Experimental procedure and subjects

Five subjects participated in the ambiguously rotating sphere experiments (Experiment 1); four of these also participated in the binocular rivalry experiments. In both groups two subjects were completely naive with respect to the aims of the experiments. Subjects ranged in age between 22 and 39 years and had normal or corrected-to-normal visual acuity. They were seated with their head restrained by a head- and chinrest and were instructed to fixate on the square in the sphere experiment, or the middle of the screen in the binocular rivalry experiments. They were instructed to press one button when the front (near) surface reversed from a rightward to a leftward direction and to press another button when the opposite occurred. Subjects could occasionally also perceive the stimulus as either two convex surfaces or two concave surfaces that are sliding on top of each other (Chen & He, 2004; Hol et al., 2003). Nevertheless, they still perceive one surface sliding in front of the other, meaning that our instruction to report the direction of the front surface is clear and unambiguous. Stimuli



were presented intermittently, with different combinations of ON-time duration ( $T_{on}$ ) and OFF-time durations ( $T_{off}$ ) (Figure 3-1B). To avoid probing both the percept-choice process at stimulus onset and the percept-switch process present with continuous viewing, we chose to study ON periods short enough to prevent spontaneous percept switches during the stimulus presentation. Four different values of  $T_{on}$  (logarithmically spaced between 0.71 and 2.0 seconds) combined with 9 different values of  $T_{off}$  (logarithmically spaced between 0.125 and 2.0 seconds) for the SFM sphere and 11 different values of  $T_{off}$  (logarithmically spaced between 0.125 and 4 seconds) for the binocular grating resulted in 36 and 44 different conditions, respectively, that were each presented twice in pseudo-random order in blocks of 2 minutes, resulting in 40 to 288 stimulus presentations and perceptual choices per condition. Percepts were reported with a button press.

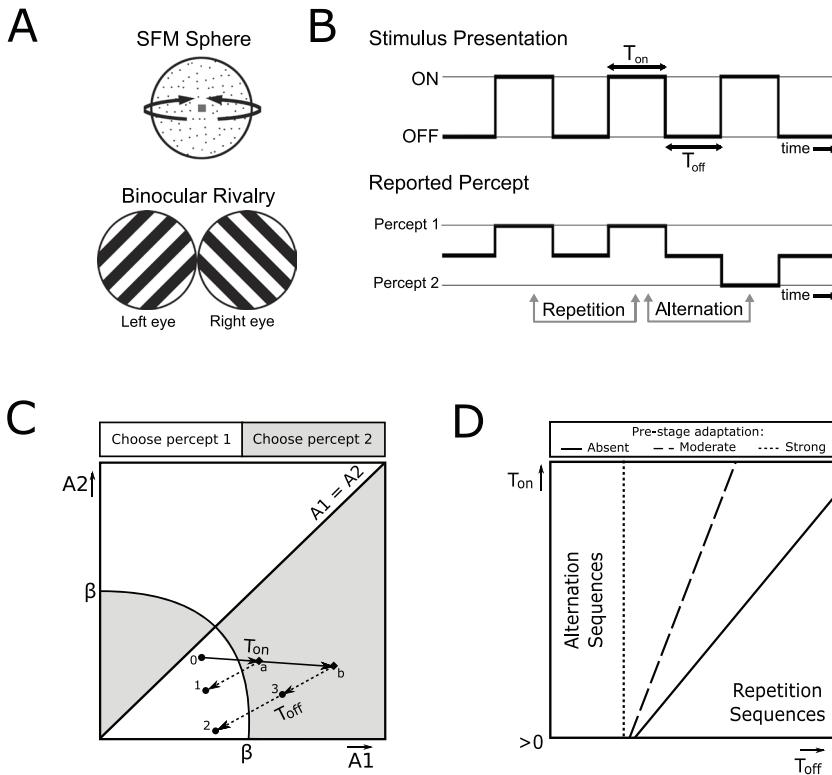
Subjects were instructed to respond only once per stimulus presentation and to report the first percept in the event a perceptual switch occurred during longer stimulus presentation (Note that our selection of ON-durations ensured us that this rarely happened. As an indication, less than 9 percent of perceptual durations with continuous viewing were shorter than 1 second). If uncertain, they were to choose the percept that appeared strongest. An extra condition was added in which the stimuli were presented continuously ( $T_{off} = 0$  seconds) for a block of two minutes that was presented twice, and subjects reported percept-switches with the same two buttons.

### 3.4.3 Data analysis

For continuous viewing, the number of percept switches per minute was calculated. For intermittent presentation, two subsequent stimulus presentations with different reported percepts was defined as an alternation (see Figure 3-1B) and both the number of alternations per minute and the fraction of the total number of trials in which an alternation occurred (alternation probability) were calculated. In these percept choice experiments, subjects were instructed not to report percept switches during the stimulus presentations. If subjects responded twice during one stimulus presentation, the second response was excluded. Trials in which the subject failed to respond were also excluded, along with their preceding and subsequent trials. Less than 3 percent of all trials were discarded based on one of these criteria.

The data were fitted with a descriptive two-dimensional cumulative Gaussian function (Equation 3-1). This function describes a surface of alternation probabilities in  $T_{on}/T_{off}$ -space, where  $x$  and  $y$  are the logarithms of  $T_{off}$  and  $T_{on}$  respectively. Parameter  $a$  represents the base of the surface,  $b$  the amplitude,  $c$  the transition point of  $T_{off}$  in which the subjects change from more perceptual alternation to a perceptual repetition regime (shift),  $d$  is the standard deviation of the cumulative Gaussian function and a measure of the steepness of this transition (sigma) and  $k$  represents the steepness of the change in the transition point for different  $T_{on}$  (slope).

$$P_{ChoiceAlt} = a + \frac{b}{2} \operatorname{erfc} \left( \frac{x - (c + ky)}{d\sqrt{2}} \right) \quad (\text{Eq. 3-1})$$



**FIGURE 3-1.** **A)** Schematic representation of the stimuli. The structure from motion (SFM) sphere consisted of two random dot patterns moving in opposite directions with a sinusoidal speed profile that gives rise to a vivid impression of a rotating three-dimensional sphere with an ambiguous direction of rotation. The binocular rivalry stimulus consisted of an orthogonal set of gratings presented to the individual eyes. **B)** Experimental procedure. Stimuli were presented intermittently with various durations of stimulus presentation ( $T_{on}$ ) and inter stimulus interval ( $T_{off}$ ) (upper panel). During each ON-time, the subject reported the current percept (lower panel). Two subsequent percepts that are opposite are defined as an alternation; two similar percepts are defined as a repetition. **C)** The model of Noest et al. (2007) predicts the percept-choice to depend on the adaptation states of the rivaling neural populations at stimulus onset and the neural baseline parameter  $\beta$ . These adaptation states in turn depend on the amount of adaptation build-up during  $T_{on}$  and decay during  $T_{off}$ . If at stimulus onset '0', percept 1 is being perceived, the corresponding adaptation state  $A_1$  will increase (solid arrow to the right) during  $T_{on}$  and decay during  $T_{off}$  (dotted arrow to the left). With short  $T_{on}$  (a), a short  $T_{off}$  suffices to get perceptual repetitions (1), while with longer  $T_{on}$ 's (b) the percept-choice depends on the length of  $T_{off}$ : short  $T_{off}$ 's give alternations (3), long  $T_{off}$ 's repetitions (2). **D)** With constant input to the rivalry resolving stage the transition from alternation to repetition sequences involves a positive correlation between  $T_{on}$  and  $T_{off}$  (solid line). Adaptation prior to the rivalry resolution stage normalizes the input and thus the amount of adaptation that is built up (the solid arrow to the right in C) would be of a fixed length). The border between alternation and repetition sequence areas is steepened with increasing preceding adaptation until it becomes vertical for complete pre-stage adaptation (dotted line).

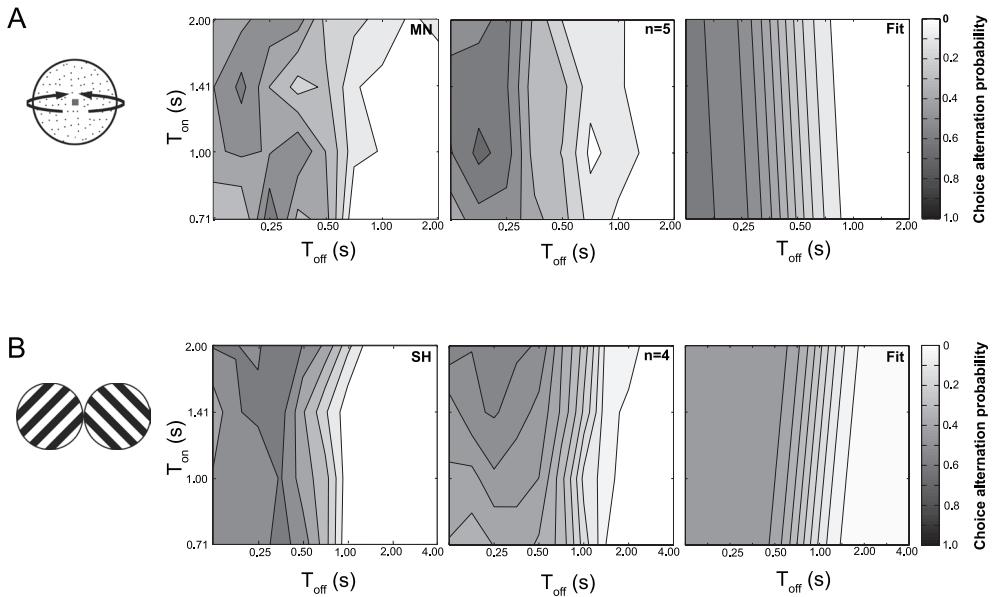


### 3.5 Results

If percept-choices crucially depend on the adaptation states of the underlying neuronal populations, this should become obvious from manipulations of the adaptation build-up during stimulus presentation ( $T_{on}$ ) and decay during the interstimulus interval ( $T_{off}$ ). Figure 3-2A illustrates how alternation probabilities for perceptual decisions depend on both the  $T_{on}$  and  $T_{off}$  for the experiment with the ambiguously rotating sphere. The black coloring indicates high alternation probabilities, while the white coloring indicates low alternation probabilities. In the left panel, the data are provided for one typical subject. The same qualitative pattern of  $T_{off}$ -dependency is present for all subjects, indicated by the average plot of all five subjects in the center panel. Short  $T_{off}$ 's lead to high alternation probabilities, whereas at longer  $T_{off}$ 's the alternation probability declines and perceptual stabilization occurs. Furthermore, these figures indicate that the change from alternation to repetition depends primarily on  $T_{off}$ , not  $T_{on}$ . A two-way ANOVA confirms this notion; for all individual subjects -and for the group data-  $T_{off}$  significantly influences the alternation probability ( $p < 0.0001$ ), whereas  $T_{on}$  does not ( $p > 0.20$ ). These findings confirm our predictions for the effective contrast normalization by pre-stage adaptation. The noisy boundary between the alternation and repetition regime areas likely results from inevitable noise in the rivalry-resolving system (Brascamp et al., 2006).

To quantify the results, we performed a descriptive fit on the data using a cumulative Gaussian function (see Methods, Equation 3-1). The fitted data ( $R^2 = 0.94$ ) of the average of all five subjects are shown in the right panel of Figure 3-2A. The transition  $T_{off}$  time for alternation to repetition regimes of percept choices ( $c$  in Equation 3-1) averages 0.48 seconds ( $\pm 0.10$  seconds). If percept-choices with intermittent presentation and percept-switches under continuous presentation are basically manifestations of the same underlying process, one might expect that the transition times from repetition to alternation regimes (intermittent presentation) and the average percept durations during continuous viewing roughly coincide, or are at least correlated. However, the transition time is roughly 14 times smaller than the mean percept duration under continuous viewing conditions for the same subjects (6.70, s.d.  $\pm 3.62$  seconds) and the absence of a correlation between the two measures ( $r^2 = 0.01$ ) adds further evidence to the suggestion that percept-choice (intermittent presentation) and percept-switch (continuous presentation) are fundamentally different processes. The average steepness of the transition point ( $d$  in Equation 3-1) for the five subjects was  $1.20 (\pm 0.73)$ , corresponding to 0.13 seconds ( $\pm 0.02$  seconds), and the average steepness of the change in transition point of  $T_{on}$  ( $k$  in Equation 3-1) was  $-0.13 (\pm 0.22)$ . This quantification indicates a vertical border and supports the statistical findings that the transition point depends on  $T_{off}$  not  $T_{on}$ .

A qualitatively similar pattern is present for the binocular gratings (Figure 3-2B). The figure provides the data from a typical subject (left panel), the averaged group data (center panel), and the fit to the data (right panel,  $R^2 = 0.95$ ). A two-way ANOVA confirmed that the effect of  $T_{off}$  on the alternation probability is significant ( $p < 0.001$ ) for all subjects as well as the group data. Alternation probabilities decrease with increasing  $T_{off}$  ( $a = 0.02 \pm 0.03$ ;  $b = 0.48 \pm 0.04$ ); these values are in the same order of magnitude as the values for the SFM sphere ( $a = 0.06 \pm 0.04$ ;  $b = 0.54 \pm 0.07$ ). The average transition moment from an alternating to a repetitive regime lies at a  $T_{off}$  of 0.71 seconds ( $\pm 0.12$  seconds), which is roughly 8 times shorter than, and not correlated ( $R^2 = 0.46$ ) with, the average percept duration during continuous viewing conditions (5.35, s.d.  $\pm 1.15$  seconds). Compared to the SFM spheres the important differ-



**FIGURE 3-2.** Alternation probabilities for intermittently presented bistable stimuli generally depend on presentation duration ( $T_{on}$ ) and inter stimulus interval ( $T_{off}$ ) for A) the ambiguous sphere and B) the binocular rivalry stimulus. The left panels depict the data from a typical single subject, center panels show the averaged group data, and right panels show the fit. The influence of  $T_{on}$  is quite small, leaving choice alternation probabilities to be determined predominantly by  $T_{off}$ .

ence is that, with binocular gratings,  $T_{on}$  has a small but significant effect on the alternation probability for all subjects as well as the group data (two-way ANOVA,  $p < 0.01$ ). The average  $k$  value is  $0.27 (\pm 0.18)$ , indicating a slightly sloped border between alternation- and repetition regimes. The effect of  $T_{on}$  is smaller than that of  $T_{off}$  ( $d = 1.12 \pm 0.34$ ), but its significance reveals that stages preceding the resolution of binocular rivalry are subject to less adaptation than the stages preceding the rivalry resolution for the sphere. Apparently the rivalry stage input signal has not been fully normalized in preceding neural stages.

While repetition probabilities reach unity for all conditions at long  $T_{off}$ , the average maximum alternation probability for the ambiguously rotating sphere and binocular rivalry stimulus at short  $T_{off}$ 's are 0.33 and 0.47, respectively. One could therefore presume that, instead of reaching an alternation regime, the system simply reaches a regime of random percept choice at short  $T_{off}$ . An analysis of the occurrence of longer sequences of alternating percepts however reveals that subjects truly perceived sequences of alternations at short  $T_{off}$  rather than random percepts (Figure 3-5A).

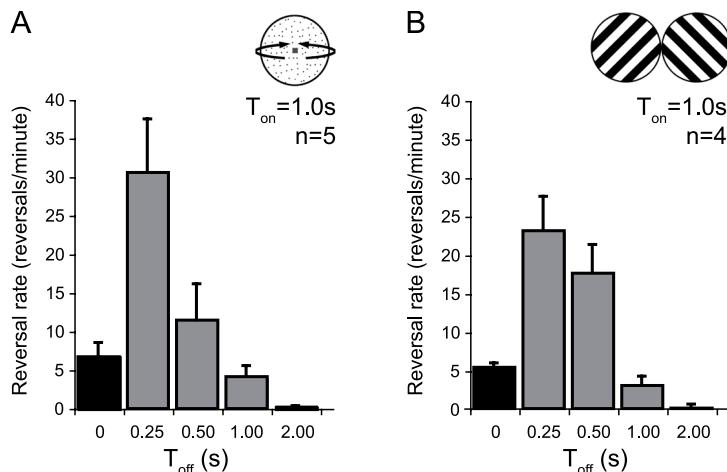
Finally, in Figure 3-3 we plotted the reversal rates expressed in alternations per minute for the continuous presentation and all the intermittent presentations with a  $T_{on}$  of 1 second to directly compare our results with a previous study that reported perceptual stabilization due to intermittent stimulus presentation (Figure 5B in Leopold et al., 2002). For both stimuli, the reversal rates for intermittent presentation decline with increasing  $T_{off}$  and reversal rates with continuous viewing lie between those acquired with  $T_{off}$ 's of one second and half a second.



Our range of  $T_{\text{off}}$ 's clearly demonstrates that intermittent presentation can result in percept-choice alternations and percept-choice repetitions (stabilization), ultimately depending on the length of the intermittent interval.

### 3.6 Experiment 2: Percept-choices and voluntary control

In this second experiment, we directly probe the effects of voluntary control. Whereas voluntary control over continuously presented bistable stimuli has been known for some time (Meng & Tong, 2004; van Ee et al., 2005), it has never been shown for percept choices at stimulus onset. Interestingly, there are reports about the influence of voluntary object-based attention on the percept-choices at the onset of binocular rivalry stimuli (Chong & Blake, 2006) and we argued that there is considerable evidence to interpret voluntary control as an attentional gain modulator that effectively biases the input of the rivalry-resolving neuronal stage towards one of two perceptual interpretations. For this interpretation, our new extended model provides some clear predictions. The topography of the ‘percept-choice-map’ as related to the adaptation states (Figure 3-1C) changes under biased inputs (Noest et al., 2007)a. The area where the favoured percept will be chosen increases in size, whereas those of the unfavoured percept shrink (Figure 3-4B). This shrinkage depends on the size of the bias and is more dramatic in the area of low adaptation levels compared to the areas of high adaptation levels. These novel predictions imply that voluntary control should interact with the adaptation states and become more effective when stimulus OFF-times increase (lower adaptation levels).



**FIGURE 3-3.** Mean number of reversals per minute (with s.e.m.) of five subjects for continuous ( $T_{\text{off}} = 0$  seconds) and intermittently presented stimuli ( $T_{\text{off}} > 0$  seconds) for A) the ambiguous sphere and B) the binocular rivalry stimulus. Stimulus presentation duration ( $T_{\text{on}}$ ) was 1 sec for the intermittently presented stimuli. The effect of intermittent presentation depends on  $T_{\text{off}}$  and can result in either lower or higher reversal rates than in the continuous viewing condition.



### 3.7 Methods

In this experiment, we used the same stimuli and experimental procedure as in the first experiment. The only difference is that each block of stimulus presentations was now preceded by one of five possible instructions with respect to the voluntary control: (1) alternate; (2) hold the current percept; (4) hold leftwards or (4) rightwards rotation/orientation and; (5) The baseline condition, exerting no control (passive condition). In our first experiment we have shown that alternation probabilities for this stimulus are independent of  $T_{on}$ , therefore we only varied the  $T_{off}$  in this experiment.  $T_{on}$  was always 1 second, and  $T_{off}$  was pseudo-randomly chosen from 11 (sphere) or 5 (binocular rivalry) possible durations in the same range as in experiment 1. All subjects' performance with all five instructions was also recorded with continuous presentation of the stimulus ( $T_{off} = 0$  seconds), presented twice in blocks of two minutes. Six subjects participated in the experiment with the sphere, including four who had also participated in experiment 1. One of these six subjects (MS) was not successful in exerting voluntary control and was excluded from the group analysis. Four subjects participated in the binocular rivalry experiment. Subjects ranged in age between 22 and 39 years and had normal or corrected-to-normal visual acuity. For both stimuli-groups, two of the subjects were naive with respect to the purpose of the experiments. As a descriptive fit to the data we used a one-dimensional version of Equation 3-1:

$$P_{ChoiceAlt} = a + \frac{b}{2} \operatorname{erfc} \left( \frac{x - c}{d\sqrt{2}} \right) \quad (\text{Eq. 3-2})$$

In this equation,  $x$  once again represents  $\log(T_{off})$ ,  $a$  is the baseline,  $b$  is the amplitude,  $c$  is the shift that represents the transition point and  $d$  is the standard deviation of the cumulative Gaussian function.

### 3.8 Results

It is known that voluntary control over perceptual states for bistable stimuli is possible when the stimulus is shown continuously (Brouwer & van Ee, 2006; Hol et al., 2003), but it was still unclear whether observers have voluntary control when stimuli are presented using varying  $T_{on}/T_{off}$  sequences. Figures 3-4C (sphere) and 3-4D (binocular rivalry) demonstrate that when subjects were instructed to view the stimulus passively, the same  $T_{off}$  dependency of alternation probability occurs as in Experiment 1. High alternation probabilities occur at short  $T_{off}$ 's and low alternation probabilities at high  $T_{off}$ . Perhaps more surprising is the presence of the same qualitative pattern for situations in which subjects were instructed to exert voluntary control to perceive either as many alternations or as many repetitions as possible. Furthermore, the average choice alternation probabilities of all subjects as a function of  $T_{off}$  clearly shifted in the direction of the instruction to repeat or alternate demonstrating the successful exertion of voluntary control. We used a balanced two-way ANOVA to test the influence of instruction and  $T_{off}$  on alternation probability for statistical significance. For the sphere, both the effects of instruction ( $p < 0.001$ ) and  $T_{off}$  ( $p < 0.001$ ) are significant, no interaction is evident between the two factors ( $p = 0.88$ ). For all three individual instructions, there is



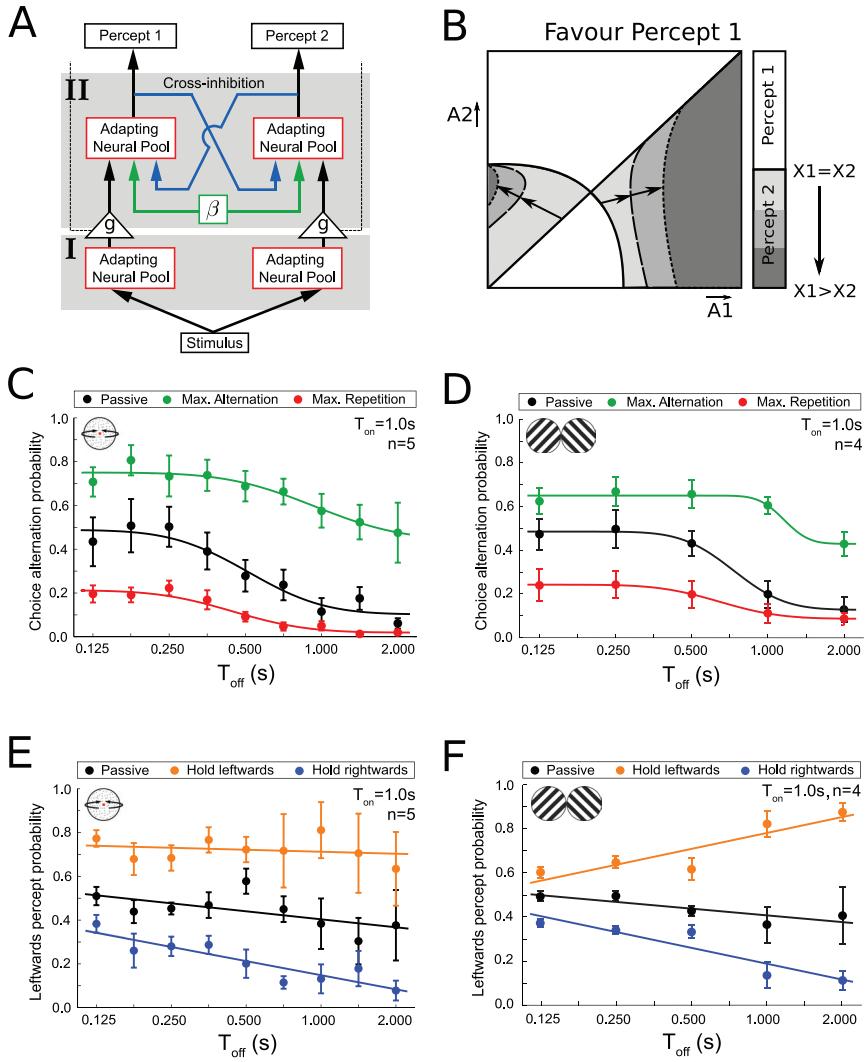
a significant effect of  $T_{\text{off}}$  ( $p < 0.01$ ) and a significant difference between individual subjects ( $p < 0.01$ ). Despite this difference between subjects, all individual subjects demonstrated a significant instruction effect ( $p < 0.01$ ); all but one (CK) had a significant  $T_{\text{off}}$  effect ( $p < 0.01$ ). Results for the binocular rivalry stimulus are highly similar. A two-way ANOVA once again revealed significant effects of instruction ( $p < 0.001$ ) and  $T_{\text{off}}$  ( $p < 0.001$ ) and the absence of a significant interaction between the two ( $p = 0.37$ ). For all the three instructions this  $T_{\text{off}}$  effect was significant ( $p < 0.001$ ). All the individual subjects demonstrated a significant instruction effect ( $p < 0.001$ ) and all but one (RW) a significant effect of  $T_{\text{off}}$  ( $p < 0.01$ ).

We fitted the averaged data to a cumulative Gaussian function. For the sphere, the average quality of the fit for individual subjects was good ( $R^2 = 0.72$ , s.d.  $\pm 0.29$ ), but the individual fits for binocular rivalry lacked statistical power. For the group data, we performed weighted fits. The individual data points received a weight-factor proportional to the inverse of their squared standard error. The average quality of this fit over the three conditions was very good for both stimuli (Sphere:  $R^2 = 0.93$ , s.d.  $\pm 0.01$ , Binocular rivalry:  $R^2 = 0.93$ , s.d.  $\pm 0.01$ ). Figures 3-4C and 4d demonstrate that for both stimuli and both types of instructions, the fitted curves shifted vertically with respect to the “*passive*” curve, toward their intended goals of maximal and minimal alternation probabilities. This can also be seen in the estimated parameters of  $a$ , representing the base of the fitted curve (Sphere:  $a_{\text{passive}} = 0.10 \pm 0.01$ ;  $a_{\text{repetition}} = 0.02 \pm 0.03$ ;  $a_{\text{alternation}} = 0.44 \pm 0.02$ ; Binocular rivalry:  $a_{\text{passive}} = 0.13$ ;  $a_{\text{repetition}} = 0.08$ ;  $a_{\text{alternation}} = 0.43$ ). We must be cautious in comparing other parameter estimates, such as the transition point ( $c$ ) or amplitude ( $b$ ), between the different instruction conditions since the shapes of the curves for repetition and alternation seem to be influenced by saturation effects.

Figures 3-4E and 3-4F show the fraction of trials in which subjects reported leftward percepts as a function of  $T_{\text{off}}$  when they were instructed to hold one specific percept: a leftwards rotating vs. a rightwards rotating sphere or a leftwards tilted vs. a rightwards tilted grating. We compared these two instructions with the passive condition. The data points indicate averaged values of leftwards percept probabilities for five subjects while error flags depict the standard errors of the mean. To fit the data, we used weighted linear regressions. Data points received a weight factor proportional to the inverse of their squared standard errors.

For the sphere, the data points for the passive condition in Figure 3-4E are positioned around 0.5, but are negatively correlated with  $T_{\text{off}}$  indicating a small bias towards rightwards percepts, which becomes obvious especially at longer  $T_{\text{off}}$ . This is the reason why the curves for ‘hold leftwards’ and ‘hold rightwards’ are not mirror images of each other in the chance-axis (probability = 0.5). An additional analysis confirmed the presence of a small bias for rightward percepts (data not provided). The curves for holding one percept are shifted in the intended direction indicating that subjects are able to exert voluntary control over their percept choice. A two-way ANOVA on the data-set reveals that the instruction effect is significant for all individual subjects ( $p < 0.02$ ), as well as for the group data ( $p < 0.01$ ). The effect of  $T_{\text{off}}$  was significant only for two subjects ( $p < 0.02$ ), and when we examined the instruction conditions separately, a significant effect of  $T_{\text{off}}$  occurred only in the “hold rightwards” condition ( $p < 0.01$ ).

From Experiment 1 we know that the number of alternations drops with increasing  $T_{\text{off}}$ , which results in the larger standard errors at longer  $T_{\text{off}}$ ’s. An analysis of longer sequences of alternations revealed that subjects are not reporting random percepts at short  $T_{\text{off}}$ ’s but true sequences of alternations or repetitions (Figure 3-5B). It also demonstrates that the fraction



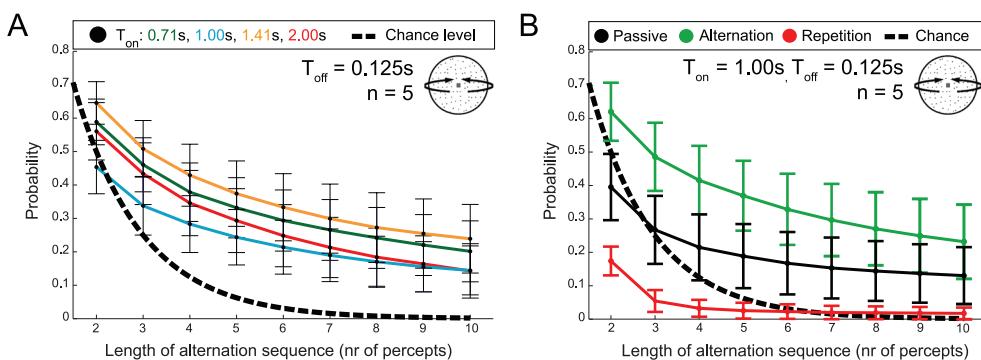
**FIGURE 3-4.** **A)** A schematic representation of our two-stage model of perceptual decisions. The model consists of a pre-processing stage (I) where adapting neural pools process perceptual representations without inhibiting each other. The second, rivalry-resolving stage (II) has similar adapting neural pools but these pools do exhibit cross-inhibition and an interaction with a neural baseline ( $\beta$ ). Top-down biases are implemented as gain modulations (dashed lines,  $g$ ) preceding the rivalry-resolving stage. **B)** An input bias ( $X_1 > X_2$ ), caused by a gain imbalance re-shapes the percept-choice map in adaptation state space. The areas in which the unfavoured percept gets chosen shrink (grey shades and arrows) in size. This shrinkage depends on the size of the imbalance and is faster in the area for lower adaptation levels (on the left side of the plot). **C/D)** Average alternation probability for the sphere C) and binocular rivalry stimulus D) for all subjects as a function of  $T_{off}$  and three different instructions: maximize number of alternations (green), maximize number of repetitions (red), and view passively (black).  $T_{on}$  was 1 second in all conditions. Thick lines are fitted one-dimensional cumulative Gaussian functions. **E/F)** Average probability of leftwards rotation (sphere, E) or orientation (binocular gratings, F) percepts for all subjects as a function of  $T_{off}$  for three different instructions: hold percept leftwards (orange), hold percept rightwards (blue), and passive viewing (black). Thick lines are linear regression fits. Error bars are the SEMs. A profound effect of voluntary control is evident in all four plots.



of alternations that is still present when subjects try to maximize the number of repetitions is predominantly due to very short repetition sequences (Figure 3-5B). The opposite is true when subjects try to maximize repetitions.

The results for the binocular rivalry stimulus, shown in Figure 3-4F, are very similar to the findings with the sphere. A two-way ANOVA revealed that the effects of instruction are highly significant ( $p < 0.001$ ) and although the  $T_{\text{off}}$  effect only approaches significance ( $p = 0.06$ ), when investigated separately all instructions did show a significant  $T_{\text{off}}$ . The interaction between instruction and  $T_{\text{off}}$  was also significant ( $p < 0.001$ ). This pattern is present with all subjects ( $p < 0.01$ ). The weighted linear regressions reveal a small bias in passive condition, but the ‘hold-curves’ are near perfect mirror images in the chance axis.

In addition, for the continuous viewing condition ( $T_{\text{off}} = 0$  seconds), all subjects were successful in exerting voluntary control over the stimuli. Mean percept durations were significantly shorter (ANOVA,  $p < 0.01$ ) when subjects tried to maximize their alternation probability compared to when they tried to minimize the alternation probability (for the sphere on average 4.05 times shorter, s.d.  $\pm 2.10$ ; for binocular rivalry on average 2.58 times shorter, s.d.  $\pm 1.27$ ). When instructed to hold one of the two percepts, all subjects were again successful. Mean percept durations were significantly longer (ANOVA,  $p < 0.01$ ) for the percepts that subjects were instructed to hold, compared to the alternative perceptual interpretation (for the sphere on average 1.81 times longer, s.d.  $\pm 0.32$ ; for binocular rivalry on average 1.47 times longer, s.d.  $\pm 0.31$ ).



**FIGURE 3-5.** When longer sequences of alternations are taken into account it becomes obvious that subjects are not reporting random percepts, but perceive true alternation sequences even if the alternation probability is around chance level at short  $T_{\text{off}}$ . The dashed line in both panels represents the chance level for the occurrence of alternation sequences of increasing length. **A)** Thin colored lines represent the probabilities of alternation sequences of increasing length in the data of Experiment 1, for the sphere with a  $T_{\text{off}}$  of 0.125 seconds. Even though the lines are around chance level for the shortest possible alternation sequence of 2 percepts, they are all well above chance for longer alternation sequences. **B)** The data from experiment 2 for three instructions; “passive viewing” (black solid line), “maximize alternations” (green) and “maximize repetitions” (red) clearly demonstrate that subjects perceive sequences of alternations rather than random percepts at a  $T_{\text{off}}$  of 0.125 seconds. It becomes clear that any alternations that are still present when subjects maximize the number of repetitions are predominantly due to very short sequences of alternations. Error bars represent SEMs.



We believe that the novel implementation of adaptive pre-processing stages with a gain factor under the influence of voluntary control is physiologically plausible, widely supported by existing evidence, and elegant in its simplicity. The tight qualitative match between the extended model predictions and our data (see the appendix for numerical simulations) suggests that the mechanism of voluntary control is a top-down, attention-driven gain modulation on the independent perceptual interpretations of a bistable stimulus (Blaser et al., 1999; Suzuki & Peterson, 2000; van Ee et al., 2006), influencing the percept-choice process at an early stage of visual processing.

### 3.9 Eye movements

It has been found that even though eye movements are not essential for perceptual alternations during continuous viewing, they can be facilitating (Brouwer & van Ee, 2007; Brouwer & van Ee, 2006; Toppino, 2003; van Dam & van Ee, 2006a; van Dam & van Ee, 2006b). Alternations could be triggered, either due to the acquisition of ‘fresh’ neural tissue after an eye movement (Blake et al., 2003) or by tracking individual dots resulting in optokinetic nystagmus. We controlled for the possible effects of eye movements on voluntary control with an extra experiment resembling experiment 2, for which we recorded eye-movements. Three subjects that also participated in the first two experiments viewed a rotating sphere consisting of white and black dots on a grey background ( $30.7 \text{ cd/m}^2$ ).  $T_{\text{on}}$  was always 1 second, but three  $T_{\text{off}}$ ’s were used (0.25, 0.71 and 2.00 seconds). All four voluntary control instructions were present (maximize alternations/repetitions, hold left/right) and each experimental condition was presented twice in random order in blocks of 60 seconds. Gaze position and eye movements were recorded using a head-mounted infrared camera based Eyelink system (SMI/SR Research, Berlin, Germany) with a sample frequency of 250 Hz. The raw gaze position data were median-filtered and converted to Fick-angles; eye movement velocity was calculated using a five point sliding window (Engbert & Kliegl, 2003; van Dam & van Ee, 2005). After blink-removal, saccades were detected and filtered using a separate threshold for horizontal and vertical velocities. Alternation probabilities were calculated the same way as in the previous two experiments.

All subjects were successful in exerting voluntary control. Both the effect of controlling the number of alternations ( $p < 0.02$ ) and choosing one of two percepts ( $p < 0.01$ ) was significant for all subjects and the group data ( $p < 0.001$ ). The average magnitude of voluntary control can be expressed as the difference in alternation- or percept probability between two opposing instructions (maximize alternations vs. maximize repetitions and hold left vs. hold right). These magnitudes were comparable in the control experiments with eye movement recording (0.47, s.d.  $\pm 0.18$  for controlling the alternation probability; 0.52, s.d.  $\pm 0.14$  for choosing a percept) and in the main experiment without eye movements monitoring (0.53, s.d.  $\pm 0.21$  and 0.52, s.d.  $\pm 0.29$  respectively). The occurrence of blinks was very low (on average 3.15 blinks/minute), making it highly unlikely that they influenced perception during the much more frequent stimulus presentations. The gaze position was split up in an  $x$ - and  $y$ -coordinate; only one subject showed a significant difference in  $x$ -coordinate between perceptual states ( $p < 0.01$ ). When she reported the sphere to rotate leftwards, the mean gaze position was 0.03 degrees left of the fixation dot, while rightwards-rotating percepts had a mean gaze



position 0.01 degrees right of the fixation dot. No significant difference in y-positions was evident for any of the subjects ( $p > 0.15$ ). None of the subjects demonstrated significant differences in gaze position over different instructions ( $p > 0.65$ ), or inter-stimulus intervals ( $p > 0.19$ ). None of the subjects showed a significant difference in saccade direction between the voluntary control-related task instructions (2-way ANOVA, AR:  $p = 0.15$ ; CK:  $p = 0.07$ ; MN:  $p = 0.14$ ) or perceptual states (AR:  $p = 0.46$ ; CK:  $p = 0.61$ ; MN:  $p = 0.95$ ). Two subjects had a significant difference in saccade directions ( $p < 0.01$ ) between  $T_{\text{off}}$ -conditions, but the mean directions remained in the same quadrant of directions (AR: 96, 108, and 132 degrees; MN: 159, 150, and 167 degrees) making it highly unlikely that it played a role of any importance in the percept-choice dynamics. Eye movements could very well play a role in the perception of bistable stimuli and the exertion of voluntary control under less constrained conditions, but the results presented here clearly indicate that the successful exertion of voluntary control over an intermittently presented sphere cannot be explained by mechanisms solely based on eye movements.

## 3.10 Discussion

### 3.10.1 Stimulus timing

We studied the mechanisms underlying perceptual choices at the onset of bistable stimuli. In accordance with previous studies (Kanai & Verstraten, 2005; Leopold et al., 2002; Noest et al., 2007; Orbach et al., 1966; Pearson & Clifford, 2004), we found that perceptual repetitions occur at stimulus OFF-times lasting longer than approximately half a second. At short OFF durations the opposite of percept repetition occurs and alternation rates increase dramatically. We demonstrate that this regime change primarily reflects the balance between true alternation and repetition sequences and not random percept-choices at short OFF durations. Within our range of stimulus ON- and OFF durations, the transition from perceptual repetition- to alternation regimes depends primarily on the stimulus OFF duration, not the ON period. The use of longer ON-times for the study of perceptual choices at stimulus onset is not possible because spontaneous perceptual switches (a fundamentally different process) will then occur within the stimulus presentation episodes.

Models of perceptual bistability invariably include two terms: 1) Mutual inhibition of two percept-related neural pools (cross-inhibition)(Bradley et al., 1998; Dodd et al., 2001; Leopold & Logothetis, 1996; Logothetis & Schall, 1989; Maier et al., 2007; Parker et al., 2002; Wilson, 2007), and 2) adaptation of these neural pools at a rate that depends on the magnitude of their activity (e.g. Blake et al., 2003; Lankheet, 2006; Wilson, 2007). These two components alone predict perceptual alternations at any OFF-time duration, and thus cannot account for percept repetitions with long OFF-times. Recently a single-stage, low-level model was proposed that introduces an interaction of the adaptation level with a fixed neural baseline in addition to adaptation and cross-inhibition (Noest et al., 2007). This model can explain both perceptual alternations with short OFF-times and perceptual stabilization with longer OFF-times. Our experiments reveal that the implementation of an extra adaptation stage with a gain modulation mechanism preceding the stage of rivalry resolution is necessary and sufficient for the model to explain our findings (Figure 3-4A) and conform to realistic constraints. The single-stage, low-level model is based on a dynamic balance between the adaptation levels



of cross-inhibiting, percept-related neural pools and an interaction of adaptation with a near-threshold neural baseline (for details see Noest et al., 2007 and the Appendix of this chapter). In essence, this interaction term introduces an additional effect of adaptation by producing an elevated baseline activity and a reduced effective membrane time-constant of the adapted population. For low adaptation levels this baseline activity causes the more adapted neural population to ‘win’ the rivalry and percept-choice repetitions occur (Figure 3-1C). This explanation for perceptual stabilization has no need for an implicit perceptual memory (Kanai & Verstraten, 2006; Leopold et al., 2002; Maier et al., 2003; Pearson & Clifford, 2004), unless one would propose that the low-level balance between the key terms of the model *is* in fact the implicit perceptual memory trace (see also van de Grind et al., 2004). Consequently, a major difference from classic accounts of perceptual memory would then be that this type of perceptual memory does not require any active monitoring of the current percept, but is a passive result of the shifting balance between neural pools.

Despite large differences in stimulus composition between a set of binocular rivalry gratings and an ambiguously rotating sphere, the relation between perceptual choices and stimulus timing was very similar. Interestingly, the only difference that we observed is the role of stimulus ON-time. Whereas the percept-choice dynamics of the SFM sphere are independent of  $T_{on}$ , the percept-choice dynamics of the binocular rivalry stimulus do show a small, but significant, ON-time dependency. Our model can account for this observation by assuming different amounts of adaptation preceding the rivalry resolution. Apparently our binocular rivalry stimulus has undergone less pre-rivalry-stage adaptation than the sphere stimulus. Even though we can only make firm statements about the amounts of adaptation, it is likely that less adaptation also indicates less neural processing. This interpretation would be consistent with previous suggestions that binocular rivalry is a much more low-level, stimulus based type of rivalry compared to the perceptual rivalry present in the SFM sphere or a Necker cube (Meng & Tong, 2004; van Dam & van Ee, 2006b; van Ee et al., 2005).

Previous studies on the perception of bistable stimuli have mainly focused on spontaneous perceptual alternations during prolonged presentation (for a review see Blake & Logothetis, 2002). Although this percept-switching process may be closely related to the percept-choice investigated here (Pearson & Clifford, 2004), the two are not necessarily based on the same mechanism (Carter & Cavanagh, 2007; Long & Toppino, 2004; Noest et al., 2007). In our study, we find a very low or absent correlation between the average percept-switch duration and the point of transition from an alternation to a repetition regime, which supports the notion that the percept-choices at stimulus onset and percept-switches during continuous viewing are fundamentally different processes.

### 3.10.2 Voluntary control

The perception of bistable stimuli can be influenced by attention. This attentional modulation can be bottom-up, stimulus driven (Carter & Cavanagh, 2007; Chong & Blake, 2006; Hancock & Andrews, 2007; Kamphuisen et al., 2007; Mitchell et al., 2004; Ooi & He, 1999) and top-down, instruction driven (Chong et al., 2005; Hol et al., 2003; Lack, 1978; Meng & Tong, 2004; Peterson & Hochberg, 1983; Toppino, 2003; van Ee et al., 2005; Von Helmholtz, 1924). There is also evidence for an interaction of bottom-up and top-down modulation (Brouwer & van Ee, 2006; Suzuki & Grabowecky, 2007). Interestingly, some evidence suggests that for



percept-choices top-down attention is equivalent to a moderate bottom-up change in stimulus contrast (Chong & Blake, 2006; Chong et al., 2005). Recent advances in psychophysics (Blaser et al., 1999; Boynton, 2005; Ling & Carrasco, 2006) and physiology (for a review, see Reynolds & Chelazzi, 2004; Treue & Maunsell, 1999) have converged to conclude that the neural mechanism underlying attention involves shifting the balance between neural gains of feature-selective neural pools, even in early stages of visual processing (Felisberti & Zanker, 2005; Saalmann et al., 2007; Treue & Martinez Trujillo, 1999; Treue & Maunsell, 1999; Wannig et al., 2007). To some extent, voluntary control and spatial attention shifts are associated with common activity in the posterior parietal cortex, suggesting voluntary control to be at least partially based on shifting the focus of attention (Slotnick & Yantis, 2005). The finding that the amount of voluntary control a subject can exert depends on stimulus features (Brouwer & van Ee, 2006; Suzuki & Peterson, 2000), and the demonstration of independent control over the two individual rivaling percepts (van Ee et al., 2006) further add to the suggestion that voluntary control influences perception by independent, attention-driven gain modulations at early stages in visual processing.

Concerning voluntary control, we demonstrated: (1) that voluntary control over perceptual decisions at the onset of bistable stimuli exists for both binocular rivalry and a bistable rotating sphere, (2) that it does not override, the  $T_{off}$  dependency shown in our first experiment but interacts with it, and (3) that it is not driven by eye-movements. Our subjects only indicated pure dominant percepts and no mixtures. This means that theoretically they may have changed their reporting criteria in the voluntary control conditions instead of their perception. We took special care in avoiding mixed percepts with our sphere stimulus (see methods in Experiment 1), but for the binocular rivalry experiment this issue cannot be completely excluded. However, given the resemblance between the results of the sphere stimulus and binocular rivalry, we are fairly certain we are probing a perceptual effect rather than a change in reporting criteria.

In the appendix of this chapter we provide numerical simulations of our model that demonstrate a strikingly close resemblance between the experimental effects of voluntary control and simulations in which we implement voluntary control as a gain modulation in the two individual pre-processing stages (conform van Ee et al., 2006)(compare Figures 3-4C&D with 3-A1D and 3-4E&F with 3-A1E). This resemblance suggests that voluntary control acts as a top-down attention-driven gain modulator in early visual processing stages. The fact that the perception of a bistable sphere and binocular rivalry gratings are similarly influenced by voluntary control, together with the suggestion that binocular rivalry is resolved at a lower level of neural processing, further supports the notion that voluntary control effects perception at very early levels of neural processing.

Attention and voluntary control both appear to influence bistable perception via early neural gain modulations. Voluntary control could thus be interpreted as the application of a non-specific attentional strategy to affect perception (it is the *reason for* rather than the *type of* attention). This would explain the resemblance between our voluntary control findings and experiments in which onset rivalry is influenced by attentional states (Chong & Blake, 2006). In their study Chong and Blake demonstrated that feature-based attention influences binocular rivalry. They asked their subjects to attend to some stimulus feature preceding rivalry in the assumption that this feature-based attention would still be present at the onset of rivalry and influence perception accordingly. In our current study, we investigate the underlying



mechanism of voluntary control without ever explicitly instructed our subjects to attend to any stimulus feature. The only instruction they received was to attempt to voluntarily control perception. It was never specified how they should do this, leaving subjects free to exploit any appropriate cognitive strategy to achieve the instructed effect on perception. The resemblance between our findings and those of Chong and Blake (2006), together with existing literature on visual attention and our own novel computational efforts, strongly suggests that our subjects used some attentional strategy to exert voluntary control and influence perception. Differences in attentional strategy between individual subjects may account for differences in the extent to which they can exert voluntary control. Although in repeated tests observers produce highly replicable data with only small variation in drift and sequential dependence (Suzuki & Grabowecky, 2007; van Ee, 2005), they also benefit from initial training to become successful in exerting voluntary control. In this view, they may simply use this initial training to develop an appropriate attentional strategy. It remains an open question what attentional strategy our subjects used to be successful in exerting voluntary control. Previous reports provided evidence that voluntary control over the sphere does not interact with perception at the level of individual elements but at a higher level of surface representations (Brouwer & van Ee, 2006), dovetailing nicely with recent single cell recordings showing that attention can select surface representations to modulate activity as early as cortical area MT (Wannig et al., 2007).

The hallmark of bistable stimuli is that there is evidence in the stimulus for mutually exclusive perceptual interpretations. Voluntary control functionally appears to increase the relative value of the evidence for one of the possible perceptual interpretations thereby biasing the system towards a specific perceptual decision. The neural mechanism by which voluntary control achieves this gain-modulation could involve feedback connections from prefrontal and posterior parietal areas back to earlier areas of visual processing (Buschman & Miller, 2007; Saalmann et al., 2007), but is a topic that has to be addressed in future research.

### 3.11 Acknowledgements

The authors would like to thank Jan Brascamp and Tomas Knapen for their help with the visual stimuli and Loes van Dam for her assistance in the eye movements experiment. Martin Lankheet and Jan Brascamp provided valuable comments on earlier versions of this manuscript. This work was supported by two VIDI grants from the Netherlands Organization for Scientific Research (NWO) awarded to RvW and RvE as well as a High Potential grant from Utrecht University awarded jointly to RvE and RvW.



### 3.12 Appendix

The fact that perceptual stabilization only occurs at longer inter-stimulus intervals was predicted and qualitatively shown by a single-stage, low-level neural model for visual rivalry (Noest et al., 2007). Here we present a data-driven analysis of the influence of stimulus timing on perceptual rivalry revealing a relative insensitivity to stimulus presentation duration that is significantly stronger than was predicted in the original single-stage model. We demonstrated that an extension of the model easily accounts for this finding if a realistic stage of adaptation prior to the rivalry resolution is included. Furthermore, our experiments with voluntary control and variable stimulus timing confirm novel model predictions for input imbalances suggesting voluntary control to act as a gain control mechanism preceding rivalry resolution. Figure 3-A1A of this appendix is a schematic representation of the model consisting of an adaptive pre-processing stage, a gain factor under top-down control and a rivalry-resolving stage (as published by Noest et al., 2007). In equations the entire model can be described by:

#### Pre-processing stage

$$\tau \partial_t \xi_i = x_i - (1 + a_i) \xi_i \quad (\text{Eq. 3-A1})$$

$$\begin{aligned} \partial_t a_i &= -a_i + \alpha S[\xi_i]; \\ i, j \in \{1, 2\}, i \neq j \end{aligned} \quad (\text{Eq. 3-A2})$$

#### Gain mechanism

$$X_i = g_i S[\xi_i] \quad (\text{Eq. 3-A3})$$

#### Rivalry-resolving stage

$$\tau \partial_t h_i = X_i - \beta - (1 + A_i) h_i - \gamma S[h_j + \beta] \quad (\text{Eq. 3-A4})$$

$$\begin{aligned} \partial_t A_i &= -A_i + \alpha S[h_i + \beta]; \\ i, j \in \{1, 2\}, i \neq j \end{aligned} \quad (\text{Eq. 3-A5})$$

The rivalry resolving stage of the model describes two neural pools that are each coding for a rivaling perceptual interpretation. They adapt, inhibit each other through cross-inhibition, and their adaptation levels interact with a neural baseline. The basic model consists of fast ‘local field’ dynamics (Equation 3-A4) and a slow shunting-type adaptation component (‘leaky integrator’, Equation 3-A5). The fast dynamics are determined by the local field activity of one of two opposing percept-coding pool of neurons ( $h$ ), which is translated into a spike-rate via a sigmoid function ( $S$ ) and depends on the neural input to the system ( $X$ ); the adaptation dynamics ( $A$ ); the amount of cross-inhibition ( $\gamma$ ) and a recently introduced term  $\beta$ , that can



physiologically be interpreted as an intraneuronal baseline interacting with the adaptation levels (for details, see Noest et al., 2007). The addition of the  $\beta$ -term to the well-established adaptation- and cross-inhibition-terms is the only real novelty compared to classic models of visual rivalry (Blake & Logothetis, 2002), but it has been shown to be crucial for understanding the underlying mechanism of percept-choice repetitions (Noest et al., 2007). The intraneuronal baseline ( $\beta$ ) effectively creates an elevated baseline activity and a reduced effective membrane time-constant of the adapted neural population. This results in a small head start for the more adapted neural population at the next stimulus onset that may overcome the effects of adaptation causing the more adapted neural population to become dominant and a perceptual repetition to occur. The intraneuronal baseline thus determines at which combination of adaptation levels perceptual repetitions will occur (see Figure 3-1C in the main text).

Since visual competition is highly unlikely to be resolved at the very first stage of neural processing, reality forces us to consider neural processing preceding the rivalry resolution. We implement these stages in their simplest possible form. Neural stages preceding rivalry resolution are assumed to have the same type of fast local field and adaptation dynamics as the rivalry resolving stage (Equations 3-A1 and 3-A2) whilst lacking cross-inhibition. The interaction with a neural baseline ( $\beta$ ) is functionally irrelevant here because the two pools do not interact (Noest et al., 2007) and it is therefore left out of the equations.

Figure 3-A1B demonstrates how, at each stimulus onset, the flow field of the fast local field dynamics (grey arrows) gives rise to trajectories (black lines) leading to one of two equilibrium points that represent the opposing percepts (P1, P2). The separatrix (dashed lines) at the intersection of the null-clines (red paired with either blue or green line) of the local field dynamics groups trajectories that are destined for either equilibrium. At stimulus onset, the starting point of a trajectory with respect to the separatrix thus determines which percept will win the rivalry. Adaptation asymmetries between the opposing neural populations scale the corresponding null-cline (green to blue solid line) and shift the separatrix (purple arrow; green to blue dashed line). If the separatrix-shift passes the starting point of a trajectory, this will change the corresponding endpoint to the less adapted percept, causing perceptual alternations. The inset demonstrates that the presence of an intraneuronal baseline ( $\beta$ ) creates a fixed offset in the starting points of the trajectories giving them a head start towards the more adapted percept. If this offset is smaller than the adaptation-driven separatrix-shift the system will end up with perceptual alternations ( $\beta_1$ ; the green solid line is the trajectory for the original green dashed separatrix, the blue solid line is the trajectory for the shifted blue dashed separatrix), but if the offset is larger than the separatrix-shift, the endpoints of the trajectories remain unchanged and perceptual stabilization occurs ( $\beta_2$ ; the black line is the new trajectory corresponding to the shifted blue dashed separatrix). With a fixed size for  $\beta$ , it is the adaptation-driven separatrix-shift that determines whether perceptual decisions alternate or repeat on subsequent presentations (see also Figure 3-1C for the relation between  $\beta$  and the adaptation levels).

Figure 3-A1C shows the result of a numerical simulation with our two-stage model and demonstrates that adaptation in pre-processing stages causes a steep border between areas of percept-choice repetitions and alternations (compare with Fig. 3C in Noest et al., 2007) confirming our experimental findings. The addition of noise to the system would make the model more realistic but blur the mentioned boundary, without influencing its shape (see Fig. A4B in Noest et al., 2007) and is therefore left out of this simulation. Apart from the fact



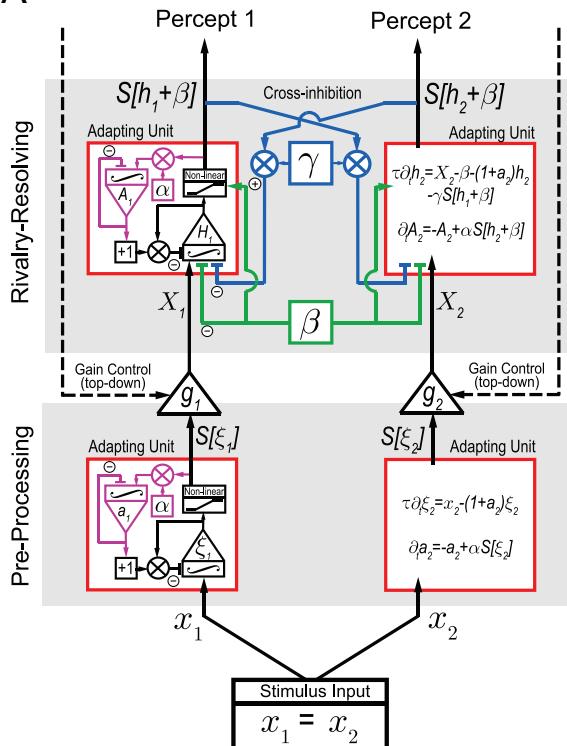
that existing empiric evidence dictates us to account for these pre-processing stages, no other simple parameter manipulations within the rivalry stage can produce the observed effect of stimulus timing.

Figures 3-A1D and 3-A1E, show the results of numerical model simulations for voluntary control. Here we use a 2% gain imbalance to model an active preference for one of two perceptual interpretations at stimulus onset and realistic noise (4% variance of the mean) is added to the adaptation levels to obtain smooth curves. As a result of an increase in the number of repetitions, the error on the data points also increases with increasing  $T_{\text{off}}$ . Not only does existing empirical evidence favour the implementation of early gain modulations, manipulations of other model-parameters, most importantly, those for the cross-inhibition and neural baseline, fail to yield results similar to our experimental data. An earlier suggestion that top-down voluntary control may be mediated through modulations of the  $\beta$ -parameter (Noest et al., 2007) are inconsistent with the vertical curve-shifts we demonstrate in our second experiment. Changing  $\beta$  would only affect the effect in the horizontal ( $T_{\text{off}}$ ) direction. The one notable difference between the simulation and our data is due the small bias for rightward percepts that is present in the data, but not in the simulation. If such a bias would be added to the simulation the simulation would match the data even more.

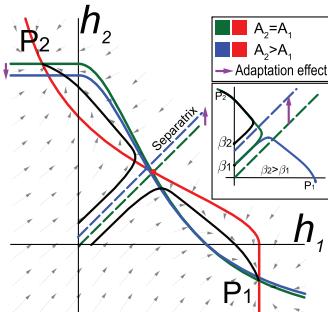
Altogether, we have presented a relatively simple yet physiologically plausible, low-level neural model that is sufficient to explain perceptual alternations, perceptual repetitions (stabilization), the influence of stimulus timing and the mechanism underlying top-down voluntary control for perceptual decisions at the onset of visual rivalry.



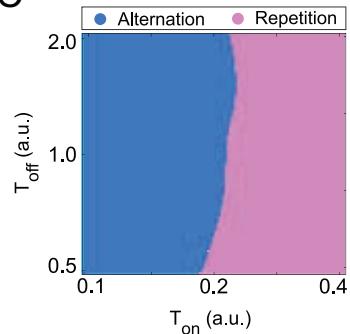
A



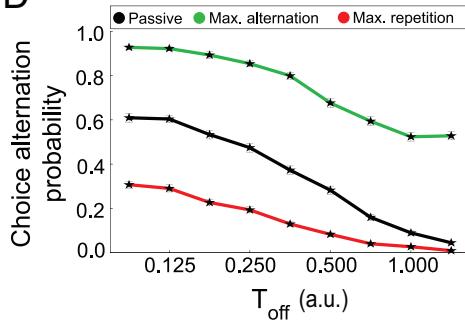
B



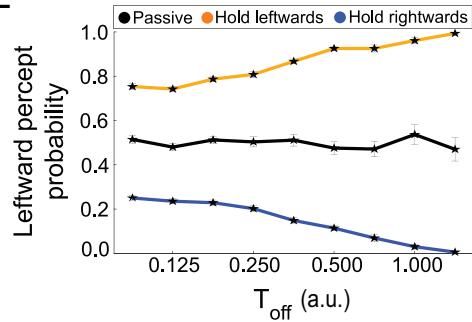
C



D



E





◀ **FIGURE 3-A1.** **A)** Schematic representation of our complete percept-choice model for visual rivalry and its voluntary control. The model consists of two stages; a pre-processing stage feeds into a rivalry-resolving stage with a gain factor that is under top-down control. The dynamics of the model are determined by adaptation (purple lines), cross-inhibition (blue lines) and a neural baseline (green lines). Both stages consist of adaptive neural pools coding for opposing percepts (red rectangles). On the left side of the scheme, the dynamics are displayed schematically, on the right side the corresponding equations are depicted. **B)** The dynamics of the low-level neural model for perceptual choices in visual competition as defined by differential equations 3-A4 and 3-A5. Depicted on the axes are the activity levels of opposing neural populations (local fields,  $h_i$ ). When the adaptation levels of both are equal ( $A_2 = A_1$ ), the red and green null-clines ( $d_t h_i = 0$ ) determine the flow fields of the fast local field dynamics creating the saddle-point and corresponding separatrix at their cross-section. An imbalance in adaptation states with  $A_2 > A_1$  scales one of the null-clines (for  $h_2$ ) in the  $h_2$ -direction (green line) and causes the intersection and attached separatrix to move slightly up- and leftwards. The intraneuronal baseline ( $\beta$  in the model) creates an offset of the starting points of the trajectories. The inset shows that trajectories (solid lines) for which the offset ( $\beta_s$ ) is smaller than the separatrix-shift (purple arrow) the trajectories are now on the other side of the separatrix (green dashed line to blue dashed line). This changes their endpoints to the less adapted percept, resulting in perceptual alternations (green to blue trajectory). If the adaptation-driven separatrix-shift is however smaller than the offset ( $\beta_s$ ) the trajectory endpoint remains unchanged (green to black trajectory), the more adapted percept prevails and perceptual stabilization occurs. Thus, with a fixed  $\beta$ , it is the size of the adaptation-driven separatrix-shift that determines whether perceptual decisions alternate or repeat on subsequent presentations. **C)** A simulation for the effect of pre-stage adaptation demonstrates a near-vertical border between alternation and repetition areas. **D)** Simulations for instructions “maximize alternations” (green), “maximize repetitions” (red) and the passive condition (black). **E)** Simulations for instructions “hold leftwards” (orange), “hold rightwards” (blue) and the passive condition (black). Voluntary control was modeled in the preprocessing stage as a 2 percent increase of the gain for the favored percept.  $T_{on}$  and  $T_{off}$  in are given in units of the adaptation time constant. Note the close resemblance between panels D and E and the data in Figures 3-4C and 3-4D. Simulations were performed with  $x_i \in \{0,3g\}$ ,  $\alpha = 5$ ;  $\beta = 4.4/(3\alpha)$ ;  $\gamma = 10/3$ ;  $g_1 = g_2 = 1.5$ ;  $\tau = 50$ . For panel D and E,  $T_{on}$  was set to 1.0, noise on the adaptation levels was added as 4% variance of the mean, and top-down gain control was modeled as a 2% gain increase for the favoured percept.



# Chapter 4

**Intermittent stimulus presentation stabilizes  
neuronal responses in macaque area MT**

*Let us assume that the persistence or repetition of a reverberatory activity (or “trace”) tends to induce lasting cellular changes that add to its stability.*

**Donald Hebb**

**Submitted as**

Klink, P.C., Oleksiak, A., Lankheet, M.J.M., & van Wezel, R.J.A. (Submitted). Intermittent stimulus presentation stabilizes neuronal responses in macaque area MT.



*This is the “blood, sweat and tears” chapter of this thesis. Murphy’s law must have been written by an electrophysiologist, because it seems that everything that can possibly go wrong with an electrophysiology set-up will go wrong over the course of a PhD project. Armed with rolls of aluminum foil and duct tape we eventually managed to create enough episodes with working equipment and a decent enough signal to be able to gather the data on which this chapter is based. I am really happy that we decided to add some extra hardware to the set-up which allowed us to save the raw neural traces for offline local field potential analysis. We have worked on gathering this data for several years, often in parallel with some of the other projects described in this thesis. One of the most important things I have learned in the process is that it helps to keep your research questions and stimuli as simple as possible. As tempting as it may seem to directly aim for explanations of higher cognitive functions, this is going to be very hard with a single piece of tungsten wire and observers that are not always as cooperative as you might want them to be. There were many times when I thought that this project was either going nowhere, or that I was simply incapable of figuring out the proper way to analyze the data. Now that we actually ended up with a sensible story, I can look back on the many hours of trying to ‘hear a neuron in the noise’ and admit that it was definitely all worth the effort.*

---



## 4.1 Abstract

Repeated stimulation impacts neuronal responses. Here we show how response characteristics of sensory neurons in macaque visual cortex are influenced by the duration of the interruptions during intermittent stimulus presentation. Spikes and local field potentials in motion-sensitive area MT both demonstrated lower response variabilities when interruption durations systematically increased from 250 to 2,000 ms. Activity fluctuations between subsequent trials and Fano Factors over full response sequences were both lower with longer interruptions, while spike-timing patterns became more regular. These changes were independent of general adaptation effects and accompanied by increases of spectral power in the high Gamma range of the local field potential. This suggests that the reduced response variability results from an increased involvement of the local cortical network. While neuronal response stabilization may be a general effect of repeated sensory stimulation, it potentially reveals itself in the perceptual stabilization of ambiguous stimuli that occurs under similar stimulation dynamics.

## 4.2 Introduction

Repetition is important for perceptual learning and memory, but it remains largely unclear how sensory neurons alter their response characteristics when they are exposed to repeated stimulation. Studies on the effects of repetition in the brain predominantly address the reduction of response magnitude known as adaptation (Grill-Spector et al., 2006). However, behavioral and computational studies on the conscious perception of intermittently presented ambiguous visual stimuli suggest that repeated stimulus presentations can have much more elaborate effects on neuronal response characteristics than might be initially expected from straightforward neuronal fatigue (e.g., Leopold et al., 2002; Noest et al., 2007). Ambiguous visual stimuli are patterns that contain equal sensory evidence for multiple, mutually exclusive perceptual interpretations. When these patterns are viewed continuously, perception lapses into fluctuations between the alternative stimulus interpretations. Since these perceptual fluctuations occur while the retinal images remain constant, the dynamics of correlated neural activity patterns can be regarded as essential determinants of conscious visual experience (Blake & Logothetis, 2002; Crick & Koch, 2003).

Temporarily removing an ambiguous pattern from view severely reduces the number of perceptual fluctuations and stabilizes perception into one of the possible stimulus interpretations (Klink et al., 2008a; Kornmeier et al., 2007; Leopold et al., 2002; Orbach et al., 1966; Pearson & Brascamp, 2008). This phenomenon is dubbed perceptual stabilization or perceptual memory, and it has been demonstrated with various distinct ambiguous patterns (Klink et al., 2008a; Pearson & Brascamp, 2008) even if several different patterns are presented in interleaved sequences (Maier et al., 2003). Over the last few years, insightful results about the nature of this perceptual memory have been obtained with behavioral studies (Brascamp et al., 2009; Brascamp et al., 2008; Carter & Cavanagh, 2007; Kanai et al., 2007a; Klink et al., 2008a; Knapen et al., 2009; Pastukhov & Braun, 2008), neural imaging (Raemaekers et al., 2008), brain stimulation (Brascamp et al., 2010) and computational approaches (Gigante et al., 2009; Noest et al., 2007; Wilson, 2007). While it currently remains unclear how the activity patterns of single neurons are altered when perception stabilizes, it is conceivable that



response characteristics are influenced in complex manners that go beyond simple fatigue.

Behavioral and computational approaches have demonstrated that the occurrence and depth of perceptual stabilization crucially depend on the duration a stimulus is removed from view (Klink et al., 2008a). Compared to the perceptual dynamics during continuous viewing, long interruptions ( $> 1.0$  s) stabilize perceptual sequences to a single interpretation, but short interruptions ( $< 0.5$  s) de-stabilize perception and cause an increase in the frequency of perceptual alternations. This correlation between removal duration and the depth of perceptual stabilization presents a window of opportunity for neurophysiological efforts aiming to reveal the correlates of perceptual stabilization at the level of a single neuron.

Here we set out to investigate how neurons in area MT of the macaque monkey change their response patterns depending on the duration of the blank period in a sequence of intermittently presented visual stimuli. We recorded spike patterns and local field potentials in two awake behaving monkeys while they were presented with sequences of ambiguously rotating cylinders and two other types of motion stimuli that were all periodically removed from view. Stimulus removal durations covered the full previously demonstrated spectrum of highly unstable, alternating percepts to well-stabilized sequences in which the same perceptual interpretation is dominant for minutes in humans (Klink et al., 2008a; Leopold et al., 2002). To avoid any confounding effects of attention (Klink et al., 2008a; Kornmeier et al., 2009; Mitchell et al., 2004), flash suppression (Sengpiel et al., 1995; Sheinberg & Logothetis, 1997; Wolfe, 1984) or flash facilitation (Brascamp et al., 2007), we did not include disambiguated stimuli in the sequences or ask the monkeys to report their percepts. While this approach has the disadvantage of not knowing which percept corresponds to which neural trace, it allows a much cleaner investigation of the neuronal consequences of long sequences of stimulus presentations with different removal period durations.

Are there any *a priori* expectations on the influence of blank durations in stimulus sequences on neuronal response patterns, apart from simple adaptation? If MT neurons take part in encoding the subjective percepts of ambiguous structure-from-motion (SFM) stimuli (Bradley et al., 1998; Parker et al., 2002), and periodically removing such stimuli from view stabilizes perception into one specific interpretation (Klink et al., 2008a; Leopold et al., 2002), then response patterns might be expected to stabilize under these circumstances as well. Few studies have investigated the effects of sequential visual stimulation on neuronal responses beyond classic neuronal adaptation (Kohn, 2007; Mayo & Sommer, 2008). Repeated exposure to movies of natural scenes increases cortical response reliability in primary visual cortex, but such an effect was not found with more abstract flashed bar stimuli (Yao et al., 2007). A recent analysis of response variability dynamics in multiple cortical areas, including MT, demonstrated that stimulus onset itself consistently reduces across-trial variability, implying that cortical circuits become more stable when driven (Churchland et al., 2010). Response variability also decreases when stimuli are actively attended (Mitchell et al., 2007). Building upon these findings, we hypothesize that the depth of perceptual stabilization may be related to neuronal stabilization of cortical circuits that manifests itself in a decrease of across-trial neural variability during intermittent stimulus presentations.



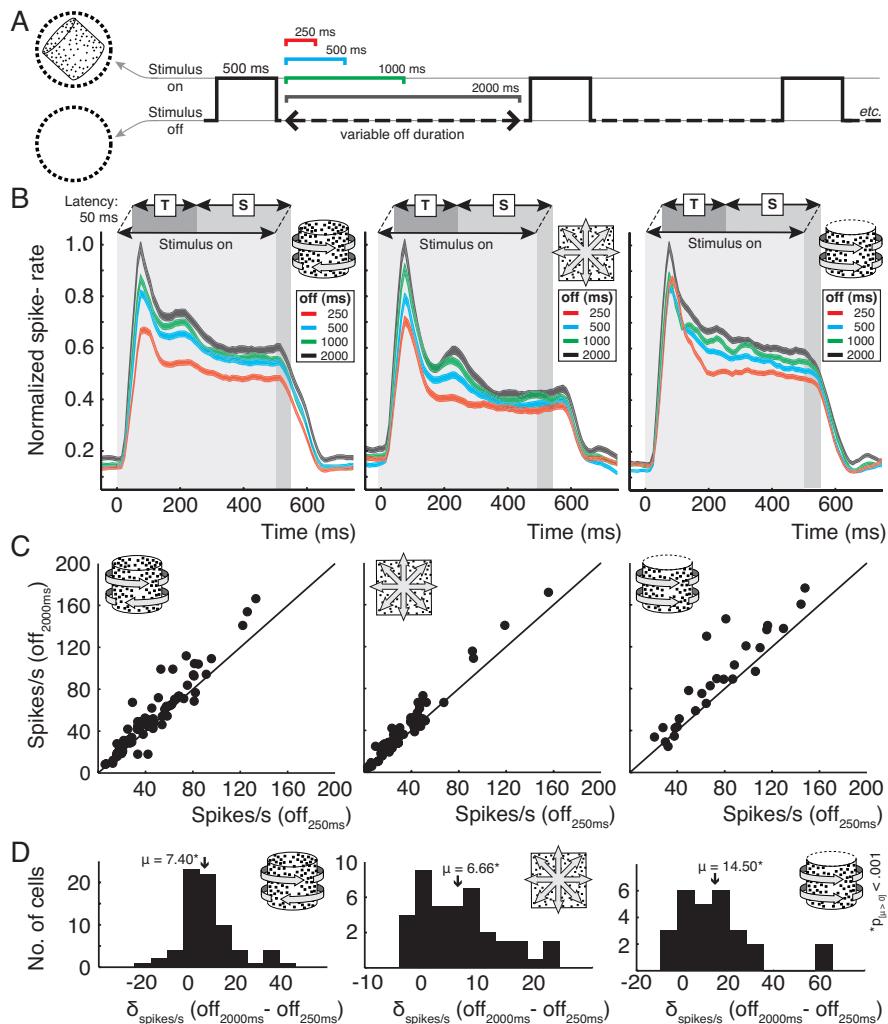
## 4.3 Results and discussion

To investigate the effects of intermittent stimulus presentation on neuronal response properties in area MT of the macaque monkey, we presented two monkeys with stimulus sequences in which the stimulus was repeatedly displayed for 500 ms (*on*) and removed from view for a fixed blank duration (*off*) that could be either 250, 500, 1000 or 2000ms (Figure 4-1A). This off duration was kept constant over a single sequence of approximately 80 stimulus repetitions. In correspondence with previous behavioral studies on perceptual stabilization, we used an ambiguously rotating cylinder stimulus for which there are two perceptual interpretations with opposite rotation directions (Klink et al., 2008a). To discern any general effect of repeated stimulus presentation from more specific effects related to the presence of a visual ambiguity, we included two additional control stimuli in our experimental protocol. These were 1) dynamic random noise patches that contained all motion direction in balanced proportions, and 2) unambiguous, ‘opaque’ cylinders (Freeman & Driver, 2006) that contained only one motion direction in the preferred direction of the recorded neuron and no perceptual ambiguity. Responses to the sequences of these stimuli were obtained from 94 single units in area MT of two macaque monkeys (46 from Monkey S and 48 from Monkey A).

### 4.3.1 Spike-rate & adaptation

When an MT neuron is continuously or repeatedly driven by the same type of input its responsiveness will decrease, a phenomenon known as neuronal adaptation (for a review, see Kohn, 2007; van Wezel & Britten, 2002). The adapted responsiveness will recover back to baseline values when stimulation stops, but this may take some time. During intermittent stimulus presentation, the length of the off period determines to what extent a neuron will recover from adaptation evoked by the previous stimulus presentation. Figure 4-1B displays the average peristimulus time histograms (PSTH) for the recorded neurons, split up for the three different stimuli (panels) and four off-durations (colors within a panel). It is immediately clear that the average response amplitude increases with increasing off-duration for all three stimuli (Spearman rank correlations of off-duration vs. average spike-rate during stimulus presentation over all trials within a stimulus type: all  $r > 0$ , all  $p < 0.001$ ). While this effect is clearly visible in both the transient phase of the response (defined as the first 200 ms after a 50 ms fixed latency, marked ‘T’ in Figure 4-1B) and the sustained phase (defined as the last 300 ms, marked ‘S’ in Figure 4-1B) for the ambiguous cylinder and dynamic random noise stimuli, it is less pronounced, yet still significant, in the transient phase for the opaque cylinders.

The difference in average response magnitude over trials is largest between the sequences with off-duration on the both ends on range of durations used. If we directly compare the average response with the shortest (250 ms) and longest (2000 ms) off-durations within single cells, we notice that for practically all recorded cells and for all stimuli, long off-durations yield higher spike-rates than short off-durations (Figures 4-1C & 4-1D)(Paired t-test,  $p < 0.001$ ). While this effect is exactly what one would expect from general adaption, it also presents a potential confound for additional analyses on the effects of off-durations on response variability. We account for these possible adaptation effects by consistently checking whether any influence of off-duration on response variability cannot be directly interpreted as an effect of response magnitude.



**FIGURE 4-1.** Experimental paradigm and averaged response patterns. **A)** Stimuli were placed in a neuron's receptive field (dotted circle) and intermittently presented with a fixed 'stimulus on' period of 500 ms and a variable 'stimulus off' period that was varied over blocks between 250, 500, 1000 and 2000 ms. **B)** Averaged peristimulus time histograms for different stimuli (panels) and off-durations (colors). Histograms were calculated in 10 ms bins aligned to stimulus onset and averaged over all trials with the same combination of stimulus and off-duration ( $n_{\text{amb},\text{cyl}} = 7662, 7244, 7001, 4741$ ;  $n_{\text{dyn,noi}} = 4712, 5033, 4557, 3644$ ;  $n_{\text{op,cyl}} = 2263, 2183, 1976, 1830$ , for the four off-durations in increasing order respectively). For plotting purposes, the histograms were normalized to the peak value at the longest off-duration of 2000 ms and smoothed with a 50 ms moving average window. Line thickness indicates the average value  $\pm$  S.E.M. The gray rectangular areas indicate the time range of stimulus presentation. The response, defined as the activity in a 500 ms window with a 50 ms latency relative to stimulus onset, was subdivided in a 200 ms initial transient response (marked 'T') and a remaining 300 ms sustained response (marked 'S'). **C)** Average spike-rates during stimulus presentations for the shortest vs. the longest off-duration split up by stimulus type (panels). Dots are individual cells. **D)** Distributions of differences in the average spike-rates during stimulus presentation with longest and shortest off-duration over all cells. Mean differences are significantly above zero for all stimuli.



### 4.3.2 Fano Factors

It has recently been shown that in a wide range of cortical areas, including area MT, stimulus onset reduces the response variability expressed in the Fano Factor (Churchland et al., 2010). The Fano Factor (*FF*) is calculated by dividing the spike count variance of a response to its mean ( $\sigma^2/\mu$ ). We calculated *FF*'s for our stimuli and off-durations both by directly dividing the variance to the mean in a 70 ms sliding windows moving in 10 ms increments, and by applying the 'mean-matched' method introduced by Churchland et al. (2010). Since these two methods resulted in qualitatively similar results, we only report the directly calculated *FF*'s that are based on more data-points.

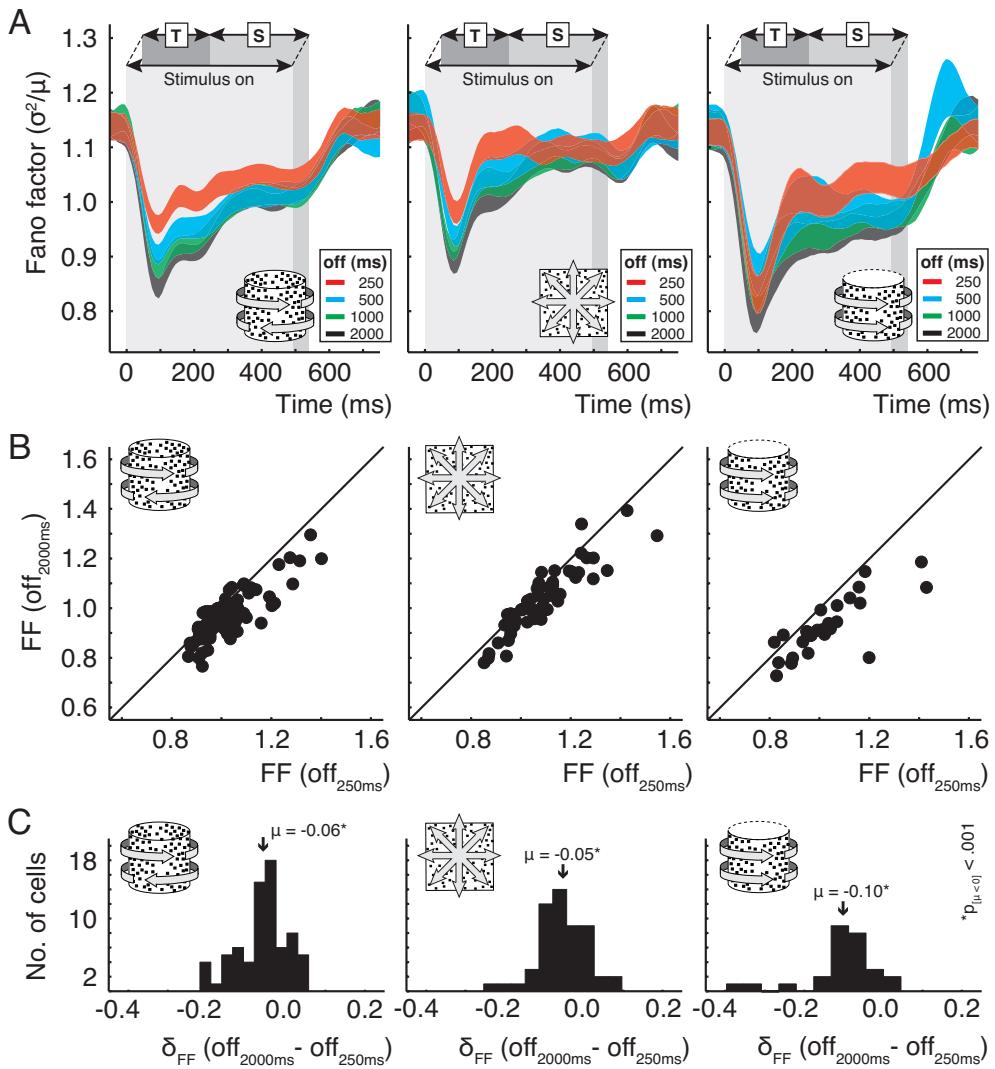
For all three stimulus types, stimulus onset clearly reduced the *FF*, but the extent to which the *FF*'s decreased depended on the duration of the off-period (Figure 4-2A). For the ambiguous and opaque cylinders this notion held for both the transient and sustained phases (Spearman rank correlation, Ambiguous:  $r_{trans} < 0$ ,  $p_{trans} < 0.01$ ;  $r_{sust} < 0$ ,  $p_{sust} < 0.02$ ; Opaque:  $r_{trans} < 0$ ,  $p_{trans} < 0.05$ ;  $r_{sust} < 0$ ,  $p_{sust} < 0.02$ ), while for the dynamic noise patches only the transient phase showed a significant off-duration dependency ( $r_{trans} < 0$ ,  $p_{trans} < 0.03$ ;  $r_{sust} < 0$ ,  $p_{sust} = 0.10$ ). A comparison of the average *FF* during stimulus presentation for the longest and shortest off-durations (Figures 4-2B & 4-2C) reveals that for virtually all cells the *FF* is lower for the long off-duration in all three stimuli (Paired t-test,  $p < 0.001$ ).

It has been argued that the drop in *FF* at stimulus onset reflects the stabilization of the cortical networks that are driven (Churchland et al., 2010). Furthermore, it has been demonstrated that the decline in *FF* when a stimulus enters the receptive field of a V4 neuron is stronger when this stimulus is attended than when it is not attended (Mitchell et al., 2007). This suggests that attention may cause an additional cortical stabilization that facilitates a maximized neuronal signal-to-noise ratio (Mitchell et al., 2007). Our results demonstrate that such a stronger decrease in neuronal response variability similarly exists for intermittent stimulus presentations without explicit attentional demands, but with increasingly longer off-durations. It is tempting to envision the perceptual stabilization known to occur with intermittently presented ambiguous stimuli at the behavioral level as a manifestation of this generic neuronal response stabilization that occurs under the same temporal stimulation conditions.

### 4.3.3 Local Activity Contrast

Whereas Fano Factors provide a nice estimate of response variability over the full sequence of stimulus presentations and the associated responses, we would also like to obtain a more temporally local measure of trial-to-trial response variability. To this end we calculated the 'Local Activity Contrast' (LAC) that we defined as the absolute spike count difference between subsequent trials divided by the mean activity of these two trials (Equation 4-1). In this equation,  $k_n$  denotes the spike count in the first trial of the pair and  $k_{n+1}$  the spike count during the second trial. The LAC can theoretically take values between zero (when there is an equal number of spikes on both trials) and two (when there are no spikes on one of the two trials).

$$LAC = \frac{|k_{n+1} - k_n|}{(k_{n+1} + k_n)/2} \quad (\text{Eq. 4-1})$$



**FIGURE 4-2.** Fano factors decrease differentially at stimulus onset. **A)** Averaged Fano factors for different stimuli (panels) and off-durations (colors). Fano factors were calculated in 10 ms bins aligned to stimulus onset and averaged over cells ( $n_{\text{amb\_cyl}} = 82, 80, 79, 67$ ;  $n_{\text{dyn\_noi}} = 56, 61, 58, 55$ ;  $n_{\text{op\_cyl}} = 29, 30, 29, 29$ , for the four off-durations in increasing order respectively). For plotting purposes, the data was smoothed using a 50 ms moving average window. Line thickness indicates the average value  $\pm$  S.E.M. The gray rectangular area marks the moment of stimulus presentation. The transient and sustained phases of the response are marked with 'T' and 'S' respectively. **B)** Average Fano factors during stimulus presentations for the shortest vs. the longest off-duration split up by stimulus type (panels). Dots are individual cells ( $n_{\text{amb\_cyl}} = 67$ ,  $n_{\text{dyn\_noi}} = 55$ ,  $n_{\text{op\_cyl}} = 29$ ). **C)** Distributions of differences in the average Fano factor during stimulus presentation with longest and shortest off-duration over all cells. Mean differences are significantly below zero for all stimuli.



Figure 4-3A demonstrates that the LAC value significantly depends on the duration of the off-interval for all stimuli, with a lower LAC when off periods are longer (Spearman rank correlations; all  $r < 0$ , all  $p < 0.05$ ). The LAC values are relatively low in general, indicating that subsequent trials do not differ in their spike count all that much. If the shortest and longest off-durations are contrasted (Figures 4-3B & 4-3C), we can clearly see that for practically all cells the longer off- duration is correlated with the lower LAC (Paired t-test, all  $p < 0.01$ ).

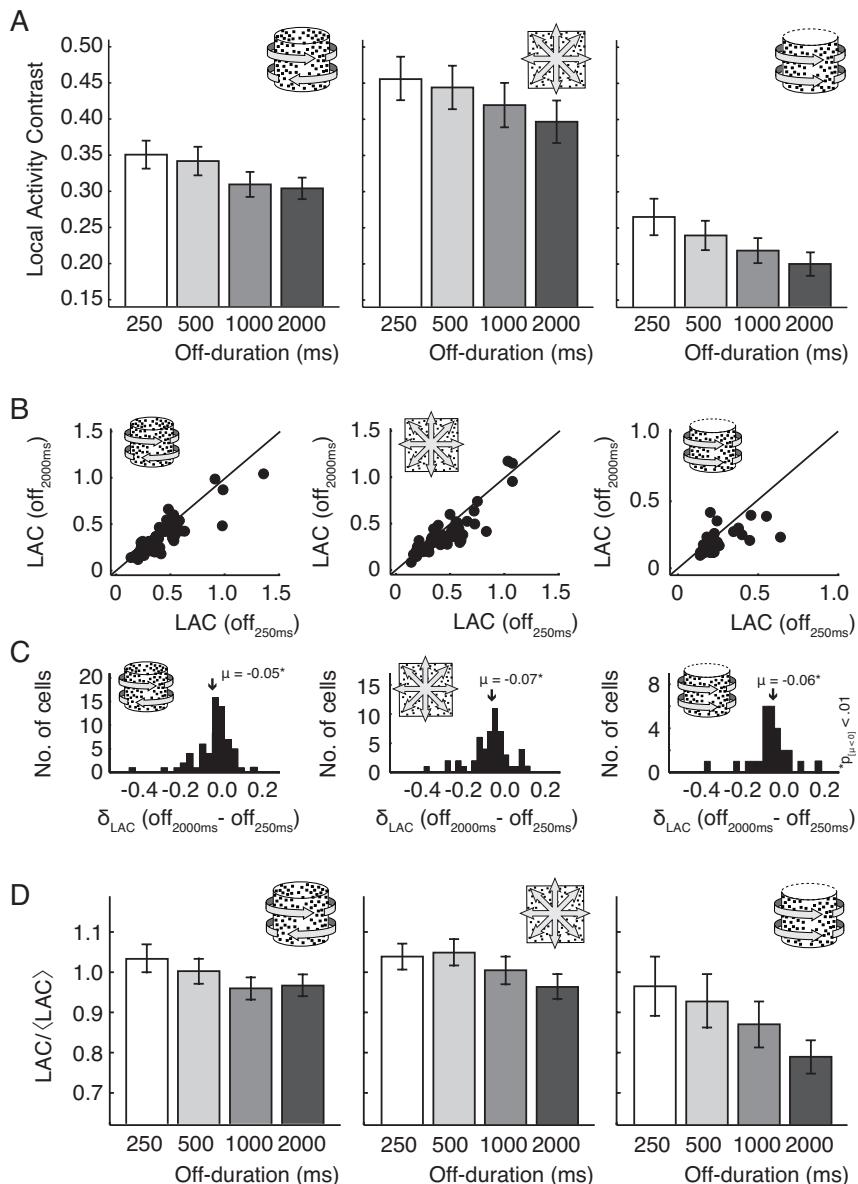
One concern with this definition of local activity contrast is that the LAC's are strongly correlated with the average spike count on the two trials (Spearman rank correlations; all  $r < 0$ , all  $p < 0.0001$ ). Since we already demonstrated that the duration of the off-interval significantly influences the average spike count, this may present a serious confound in our indication of trial-to-trial variability. However, if we consider the average spike count as a continuous grouping variable, an additional analysis of variance reveals that there is still a significant influence of the off-period on the LAC when average activity is taken into account (ANOVA,  $P_{amb\_cyl} < 0.001$ ,  $P_{dyn\_noi} < 0.05$ ,  $P_{opa\_cyl} < 0.001$ ). The relation between the average spike count and the LAC can be approximated if we assume that our neuron's spike counts ( $k$ ) resemble independent samples from the same Poisson distribution. The Poisson-predicted LAC values ( $\langle LAC \rangle$ ), can then be derived to follow Equation 4-2 (for details on the derivation see Methods & Materials).

$$\langle LAC \rangle = \frac{2}{\sqrt{\pi(k_1 + k_2)/2}} \quad (\text{Eq. 4-2})$$

With this predicted LAC, we can calculate the ratio of the measured and predicted LAC for every pair of trials individually ( $LAC/\langle LAC \rangle$ ). A value smaller than 1.0 would indicate less variability in trial-to-trial spike counts than predicted for a Poisson process, whereas a value larger than 1.0 would imply more variability. Consistent with our Fano Factor findings, Figure 4-3D demonstrates that  $LAC/\langle LAC \rangle$  ratios decrease with increasing off-duration indicating that indeed trial-to-trial variability is lower when stimulus sequences are presented with longer intermittent blank intervals. For ambiguous cylinders and dynamic noise, the  $LAC/\langle LAC \rangle$  ratio is above 1.0 for the shortest off interval and below 1.0 for the longest off periods. The resolution of perceptual ambiguity in the former and broad motion direction content of the latter could be responsible for this higher initial trial-to-trial variability that appears to stabilize when blank durations increase. The opaque cylinders that contain only one motion direction and no inherent ambiguities show an overall lower  $LAC/\langle LAC \rangle$  ratio, but the stabilizing effect of increasing off-duration is also clearly present there.

#### 4.3.4 Spike timing precision

Both the Fano Factor analysis and the Local Activity Contrast analysis consider merely the number of spikes in a given time interval, not their exact moments of incidence. Decreases in neuronal response variability may however also be reflected in an increasing regularity of spike times in spike trains evoked by subsequent stimulations (Maimon & Assad, 2009). To investigate this possibility, we calculated a Spike Time Deviation Index (STDI) for every



**FIGURE 4-3.** Local Activity Contrast decreases with off-duration. **A)** The Local Activity Contrast, calculated as the absolute spike count difference between subsequent stimulus presentations divided by the average spike count on these subsequent trials, averaged over the same cells as presented in Figure 4-2 and plotted against the off-duration for the three different stimuli. Error bars are S.E.M. **B)** LAC-values on the shortest off-duration of 250 ms contrasted with the longest off-duration of 2000 ms for individual cells (dots) and different stimuli (panels) ( $n_{amb\_cyl} = 67$ ,  $n_{dyn\_noi} = 55$ ,  $n_{op\_cyl} = 29$ ). **C)** The distributions of LAC differences between trials with the longest and shortest off-duration within a single cell. **D)** The ratio of the measured LAC and the Poisson-predicted LAC values ( $\langle LAC \rangle$ ) plotted for the different stimuli and off-durations.



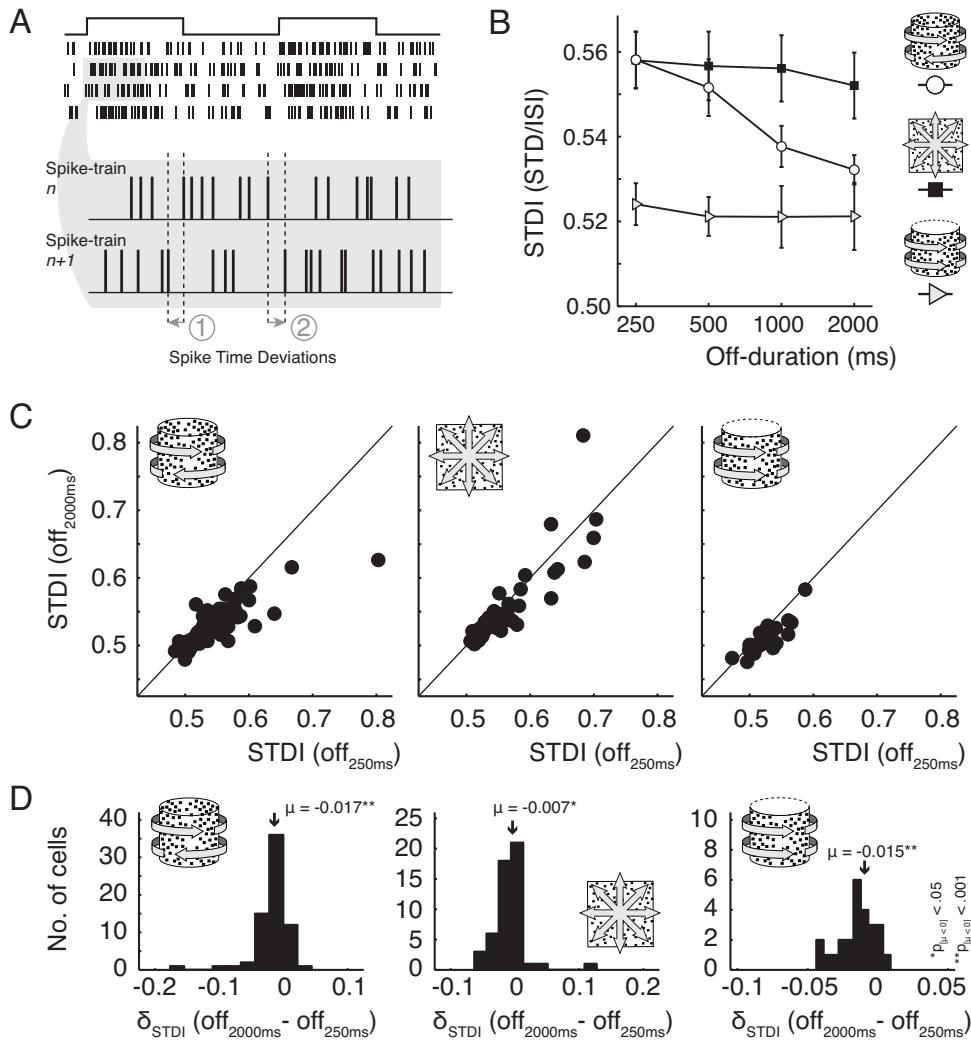
spike that occurred during stimulus presentation (Figure 4-4A). To this end, all spiketimes were first expressed relative to stimulus onset. We then determined the Spike Time Deviation (STD) as the absolute temporal distance in milliseconds between each spike and its closest temporal neighbor in the subsequent spiketrain. These closest neighbors could either occur earlier or later than the reference spike relative to stimulus onset. To account for the fact that a higher spike density will automatically cause lower STD values, we divided the STD's by the average interspike interval (ISI) of the two subsequent spiketrains to obtain our STDI. This normalization step was successful in preventing a correlation between average activities and STDI's (Spearman rank correlation, all  $p > 0.28$ , except for opaque cylinders with a blank duration of 250 ms,  $p = 0.03$ ).

Figure 4-4B plots the STDI for the different stimuli and off-durations, averaged over all stimulus presentations. Correlating the off-durations with the STDI revealed a significant negative correlation for the ambiguous cylinder stimuli (Spearman rank correlation,  $r < 0$ ,  $p < 0.001$ ), but not for dynamic noise patches ( $p = 0.60$ ) or opaque cylinders ( $p = 0.76$ ). Interestingly, the average STDI's for the unambiguous opaque cylinder are a lot lower than for the dynamic random noise patches, while STDI's for ambiguous cylinders are similar to those of dynamic noise at short blank interval but more similar to those of opaque cylinders at long blank periods. This transition may be related to the fact that intermittently presented ambiguous stimuli are perceptually unstable at short off-periods (high perceptual alternation probability) but become increasingly stable with longer stimulus interruptions (high perceptual repetition probability).

When the average STDI's of the shortest and longest used off-periods are contrasted for individual cells (Figure 4-4C & 4-4D), the STDI for the longest off-period is lower in most cells for all three stimuli, an effect that is significant for the population of recorded neurons (Paired t-tests,  $p_{\text{amb\_cyl}} < 0.001$ ,  $p_{\text{dyn\_noi}} < 0.05$ ,  $p_{\text{opa\_cyl}} < 0.001$ ).

#### 4.3.5 Local Field Potentials

All the spike-based analyses indicated that neuronal responses are more stable when intermittently presented stimuli are separated by longer off-intervals. However, the mechanisms by which this stabilization is established remain unclear. A straightforward hypothesis would be that the neurons in the local cortical network increase their processing coherence, thereby stabilizing the response patterns of the single units involved in the network (Churchland et al., 2010; Kelly et al., 2010; Sussillo & Abbott, 2009). Analysis of the local field potentials (LFP's) that we recorded simultaneously with our spiking data may provide some clues about the role of the local cortical network (Gawne, 2010; Katzner et al., 2009). To investigate the magnitude of activity in the LFP, we calculated the RMS value of the LFP amplitude (Liu & Newsome, 2006). This measure is analogous to the spike-count PSTH and was found to contain similar adaptation effects. Higher average RMS values with off-durations of 2000 ms than with off-durations of 250 ms were highly significant (Paired t-test,  $p < 0.001$ ) for ambiguous cylinders and dynamic noise patch and marginally significant for opaque cylinders ( $p = 0.06$ ). Moreover, an analysis comparable to the Fano Factor calculations for spike data revealed that the LFP magnitude variability also decreased as a function of increasing off-duration (not plotted but comparable to the spike results in Figure 4-2A). These findings imply that the neuronal stabilization effect observed in single neurons can also be observed at the level of



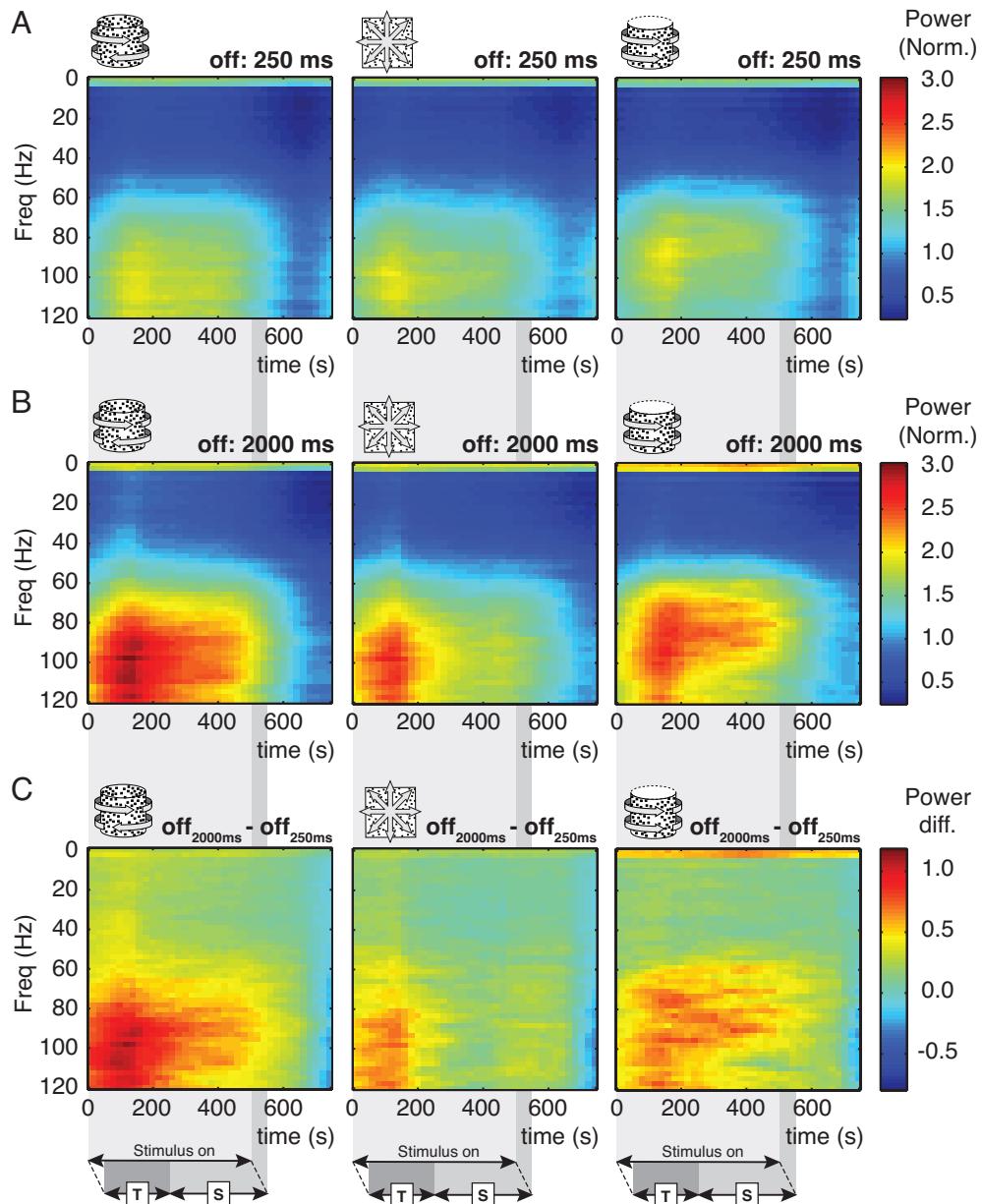
**FIGURE 4-4.** Regularity of spike timing. **A)** Spike time deviations are calculated by temporally aligning all response spiketrains to stimulus onset and determining the closest temporal distance between each spike on trial  $n$  and a spike on the subsequent trial  $n+1$ . Since spike time deviations can be both negative (example 1) and positive (example 2), absolute values are used in further analysis. **B)** Spike time deviation indices (STDI's), calculated as the average spike time deviation divided by the average interspike interval for each pair of subsequent presentation. STDI's are averaged over the same cells as presented in Figures 4-2 and 4-3 and plotted against off-duration for the three different stimuli. Error bars are S.E.M. **C)** STDI's with the shortest off-duration contrasted against those with the longest off-duration for individual cells ( $n_{\text{amb\_cyl}} = 67$ ,  $n_{\text{dyn\_noi}} = 55$ ,  $n_{\text{op\_cyl}} = 29$ ). **D)** Distributions over cells of the differences plotted in C.



the local cortical network. The local network could increase its response reliability through a tighter correlation between its constituent neurons' activity patterns. Since our data was recorded with single electrodes we cannot directly correlate the activity patterns of multiple simultaneously active neurons, but we can look at the frequency content of our LFP in search for clues of increased network coherence (Belitski et al., 2008; Fontanini & Katz, 2008; Fries et al., 2007; Logothetis, 2003; Singer & Gray, 1995).

A time-frequency analysis of the LFP data demonstrated the presence of stimulus-evoked power increases in both the Delta (0-4 Hz) and the high Gamma frequency ranges (60-120 Hz). For ambiguous cylinders, these power increases were highly significant in both frequency ranges for all off-durations and for both the transient and sustained phases of the response (cell-based paired t-tests on the average power before and during stimulus presentation: all  $p < 0.001$ ). For opaque cylinders and dynamic noise patches the power increases in the Delta range were modest but significant (all  $p < 0.05$ ), while the power increases in the high Gamma range were more profound (all  $p < 0.001$ ).

Figure 4-5 illustrates the response patterns for all three stimuli and both the shortest (250 ms, Figure 4-5A) and longest (2000 ms, Figure 4-5B) off-durations. The pattern that remains after subtracting the results obtained with the shortest off-duration from those with the longest off-duration implicates any effects of blank duration timing (Figure 4-5C). Such an effect is predominantly present in the high Gamma range. Figure 4-5C contains significant hotspots in this frequency band for all stimuli, albeit only in the transient phase of the response for dynamic noise patches ( $p < 0.01$ ). No stimulus-evoked off-effects were present in the Delta band for dynamic noise patches or opaque cylinders, but a small yet significant ( $p < 0.05$ ) effect could be observed for ambiguous cylinders. Whereas the slow Delta waves are classically observed during slow wave sleep, they are also reported in the context of sustained attention (Kirmizi-Alsan et al., 2006). However, since the stimuli in our experiment were presented with temporal frequencies in the Delta range, it is also conceivable that the observed effect in the Delta range is merely a stimulus presentation artifact. LFP activity in area MT in the higher Gamma range has been shown to reflect the motion direction and speed tuning properties that can also be observed in spiking data of single unit recordings (Khayat et al., 2010; Liu & Newsome, 2006). Furthermore, oscillations in the Gamma range are often considered to be related to the synchronization of activated neuronal ensembles (e.g. Fries et al., 2007). The increase of power in the Gamma range observed with longer off-durations is thus consistent with the idea of increased network stabilization. If the neurons in the local network would synchronize their firing patterns to oscillations in the LFP signal, this could also explain the increased regularity in precise spike timing as indicated by decreased spike time deviations. Both the off-dependent power differences and the off-dependent changes in STDI are largest for ambiguous cylinder stimuli, providing indirect support for the hypothesis of a network-driven increase in spiking regularity.



**FIGURE 4-5.** Local field potential power spectra. **A)** Cell-averaged power spectrum during presentation of three different stimuli (panels) with an intermittent off-duration of 250 ms. The power spectra for individual cells ( $n_{\text{amb\_cyl}} = 64$ ,  $n_{\text{dyn\_noi}} = 54$ ,  $n_{\text{op\_cyl}} = 29$ ) were normalized by the average power during stimulus presentation indicated by the gray rectangular area. **B)** Similar as A, but now for off-durations of 2000 ms ( $n_{\text{amb\_cyl}} = 62$ ,  $n_{\text{dyn\_noi}} = 54$ ,  $n_{\text{op\_cyl}} = 29$ ). **C)** The difference between the panels in A and B, indicating any off-duration dependent modulations of the local field potential power spectrum.



## 4.4 Conclusions

Intermittent presentation of motion stimuli has profound effects on neuronal response variability in area MT of the macaque monkey. If interruptions between stimulus presentations become increasingly longer, the responses of neurons become increasingly stable. These stabilization effects can be observed both over sequences of tens of stimulus presentations and in subsequent trials. They are likely caused by an increase in coherence of neurons that are part of a local cortical network. Hints for such network-based stabilization can be found in the blank duration dependent changes of local field potential power in the high Gamma range. Such stabilization through increased network coherence may be the cause of an additionally observed stabilization of spike timing relative to stimulus onset.

When confronted with an ambiguous visual stimulus, the brain's initial 'choice' for one of the two perceptual interpretations is generally thought to depend on small random fluctuations in the activity patterns of the neurons that represent the two competing percepts (Noest et al., 2007). A reduction of these minor neuronal activity fluctuations through response stabilization may thus have profound stabilizing effects on perception. On this account, it is intriguing to note that the response stabilization demonstrated here depends similarly on off-durations as the perceptual stabilization that was demonstrated previously (Klink et al., 2008a; Noest et al., 2007). Whereas current computational models can explain a wide range of experimental findings on perceptual stabilization (Pearson & Brascamp, 2008), they do not yet contain enough detail to reproduce our measures of variability. Taking these models to the next level would require the incorporation of complex local network interactions and subthreshold neuronal dynamics.

Since all the basic effects we report in this study are seen with all three stimuli, neuronal response stabilization might be a general neural mechanism that facilitates the efficient encoding of repeated stimulus presentations by increasing the signal-to-noise ratio. The fact that response stabilization is stronger for longer stimulus interruptions could be the signature of a time-consuming network-driven stabilization mechanism the dynamics of which is yet to be unraveled.

## 4.5 Experimental procedures

Responses were obtained from 94 neurons in two adult male rhesus macaque monkeys (*Macaca mulatta*, 46 cells from one monkey and 48 from the other) weighing 8-12 kg. Before the experiments, both monkeys were surgically implanted with a head-holding device and stainless-steel recording chamber (Crist Instruments, Hagerstown, MD) that provided access to area MT. Monkeys were trained to sit in a custom-made primate chair with their head fixed and maintain their gaze within a 2° diameter virtual fixation window surrounding a white rectangular spot (0.5° × 0.5°, 102 cd/m<sup>2</sup>) on the screen that was located at a distance of 65.0 cm. Upon successful fixation the fixation spot turned red (19 cd/m<sup>2</sup>) and the monkey was periodically rewarded with liquid drops. Eye-positions were continuously recorded with an infrared video eye tracker sampling at 500 Hz (EyeLink; SR Research, Osgoode, Ontario, Canada). Breaking fixation resulted in interruption of stimulus presentation and no reward. All surgical and experimental procedures complied with Dutch and European laws and guidelines and were approved by the Utrecht University Animal Experiments Review Committee.



#### 4.5.1 Electrophysiology

For each recording session, a parylene-insulated Tungsten microelectrode (0.1-2.0 M $\Omega$  impedance) was manually inserted through a stainless-steel guide tube that passed the dura. This electrode was then slowly advanced with a hydraulic micropositioner (David Kopf Instruments, Tujunga, CA). Area MT was identified by the recording position and depth, the transitions between gray matter, white matter and sulci, and the functional properties of encountered neurons along the electrode track. The extracellularly recorded signal was pre-amplified 1,000 times (Bak Electronics, Germantown, MS) and 50Hz noise was removed using an adaptive filter (HumBug, Quest Scientific, Canada). The raw signal was then split into two streams, one of which was sampled and saved at 4 MHz using a CED 1401 data acquisition system together with the Spike2 software (Cambridge Electronic Design, Inc) to allow offline analysis of the local field potential trace. The other stream was bandpass-filtered (using a Krohn-Hite 3362 filter) between 1.0-2.0 kHz after which action potentials from single neurons were detected with a window discriminator (Bak Electronics, Germantown, MS). Spiketimes were collected at a 2 kHz resolution using a Macintosh G4 computer equipped with a National Instruments PCI-1200 data acquisition board.

When an MT neuron was isolated, direction selectivity was initially assessed with a contra-laterally presented wide-field motion stimulus comprised of high-density white dots, that were coherently moving against a black background in 8 different, pseudorandomly alternating directions. The size and location of the neuron's receptive field were then mapped by manually moving a bar of light across the visual field. Finally, the precise direction tuning and preferred speed of the neuron were determined by presenting additional moving dot patterns within the determined receptive field.

#### 4.5.2 Visual Stimuli

Stimuli were generated with custom-written software on a Macintosh G4 computer and presented on a 21" monitor running at a resolution of 1024  $\times$  768 pixels and a refresh rate of 100 Hz. All stimuli consisted of 200 white dots (102 cd/m<sup>2</sup>) presented on a black background ( $\sim$  0.01 cd/m<sup>2</sup>). The individual dot size was 0.2°  $\times$  0.2° and the entire stimulus size was 6.7°  $\times$  6.7°. All stimuli were presented in sequences of up to 120 trials, where they were visible for periods of 500 ms and removed from view for a pseudo-randomly assigned blank duration of 250, 500, 1000 or 2000 ms (Figure 4-1A). For the ambiguous structure-from-motion (SFM) cylinder stimulus, dots were randomly positioned at each stimulus onset to mimic the two-dimensional projection of a three-dimensional transparent cylinder. During the 500 ms that this cylinder was presented the dots coherently moved with an unlimited dot lifetime simulating a rotating cylinder. The speed of the simulated rotation was chosen to match the neuron's preferred speed and the axis of rotation was orthogonal to the neuron's preferred and null directions causing one half of the dots to move in the preferred direction and the other half in the null direction. Two secondary stimuli were 1) a dynamic random dot pattern (coherence 0%) for which all dots had a single frame dot lifetime and a randomized starting position on every presentation, and 2) an opaque SFM cylinder that was similar to the primary ambiguous cylinder, but here only the 100 dots that moved in the neuron's preferred direction were presented on the screen.



### 4.5.3 Data analysis

Raw Fano Factors were obtained per cell for every combination of stimulus and off-duration by dividing the mean spikecount by the variance in a 70 ms window sliding in 10 ms increments (one PSTH bin) over the individual stimulus responses. Two additional alternative methods of Fano Factor calculation were performed. In the first method, a linear regression was calculated between the mean spikecount and the spike count variance of all cells for each window position (see also Churchland et al., 2010). The slope of this regression gives one an estimate of the variance to mean ratio. This regression analysis was also performed on a randomly picked subset of the cells that had the same overall mean spikecount for all off-durations within one of the stimulus types. Such a ‘mean-matched’ Fano Factor (Churchland et al., 2010) is strongly resistant against Fano Factor changes that are primarily driven by variations in the mean spike count, rather than in variability. While some quantitative differences in Fano Factor dynamics were present between the three different analyses, the qualitative pattern was highly consistent. We therefore report only the results from the first, cell-based, Fano Factor determination.

The Local Activity Contrast (LAC) was calculated by dividing the absolute difference in spikecount between subsequent trials by their mean spike count. Since this measure turned out to be strongly correlated with the average spike count on the included trials, we derived the theoretical relationship between these two variables under the assumption of Poisson spiking (Bair et al., 1994; Buracas & Albright, 1996; Maimon & Assad, 2009). This relationship was then used to divide each recorded LAC value by the expected value of LAC ( $\langle \text{LAC} \rangle$ ) at each level of activity. If this ratio would be 1.0, the observed LAC corresponds to the expected  $\langle \text{LAC} \rangle$  for Poisson spiking.  $\text{LAC}/\langle \text{LAC} \rangle$  values larger than 1.0 indicate more trial-to-trial spiking variability than expected from a Poisson process, whereas  $\text{LAC}/\langle \text{LAC} \rangle$ -values lower than 1.0 indicate that the spike counts in subsequent trials are more regular than expected from a Poisson process.

To calculate the behavior of LAC under the null-hypothesis of Poisson firing (Equation 4-3) with spike rate  $\lambda$ , we focus on the absolute spike count difference  $\delta$  between subsequent trials at a fixed total spike count  $n$  (Equations 4-4 & 4-5). Expressed in these terms, the LAC can be defined as denoted in Equation 4-6.

$$p[k] = \lambda^k e^{-\lambda} / k! \quad (\text{Eq. 4-3})$$

$$\delta = k_1 - k_2 \quad (\text{Eq. 4-4})$$

$$n = k_1 + k_2 \quad (\text{Eq. 4-5})$$

$$\text{LAC} = 2|\delta|/n \quad (\text{Eq. 4-6})$$

A useful classical result is that the conditional spike count distribution for a trial ( $k_i$ ) under a fixed total spike count  $n$  is the binomial distribution of  $k_i$  hits in  $n$  attempts with probability 0.5 (Equation 4-7). Fixed  $n$  also implies that the spike count difference  $\delta$  is directly related to both the single trial spike count and the total spike count (Equation 4-8). As a result, the distribution of  $\delta$  is a scaled and shifted version of the trial spike count binomial, such that it has mean zero and variance  $n$ .



$$\mathbb{P}[k_1|n] = 2^{-n} \binom{n}{k_1} \quad (\text{Eq. 4-7})$$

$$\delta = 2k_1 - n \quad (\text{Eq. 4-8})$$

For reasonably large spike counts, this  $\delta$  distribution becomes Gaussian, so we can easily compute the expected spike count difference (Equation 4-9), and the expected  $\langle LAC \rangle$ , conditional upon the total spike count (Equation 4-10). Trivially, one can also express the result in terms of the spike count mean:  $m = n/2$ .

$$\langle |\delta| \rangle = 2 \int_0^{\infty} \frac{\delta}{\sqrt{2\pi n}} \exp\left(-\frac{\delta^2}{2n}\right) d\delta = -\sqrt{\frac{2n}{\pi}} \exp\left(-\frac{\delta^2}{2n}\right) \Big|_0^{\infty} = \sqrt{\frac{2n}{\pi}} \quad (\text{Eq. 4-9})$$

$$\langle LAC \rangle = 2 \sqrt{\frac{2}{\pi n}} = \frac{2}{\pi m} \quad (\text{Eq. 4-10})$$

For the analysis of precise spike timing the timestamps of the spikes were used relative to stimulus onset. Comparisons were made between spiketimes in subsequent trials by calculating the Spike Time Deviation (STD), defined as the temporal distance between a spike in the first spiketrain and its closest neighbor in the subsequent spiketrain (Lankheet et al., 2005) (Figure 4-4A). Since higher spike densities would automatically result in smaller STD, the STD's were normalized by the average interspike interval of the two constituent spiketrains, yielding a Spike Time Deviation Index (STDI).

Local Field Potential (LFP) magnitudes were expressed as the root-mean-square (RMS) of the LFP signal, which was calculated by taking the square root of the mean of the squared LFP trace amplitudes for all trials in a sequence (Liu & Newsome, 2006). An LFP-analogue of the Fano Factor was obtained by dividing the square root of the variance in the squared LFP amplitude over trials by the square root of mean of these values. The power spectra of the LFP between 0 and 120 Hz were calculated in 2.4 Hz bins using a discrete Fourier transform in a 200 ms sliding window that moved with 20 ms increments. Calculated power spectra were normalized to the average power during stimulus presentation over a full stimulus sequence within individual cells. Stimulus-evoked changes in power were statistically tested by contrasting the average power in the 100 ms directly following stimulus presentation with the power during stimulus presentation with a cell-based paired t-test.

## 4.6 Acknowledgements

We thank André Noest for discussions and advice about data analysis, Jacob Duijnhouwer for assistance with the custom software and Theo Stuivenberg and Hans Borgeld for technical support. Bert van den Berg and Joost le Feber commented on earlier versions of this manuscript. This work was supported by an NWO-VIDI grant and a UU High Potential grant.

# Chapter 5

**Occlusion-related lateral connections stabilize kinetic depth stimuli through perceptual coupling**

*An individual neuron is in fact rather dumb. It is the intricate  
of many of them together that can do such marvelous things.*

Francis Crick

**Published as**  
Klink, P.C., Noest, A.J., Holten, V., & van Wezel, R.J.A. (2009). Occlusion-related lateral  
connections stabilize kinetic depth stimuli through perceptual coupling.  
*Journal of Vision*, 9(10):20, 1-20.



When André Noest's theoretical model for the neural mechanisms underlying perceptual choices at the onset of ambiguous visual stimuli was published (Noest et al., 2007), it started a period in which the temporal aspects of rivalry were extensively investigated in our group. This chapter describes the results of an attempt to expand our horizon and investigate some spatial aspects of rivalry resolution. Before we performed the experiments that eventually ended up in this chapter, I had two ideas for potential experiments. First, I wanted to see what would happen if multiple ambiguously rotating spheres would be simultaneously presented to an observer. Would all perceived rotation directions be the same, or would they be independent? Furthermore, if we would use depth cues to disambiguate one or several stimuli, how would such a manipulation influence the other, ambiguous stimuli? Second, I wanted to display an ambiguously rotating sphere against a background of moving dots to see whether this could evoke any contextual, center-surround-like effects. As it turned out, both these experiments had already been performed and published by other people several years earlier (Gilroy & Blake, 2004; Sereno & Sereno, 1999). We could however still use the experimental code that I had already programmed as an introduction to a student's project on spatial interactions in the perceptual resolution of an ambiguously rotating sphere. While we were doing experiments with two simultaneously presented coaxial cylinders, we were pointed to some existing Gestalt psychology literature on the perceptual differences between objects when they are either located in the foreground or in the background. The incorporation of these ideas in our rivalry paradigm eventually resulted in the series of experiments and the neural network model that are presented in this chapter.

---



## 5.1 Abstract

Local sensory information is often ambiguous forcing the brain to integrate spatiotemporally separated information for stable conscious perception. Lateral connections between clusters of similarly tuned neurons in the visual cortex are a potential neural substrate for the coupling of spatially separated visual information. Ecological optics suggests that perceptual coupling of visual information is particularly beneficial in occlusion situations. Here we present a novel neural network model and a series of human psychophysical experiments that can together explain the perceptual coupling of kinetic depth stimuli with activity-driven lateral information sharing in the far depth plane. Our most striking finding is the perceptual coupling of an ambiguous kinetic depth cylinder with a coaxially presented and disparity defined cylinder backside, while a similar frontside fails to evoke coupling. Altogether, our findings are consistent with the idea that clusters of similarly tuned far depth neurons share spatially separated motion information in order to resolve local perceptual ambiguities. The classification of far depth in the facilitation mechanism results from a combination of absolute and relative depth that suggests a functional role of these lateral connections in the perception of partially occluded objects.

## 5.2 Introduction

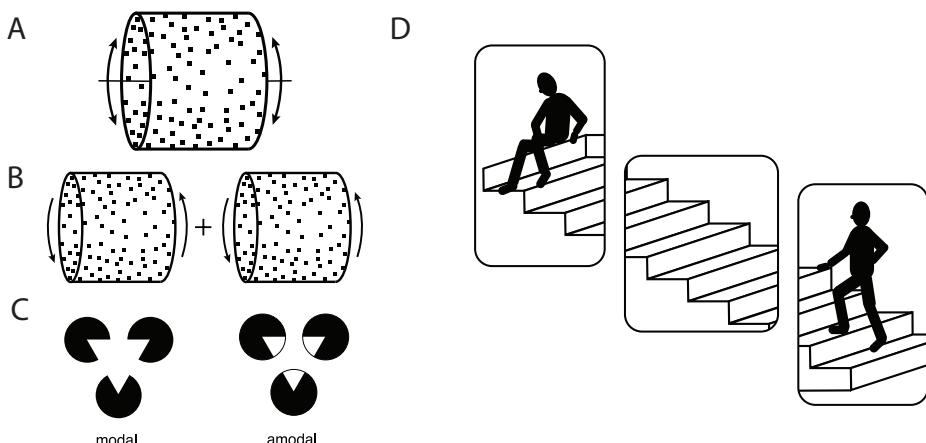
Local visual information is massively ambiguous, but fortunately the visual system does not base conscious perception on local information alone. Spatial and temporal contexts are highly effective in disambiguating local visual information, which results in a perceptual system that is relatively stable and able to interpret sensory input more globally. When the brain reconstructs the three-dimensional world from a two-dimensional projection on the retina it uses a multitude of cues such as stereoscopic disparity, occlusion, shading or (relative) motion patterns (for an extensive review see Howard & Rogers, 2002). A nice example of how context shapes the three-dimensional interpretation of two-dimensional images can be found in the famous lithograph '*Relativity*' by M.C. Escher (Escher, 1992). It depicts a world with multiple gravity sources in which the depth interpretation of a room is disambiguated by the presence of people going up or down a set of stairs.

In the laboratory most of the contextual information is often removed from visual stimuli to study highly specific mechanisms of visual processing (Rust & Movshon, 2005). The inference of three-dimensional structure from contextual cues for example can be studied with stimuli that lack explicit depth cues, but whose motion pattern gives rise to the perception of a three-dimensional object. A vivid example of such a stimulus is the two-dimensional projection of a rotating transparent cylinder covered with points, constructed from two layers of randomly positioned dots moving in opposite directions (e.g. Andersen & Bradley, 1998; Kourtzi et al., 2008)(Figure 5-1A). In the absence of an explicit depth ordering of the two dot layers, this stimulus is bistable with respect to its rotation direction. Bistable stimuli in general offer equal sensory evidence for two mutually exclusive perceptual interpretations causing conscious perception to alternate between the possible interpretations while the stimulus remains the same (for reviews, see Blake & Logothetis, 2002; Leopold & Logothetis, 1999). In the case of the bistable cylinder this means that upon prolonged viewing the rotation direction is perceived to switch every few seconds (Andersen & Bradley, 1998; Nawrot & Blake,



1989; Treue et al., 1991).

Adding context or depth cues to a bistable cylinder can overcome the rotation direction ambiguity and bias the stimulus towards one, more or less, stable perceptual interpretation. These cues can be part of the stimulus itself (e.g. Dosher et al., 1986; Klink et al., 2008b; van Ee et al., 2002) acting on a local scale or they can be an added context that influences perception in a global manner. Examples of global contextual influences are center-surround interactions between the cylinders and surrounding motion patterns (Sereno & Sereno, 1999), an apparent friction effect when two spheres rotating around parallel axes appear to touch (Gilroy & Blake, 2004) or the perceptual coupling of multiple coaxially rotating stimuli (Eby et al., 1989; Freeman & Driver, 2006; Grossmann & Dobbins, 2003)(Figure 5-1B). The last case is particularly interesting, since it shows that even an ambiguous context can have strong rivalry resolving effects. It suggests that the visual system combines spatially separated information to minimize the degree of visual conflict in the scene (e.g. Attneave, 1968; Freeman & Driver, 2006; Ramachandran & Anstis, 1983). It has been shown that the extent of this perceptual coupling is largest for two ambiguous cylinders, but coupling also occurs if one of the two stimuli is rendered less ambiguous by either adding disparity or a luminance gradient (Freeman & Driver, 2006; Grossmann & Dobbins, 2003). However, whereas for a full disparity defined cylinder the coupling persists, it is strongly reduced -or absent- for stimuli with a



**FIGURE 5-1.** A) Schematic representation of a kinetic depth cylinder stimulus. The spatial distribution and speed profile of the dots create the vivid impression of a three-dimensional cylinder rotating around a vertical axis. Without explicit depth cues the rotation direction is ambiguous and bistable. The axis drawn here was not present in the actual stimulus. B) Two coaxially presented stimuli have a strong tendency to be perceived as rotating in the same direction. C) Examples of modal and amodal completion with Kanizsa triangles (Kanizsa, 1979). In the top image, a white triangle appears to float in front of black circles. The illusory triangle surface is constructed through modal completion. The lower image's white triangle is perceived as through a set of apertures in a white 'foreground' (amodal completion) while the black shapes are perceived as part of an occluded black 'background'. D) Amodal spatial facilitation can resolve local ambiguities. An image of an occluded Schröder's staircase, looked at through three apertures. The image in the middle aperture has ambiguous depth information whereas the left and right are disambiguated by contextual information. If the middle aperture is combined with only one of the two flanking apertures, amodal facilitation disambiguates the depth structure in the middle aperture.

---



maximal luminance gradient (Freeman & Driver, 2006). Such a maximal luminance gradient effectively reduces a cylinder to a single layer of dots. This has led to the proposition that perceptual coupling between cylinders depends on the presence of both surface layers of the two cylinders (Freeman & Driver, 2006). The functional mechanism of perceptual coupling however still remains unclear.

Here we present an alternative explanation for perceptual coupling that has not been previously considered or studied. We hypothesize that perceptual coupling reflects a more common neural mechanism involved in the perception of partially occluded objects or scenes. The visual system can resolve local ambiguities by combining information from different spatially separated locations (e.g. Georgeson et al., 2008; Spillmann & Werner, 1996; van der Smagt & Stoner, 2008; Watanabe & Cole, 1995; Yang & Blake, 1995). In real life situations this is particularly useful when objects are partially occluded. When we encounter occlusion, the brain binds the separate chunks of visual information and we perceive a single occluded object rather than multiple separate objects. This perceptual construction of objects that are partially occluded or seen through an aperture is known as *amodal* completion as opposed to the construction of illusory contours and surfaces in the foreground, which is termed *modal* completion (e.g. Anderson et al., 2002; Kanizsa, 1979)(Figure 5-1C). Amodal completion is thought to be a hardwired mechanism by which spatial facilitation resolves locally ambiguous visual information (e.g. Driver et al., 2001)(Figure 5-1D). Amodal completion has been shown in a multitude of species such as domestic chicks (Forkman, 1998), pigeons (Nagasaka & Wasserman, 2008), mice (Kanizsa et al., 1993) and baboons (Fagot et al., 2006) as well as for a broad range of stimulus dimensions such as shape (e.g. Anderson et al., 2002), color (Pinna, 2008) or sound (Miller et al., 2001). The widespread occurrence of amodal completion combined with the strong contrast between the apparently effortless perception of partially occluded objects and the difficult detection of camouflaged objects suggests that the visual system is better equipped for sharing spatially separated information in the far depth plane (amodal) than in the near depth plane (modal). The findings that human observers are better in judging the relative alignment of two gratings (Anderson et al., 2002) and in recognizing faces (Nakayama et al., 1989) if they are presented in an amodal rather than modal fashion add further evidence to this suggestion.

In the current study we present a neural network model and a series of human psychophysical experiments that together provide experimental and computational evidence in support of our explanation of the perceptual coupling of kinetic depth stimuli based on amodal spatial facilitation. Our model is a straightforward extension of existing models for perceptual rivalry (Brascamp et al., 2008; Klink et al., 2008a; Noest et al., 2007) and kinetic depth (Andersen & Bradley, 1998; Nawrot & Blake, 1989). It incorporates amodal spatial facilitation through lateral connections between neuronal populations with similar tuning properties that code for spatially separated stimuli situated in the far depth plane. In our psychophysical experiments we use a percept-choice paradigm in which the kinetic depth stimuli are presented intermittently with a temporal profile that would normally cause perceptual stabilization (Klink et al., 2008a; Leopold et al., 2002; Maier et al., 2003; Noest et al., 2007; Pearson & Brascamp, 2008).

In two initial experiments that are added as appendices to this manuscript, we replicate previous findings (e.g. Freeman & Driver, 2006) with our new experimental paradigm and demonstrate: 1) How dot luminance and stereoscopic disparity influence the perceptual inter-



pretation of single kinetic depth cylinders; 2) That perceptual coupling between coaxial cylinders occurs for all disparity biases, but collapses for large luminance gradients, and; 3) That the direction of information sharing is not necessarily from the cylinder with depth cues to the ambiguous cylinder, but rather from the ‘more certain’ to the ‘least certain’ representation.

The experiments described in the main text of the manuscript further aim to unravel the nature of the perceptual coupling mechanism. Experiment 1 demonstrates perceptual coupling between disparity defined single surface ‘backsides’ and complete ambiguous stimuli for both cylinders and spheres. These findings demonstrate that spatial facilitation takes place in the background and cannot be simply attributed to surface continuation. Experiment 2 demonstrates that the collapse of perceptual coupling with increasing dot luminance gradients scales with the distance between the two cylinders. This finding supports the idea that the shared information decays over traveled distance and stronger signals in the background are needed to establish perceptual coupling across larger gaps. Experiment 3 investigates the nature of the spatial facilitation signal with asynchronously presented stimuli and reveals that perceptual coupling must occur on a fast activity-driven, rather than a slow adaptation-driven timescale. Our fourth experiment aims to unravel the roles of absolute and relative depth in spatial facilitation. In other words, does coupling occur between backsides (relative depth) or ‘far depth’ surfaces (absolute depth)? The results of this experiment indicate that the coupling mechanism depends on a mixture of absolute and relative depth that is functionally very well suited to deal with occlusion.

The model and experiments were both developed to test our functional hypothesis that the perceptual coupling of kinetic depth stimuli relies on spatial facilitation in the far depth plane. Even though the two approaches form a coherent argument in favor of this hypothesis they might be read independently of each other. The amount of detail in the neural network section of this paper is not strictly necessary to understand the psychophysical results. The experimental results on the other hand may facilitate a better understanding of the model section, but they are also not strictly necessary for it.

### 5.3 A neural network model

Classic models of bistable stimuli such as the kinetic depth cylinder are based on competing neuronal populations coding for two mutually exclusive perceptual interpretations. These neuronal populations are subject to adaptation and they are generally believed to interact via cross-inhibitory connections. The model we present here is based on a recently developed physiologically plausible, single-stage model of visual competition (Noest et al., 2007). This model was developed specifically to gain insight in the mechanism that selects a conscious percept at the onset of a visual rivalry stimulus. The model describes this selection process as a classic competition between mutually inhibitory, percept-coding neural populations. During dominance of a given percept, the response properties of the neurons coding for this percept are altered in a way that does not immediately revert when dominance ends. These continued altered response characteristics thus carry a memory trace of prior dominance (Brascamp et al., 2009). In the percept-choice paradigm, the intermittent presentation of visual rivalry stimuli offers a window on these implicit memory traces. Long interruptions (up to seconds) between stimuli result in sequences of repeated dominant percepts (Klink et al., 2008a; Leo-



pold et al., 2002; Noest et al., 2007), whereas shorter interruptions (less than half a second) result in perceptual alternations on subsequent presentations (Klink et al., 2008a; Noest et al., 2007). The Noest-model can account for these findings with an interaction of a neural baseline ( $\beta$  parameter in the model) with the adaptation dynamics of the percept coding neural populations. This interaction functionally creates a head start in the neural competition for the more adapted population at the next stimulus onset (at the level of its near-threshold field potential).

If the adaptation levels are high, they will easily overcome the small head start, causing the least adapted neural population to 'win' the competition resulting in a classic perceptual alternation. However, if the adaptation levels are too low to overcome the head start, the more adapted neural population will become dominant again at the next stimulus onset, causing perceptual repetitions. The adaptation levels of the competing populations build up during stimulus presentation and decay during the intermittent blank periods. Consequently, short interruptions will allow little decay of adaptation and the resulting high adaptation levels lead to perceptual alternations. Longer interruptions on the other hand, allow much more adaptation decay, resulting in lower adaptation levels at the next stimulus onset and thus in perceptual repetitions.

### 5.3.1 Kinetic depth

Our current model has the same internal dynamics as the original Noest-model, but for the kinetic depth cylinders we have split up the percept-coding neural populations in surface-coding neural populations (Andersen & Bradley, 1998; Nawrot & Blake, 1991b). There is sufficient reason to assume that the percept of kinetic depth cylinders is constructed through the depth ordering of the two dot-layers that constitute the concave/convex front and backside of the cylinder (Klink et al., 2008b; Li & Kingdom, 1999; Nawrot & Blake, 1991b; Treue et al., 1995). This leaves us with a set of four neural populations, each coding for a combination of depth order and motion direction, eventually giving rise to the percept of a bistable rotating cylinder (Figure 5-2A, Equations 5-1 & 5-2). Classic cross-inhibitory connections are assumed between neural populations coding for opposite directions at the same depth level and same directions at different depth levels. Weak facilitatory connections are assumed between opposite directions at different depth planes for considerations of surface continuity (even though they do not crucially change the model's behavior). Since fully opaque kinetic depth stimuli (only one motion direction visible) are predominantly perceived as convex (see for example our results in Appendix 1, Figure 5-A1A), we incorporate a small positive bias for 'near' over 'far' surfaces. This manipulation is also in agreement with the idea that relatively small stimuli that are surrounded by a uniform, differently colored region are interpreted as 'figure' or foreground and thus perceived as closer to the observer (Rubin, 2001). The quantitative predominance of neurons tuned for near depth over those tuned for far depth that has been demonstrated in many visual cortical areas (area V3: Adams & Zeki, 2001; Bradley & Andersen, 1998; DeAngelis & Uka, 2003; Gonzalez et al., 2001; Hinkle & Connor, 2005; area MT: Maunsell & Van Essen, 1983; area MST: Roy et al., 1992; area IT: Uka & Deangelis, 2003; area V2: von der Heydt et al., 2000; area V4: Watanabe et al., 2002) could also be an indication of a bias for near over far surfaces.



Mathematically the model can be written as:

**Kinetic depth** (Figure 5-2A)

$$\tau \partial_t h_{i,m} = X_{i,m} - \beta - (1 + A_{i,m}) h_{i,m} - \gamma_D S[h_{i,n} + \beta] + \varepsilon S[h_{j,n} + \beta] \quad (\text{Eq. 5-1})$$

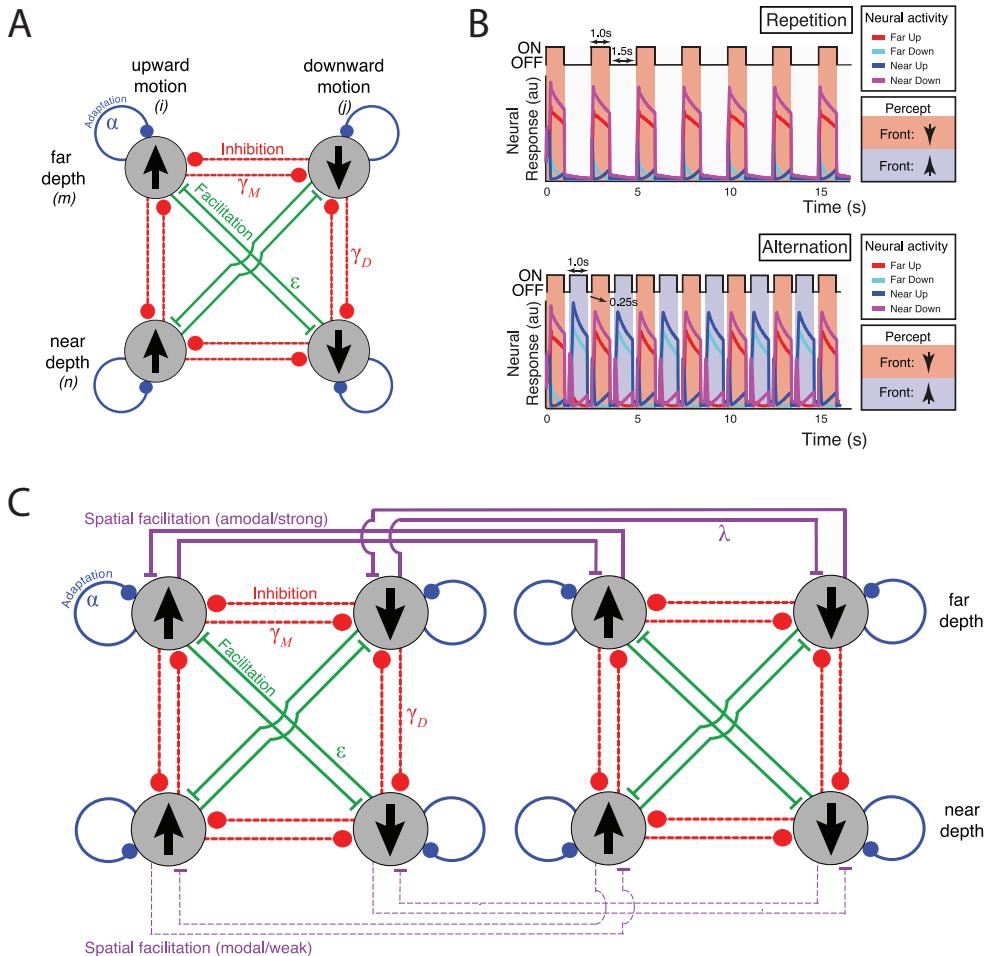
$$\partial_t A_{i,m} = -A_{i,m} + \alpha S[h_{i,m} + \beta] \quad (\text{Eq. 5-2})$$

$i, j, m, n \in \{1, 2\}, i \neq j, m \neq n$

In these equations  $X$  represents the visual input,  $i$  &  $j$  represent the two motion directions,  $m$  &  $n$  represent the two depth levels,  $\alpha$  is the strength of neuronal adaptation,  $\beta$  can be regarded as an intraneuronal baseline (for details, see Noest et al., 2007);  $\gamma_D$  represents the strength of classic cross-inhibition between depth levels and  $\gamma_M$  that between motion directions and  $\varepsilon$  represents the strength of a surface continuity facilitation (Figure 5-2A). The fast ‘local field’ activity ( $h$ ) of the neural populations is translated into a spike rate by a sigmoid function  $S$  and undergoes a slow shunting type adaptation (Equation 5-2). For more details on the internal dynamics of the neural populations see Noest et al. (2007) or Klink et al. (2008a).

Our four-population version of the model reproduces the aforementioned findings about the timing of intermittent presentation that were demonstrated with the original two-population version of the model (Klink et al., 2008a; Noest et al., 2007): short interruptions cause perceptual alternations, longer interruptions cause perceptual repetitions (Figure 5-2B). The neural populations, that together code for a coherent cylinder percept, modulate their activity in synchrony with the near surfaces having stronger responses than far surfaces. The parameters we used in these simulations were taken from the original publication of the Noest model (Noest et al., 2007). This gave us  $\alpha = 5$ ,  $\beta = 4/15$  and  $\tau = 1/50$ . Since we doubled the number of populations involved in constituting a percept compared to the original Noest-model interpretation, we end up with twice the number of cross-inhibitory connections. To stay in accordance with the original parameter-set we divided the strength of the original cross-inhibitory connections by two, leaving us with  $\gamma_D = \gamma_M = 5/3$ . The small surface continuity facilitation that we propose was set to  $\varepsilon = 0.1$ , but setting it to zero did not significantly change the simulation results. The basic input to the model was set as  $X_{near} = 1$ , while the advantage of near over far surfaces was incorporated as  $X_{far} = 0.75 * X_{near}$ .

To convert simulated neural responses to percepts we calculated the average activity of all four populations during the entire presentation-period of the stimulus and determined a single combination of dominant front and back directions via a winner-take-all mechanism. By using the average activity over the whole presentation epoch, we mimic the perceptual decision process of our human observers that are also allowed to use the entire presentation duration to reach a decision about their percept. This approach could in principle lead to four different percepts: 1) A rotating cylinder with the front moving upwards, 2) A rotating cylinder with the front moving downwards, 3) Two convex surfaces moving in opposite directions, and 4) Two concave surfaces moving in opposite directions (Hol et al., 2003). In our simulations, we only encountered the two consistent cylinder percepts.



**FIGURE 5-2.** A) A neural network model of kinetic depth. Four adaptive ( $\alpha$ ) populations of neurons each code for a combination of depth and motion direction. Facilitatory connections between the populations establish surface continuity ( $\varepsilon$ ) while inhibitory connections ( $\gamma_D, \gamma_M$ ) are crucial for visual rivalry characteristics. B) Simulations of the neural network confirm earlier findings (Noest et al., 2007) for percept-choice dynamics with interrupted stimuli. Colored lines represent the simulated response of the four neural populations and colored shading represent percepts, inferred from the neural responses via a winner-take-all mechanism. Long interruptions cause repetitions (perceptual stabilization), short interruptions cause percept-choice alternations. Parameters used are:  $\alpha = 5$ ,  $\beta = 4/15$ ,  $\tau = 1/50$ ,  $X_{near} = 1$ ,  $X_{far} = 0.75 * X_{near}$ ,  $\gamma_D = \gamma_M = 5/3$ ,  $\varepsilon = 0.1$ . C) A neural network model for (amodal) spatial facilitation in kinetic depth. Two 'single cylinder' networks are coupled by facilitatory connections between similarly tuned neural populations. Note that these lateral connections are stronger in the far depth plane (*amodal*) than in the near depth plane (*modal*).



### 5.3.2 Spatial facilitation

Two cylinders that are presented simultaneously can be modeled with two sets of four neuronal populations, each with their own inhibitory and facilitatory connections as described above. The principles of modal and amodal completion suggest that there may be lateral connections between similarly tuned populations of neurons coding for the different cylinders (Figure 5-2C, Equation 5-3 in Appendix 3). As we noted in the introduction, ecological optics (Gibson, 1950) would suggest that the visual system is better equipped to deal with occlusion than with camouflage, leading us to assume that the facilitatory connections are stronger in the far (*amodal*) than in the near (*modal*) depth plane (purple lines in Figure 5-2C). These considerations result in the mathematical description:

#### Spatial facilitation (Figure 5-2C)

$$\begin{aligned} \tau \partial_t h_{i,m,p} = X_{i,m,p} - \beta - (1 + A_{i,m,p}) h_{i,m,p} - \gamma_D S[h_{i,n,p} + \beta] \\ - \gamma_M S[h_{j,m,p} + \beta] + \varepsilon S[h_{j,n,p} + \beta] + \lambda_m S[h_{i,m,q} + \beta] \end{aligned} \quad (\text{Eq. 5-3})$$

$$\begin{aligned} \partial_t A_{i,m,p} = -A_{i,m,p} + \alpha S[h_{i,m,p} + \beta] \\ i, j, m, n, p, q \in \{1, 2\}, i \neq j, m \neq n, p \neq q \end{aligned} \quad (\text{Eq. 5-4})$$

The only difference between the fast dynamics Equations 5-1 & 5-3 is that the newly introduced  $p$  &  $q$  represent the two coupled sets of neural populations and the  $\lambda$  term indicates the strength of the spatial facilitation. Equation 5-2 & 5-4 that denote the adaptation dynamics are identical.

We developed our model in order to account for the existing experimental data that demonstrated that perceptual coupling occurs between two coaxial ambiguous cylinders and between a disparity defined and an ambiguous cylinder, but not between a fully luminance defined and an ambiguous cylinder (Freeman & Driver, 2006). Simulations with the model of amodal spatial facilitation were performed to investigate the properties of perceptual coupling. Depending on the duration of the blank period, two bistable cylinders either stabilize or alternate together (Figure 5-3A).

In our biased cylinder simulations, a temporal presentation profile was used that would normally give rise to sequences of repeated percepts (1.0 second presentations with 1.5 seconds blank periods; see also Figure 5-2B). Depth cues were incorporated in the model by multiplying the input the neural populations with a modulation factor  $M$ . Introducing depth cues to the two cylinders with dot luminance or stereoscopic disparity has different effects on the activity of the surface-coding neuronal populations (Table 5-1). Dot luminance manipulations result in biases that are based on motion direction only and will thus affect the activity of the two populations coding for the same direction of the manipulated dots regardless of their depth assignment (vertically positioned pairs of populations in Figure 5-2A). The small positive bias for ‘near’ over ‘far’ surfaces ensures that the brightest dots are perceived as the ‘near’ side of the cylinder. Stereoscopic disparity manipulations on the other hand, result in biases based on combined motion and depth information and will consequently affect the relative



activities of the two pairs of populations coding for a coherent cylinder percept (diagonally positioned pairs of populations in Figure 5-2A). To visualize the effect of the depth cues, we simulated a switch in cue direction on each consecutive presentation (grey and green M-lines in Figures 5-3B to 5-3D). If full disparity cues are used, the two coupled cylinders together follow the biased direction (Figure 5-3B), but with full dot luminance biases perceptual coupling collapses (Figure 5-3C), which is in agreement with the existing data. When stimulus biases are relatively small, they are no longer the strongest percept-determining feature. The perceptual stabilization that arises from the intermittent stimulus presentation with long blank periods (Klink et al., 2008a; Leopold et al., 2002; Maier et al., 2003; Noest et al., 2007; Pearson & Brascamp, 2008) is now more effective and the two cylinders appear coupled but their rotation direction stabilizes (Figure 5-3D) despite the alternating depth cue biases.

The strengths of spatial facilitation in our simulations was chosen to reproduce the dissociation in coupling between luminance and disparity biases, and to reflect our hypothesis that coupling is stronger in the far than near depth field, resulting in  $\lambda_{far} = 0.4$  and  $\lambda_{near} = \lambda_{far}/5$ . The depth biases were simulated by multiplying  $X_{near}$  with a modulation factor  $M$  so that  $X_{near\_mod} = M * X_{near}$ . To demonstrate the effect of strong luminance and depth cues we modulated  $M$  between 0.5 and 1. A weak modulation of  $M\{0.9-1\}$  was used to reveal the occurrence of “reverse coupling” with the two cylinders being stabilized together despite a depth cue that is alternating in direction.

In a series of psychophysical experiments we also tested the hypothesis that perceptual coupling is driven by connections in the far depth plane (Experiment 1). If such spatial facilitation through lateral connections exists, it is likely to exhibit a certain decay of signal strength with increasing interstimulus distance. A second experiment (Experiment 2) investigates whether the strength of spatial facilitation is indeed a function of the distance between the two cylinders. Our third and fourth experiment shed light on the nature of the facilitatory mechanism and the roles of absolute (depth relative to the plane of fixation) and relative depth (front or back side of the cylinder) respectively.

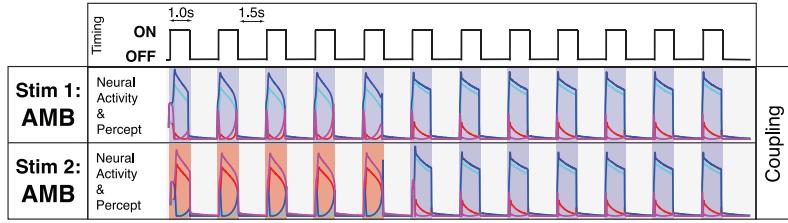
Population	Ambiguous	Luminance depth	Disparity depth
Near/Up	$X = X_{near}$	$X = M_1 * X_{near}$	$X = M_1 * X_{near}$
Near/Down	$X = X_{near}$	$X = M_2 * X_{near}$	$X = M_2 * X_{near}$
Far/Up	$X = 0.75 * X_{near}$	$X = M_1 * 0.75 * X_{near}$	$X = M_2 * 0.75 * X_{near}$
Far/Down	$X = 0.75 * X_{near}$	$X = M_2 * 0.75 * X_{near}$	$X = M_1 * 0.75 * X_{near}$

**TABLE 5-1.** The effect of simulated input modulations on the effective input to the neural populations in our model. To account for the preference for single surfaces to be perceived as being near rather than far we state that  $X_{far} = 0.75 * X_{near}$ . Depth cue modulations affecting the different neural populations of the model are denoted as gain factors  $M_1$  and  $M_2$  (green and grey lines in Figure 5-3). For simulated luminance manipulations the two populations coding for the same motion direction have the same modulation gains, while for simulated disparity modulations the two populations that code for a consistent cylinder (different depth, opposite directions) receive the same gain factor.

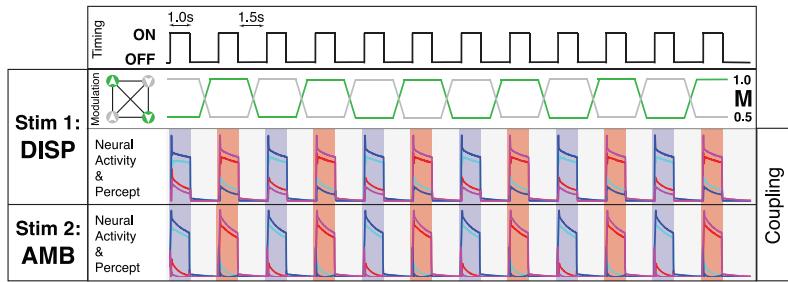


General Legend		
Neural activity	Percept	$X_{\text{near}} = 1.0$ and $X_{\text{far}} = 0.75 \cdot X_{\text{near}}$
<ul style="list-style-type: none"><li>— Far Up</li><li>— Far Down</li><li>— Near Up</li><li>— Near Down</li></ul>	Front: ↓	When biased: $X = M \cdot X$ E.g. $X_{(\text{near},\text{up})} = M \cdot X_{(\text{near})}$ $X_{(\text{far},\text{up})} = M \cdot X_{(\text{far},\text{up})} = M \cdot 0.75 \cdot X_{(\text{near},\text{up})}$
	Front: ↑	

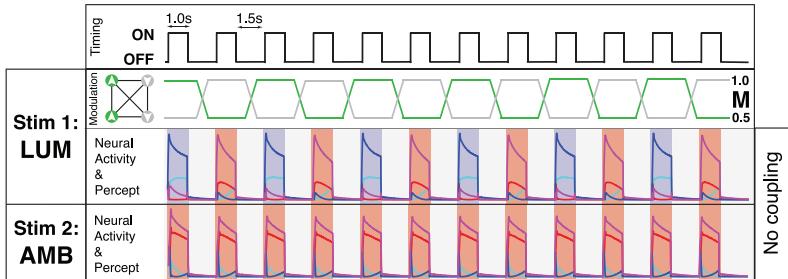
A



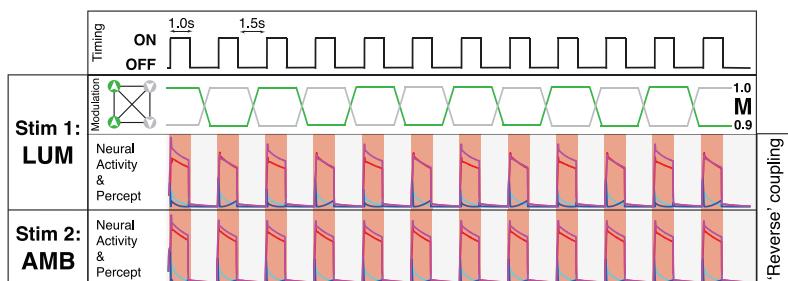
B



C



D





◀ **Figure 5-3.** Simulations of the spatial facilitation model (Figure 5-2C). Colored lines represent the simulated response of the four neural populations and colored shading represent percepts, inferred from the neural responses via a winner-take-all mechanism (see text). **A)** Perceptual coupling between two ambiguous cylinders. During the first few presentations the cylinders are individually stabilized, but later they couple and they stay coupled. The moment coupling kicks in depends on the strength of the spatial facilitation parameter. **B)** Perceptual coupling between an alternating disparity biased and an ambiguous cylinder. The ambiguous cylinder no longer stabilizes but follows the alternating disparity-defined percept of the biased cylinder. The strength and direction of the depth cue bias is given as a modulation parameter  $M$  (see text) and visualized with the green en grey lines that correspond to the green and grey neural populations in the schematic model icon (corresponding to Figure 5-2A) next to it. **C)** No perceptual coupling between an alternating luminance biased and an ambiguous cylinder. The dominant percept of the ambiguous cylinder stabilizes while the luminance-defined percept of the biased cylinder alternates. **D)** Perceptual coupling between a weakly luminance biased and an ambiguous cylinder. The luminance bias alternates direction on consecutive presentations, but is overruled by the perceptual stabilization that couples from the ambiguous to the biased cylinder. Parameters used in the simulation are:  $\alpha = 5$ ,  $\beta = 4/15$ ,  $\tau = 1/50$ ,  $X_{near} = 1$ ,  $X_{far} = 0.75 X_{near}$ ,  $\gamma_D = \gamma_M = 5/3$ ,  $\varepsilon = 0.1$ ,  $\lambda_{far} = 0.4$  and  $\lambda_{near} = \lambda_{far}/5$ . Modulation in **B & C**:  $X_{near} \{0.5-1\}$ , modulation in **D**:  $X_{near} \{0.9-1\}$ .

## 5.4 Methods

### 5.4.1 Observers

Five observers participated in Experiment 1 and 4, four observers in Experiments 2 and 3. In each experiment, one of these observers was an author while the others were naive about the purpose of the study. All observers had normal or corrected to normal visual acuity. After we explained the task and showed the stimuli to the observers we obtained their informed consent.

### 5.4.2 Apparatus

Visual stimuli were generated on a Macintosh computer in MATLAB (Mathworks, Natick, MA) using the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997) and presented on a 22 inch CRT monitor with a resolution of 1600×1200 pixels and a refresh rate of 100 Hz. Observers viewed the stimuli through a mirror stereoscope from a distance of 100 cm.

### 5.4.3 Stimuli

In all experiments, stimuli were kinetic depth cylinders or spheres (only in experiment 1), consisting of white dots on a black background ( $\sim 0$  cd/m<sup>2</sup>), rotating around a horizontal axis with 120 deg/s. Cylinders or spheres were 3×3 deg and the individual dots were 0.11 deg in size. Stimuli without disparity cues were presented monocularly to prevent explicit ‘flatness’. Disparity biases were implemented by horizontally shifting the dots presented to the individual eyes in fractions of the ‘realistic’ disparity (0, 20, 40, 70 and 100%). In the luminance biased condition, the ‘nearest’ dots always had full luminance (69.7 cd/m<sup>2</sup>) while the other dots’ luminance was modulated down to fractions of the full luminance (0, 25, 60, 90 and



100% modulation) depending on their simulated depth. Ambiguous (0% modulation) and disparity defined cylinders thus consisted of dots that were all 69.7 cd/m<sup>2</sup>, whereas e.g. 100% luminance modulated cylinders contained dots ranging in luminance between 0 cd/m<sup>2</sup> (the ‘farthest’ dots in the middle of the back surface) and 69.7 cd/m<sup>2</sup> (the ‘nearest’ dots in the middle of the front surface). Stimuli were presented on the screen for one second separated by 1.5 seconds inter stimulus interval. Blocks of stimulus presentations lasted 120 seconds and conditions were picked in pseudo-random order. During the entire duration of a block there was a fixation cross (6×6 pixels, 69.7 cd/m<sup>2</sup>) at the center of the screen.

#### *5.4.4 Procedure experiment 1: Information sharing in the near and far planes.*

Two coaxial cylinders or spheres were presented spatially separated by a gap of 0.5 degrees. The rightmost stimulus was always completely ambiguous. Only one of the two dot layers of the left stimulus was displayed. This layer could be the far or near side of a cylinder as defined by its luminance gradient or disparity information. We performed a short selection experiment to test whether these cues were sufficient for our observers to impose the specified percept. Only observers that perceived the biases in the veridical direction more than 75% of the time (80 presentations with random bias direction) were selected for this experiment (7 out of 8 observers passed this test). They then performed the experiment in which they only reported the perceived direction of the near/front surface of the full, ambiguous stimulus by pressing a button on the keyboard. The ‘half’ stimulus had a 40% probability of changing its direction on consecutive presentations.

#### *5.4.5 Procedure experiment 2: The spatial decay of perceptual coupling.*

Two coaxial cylinders were presented on each side of the fixation cross. The distance between the cylinders was variable over blocks (0.25, 0.5, 1.0 or 2.0 degrees). The rightmost cylinder was always completely ambiguous whereas the left could have a luminance bias. Observers indicated the perceived direction of the near/front surface of both cylinders by pressing buttons on the keyboard. Any possible stimulus bias (disparity or luminance) again had a 40% probability of changing its direction on consecutive presentations.

#### *5.4.6 Procedure experiment 3: Asynchronous presentation.*

Two coaxial cylinders were presented on both sides of a central fixation cross, spatially separated by a gap of 0.5 degrees. The rightmost stimulus was always completely ambiguous while the other was fully disambiguated by stereoscopic disparity (changing direction with a 40% probability). There was a temporal offset of 1.25 seconds between the presentation of the two cylinders causing each cylinder to be on the screen only during the other cylinder’s inter stimulus interval (Figure 5-6A). The presentation of these alternating cylinders thus had a residual true blank period of 250 milliseconds. Observers indicated the perceived direction of the near/front surface of both cylinders by pressing buttons on the keyboard.

#### *5.4.7 Procedure experiment 4: Relative vs. absolute depth.*

Two coaxial cylinders were presented on both sides of a central fixation cross, spatially separated by a gap of 0.5 degrees. The rightmost stimulus was always completely ambiguous while



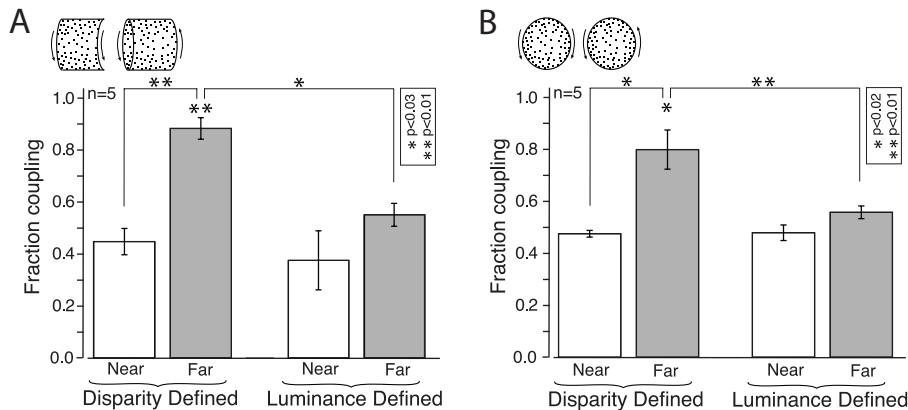
the other was fully disambiguated by stereoscopic disparity (changing direction with a 40% probability). The disparity defined cylinder could either be fully displayed or be restricted to its near or far side. The set of cylinders were defined to have a location in depth with their axis of rotation either one diameter closer to the observer than the plane of fixation or one diameter further away from the observer than the plane of fixation. Their size on the screen was maintained the same for both situations. In the plane of fixation we added a framework of three vertical and two horizontal grey bars ( $25.3 \text{ cd/m}^2$ ) with a width of 0.5 degrees to aid depth discrimination (see schematic representation in Figure 5-7). This addition caused some of the dots on the left and right sides of the cylinder to be either (partially) occluded by or on top of the this null-plane framework which, combined with the disparity information, resulted in a vivid percept of the cylinders being behind or in front of the plane of fixation. As in Experiment 1, observers reported the perceived direction of the near/front surface of the full, ambiguous stimulus by pressing a button on the keyboard. The 'half' stimulus had a 40% probability of changing its direction on consecutive presentations.

## 5.5 Results

### 5.5.1 Experiment 1: Information sharing in the near and far planes.

Our hypothesis for amodal spatial facilitation in perceptual coupling predicts that the difference in perceptual coupling between luminance and disparity depth cues results from the existence of lateral connections between neural populations involved in the representation of the two individual cylinders or spheres (Figure 5-2C). In particular, we argue that the principle of amodal completion of occluded objects suggests that these facilitatory lateral connections are only present in the far depth plane or in any case much stronger than in the near depth plane. This implies that previous assumptions about the necessity of both dot layers (or 'sides') of a kinetic depth cylinder for perceptual coupling (Freeman & Driver, 2006) may have been premature. It could very well be that one dot layer is enough to establish coupling as long as it explicitly constitutes the 'far half' of the cylinder.

In this experiment we test this hypothesis by using fully biased half cylinders and spheres that are defined by luminance or disparity to be either far or near sides of a kinetic depth stimulus. We included spheres here to investigate whether any possible coupling effect should be attributed solely to surface continuation, which could drive coupling between coaxial cylinders but not between spheres. The results convincingly demonstrate that perceptual coupling can occur between an ambiguous stimulus and a coaxial half stimulus as long as the latter is a disparity defined far side (Figure 5-4A for cylinders, T-test:  $p<0.001$ ; Figure 5-4B for spheres, T-test:  $p<0.02$ ) and because the effect is present for both cylinders and spheres it cannot be solely attributed to surface-continuation. Disparity defined near sides (T-test:  $p_{\text{cylinders}}=0.36$ ;  $p_{\text{spheres}}=0.12$ ), luminance defined far (T-test:  $p_{\text{cylinders}}=0.34$ ;  $p_{\text{spheres}}=0.50$ ) or near sides (T-test:  $p_{\text{cylinders}}=0.31$ ;  $p_{\text{spheres}}=0.09$ ) do not couple with an ambiguous cylinder. Furthermore, disparity defined far sides couple significantly better than disparity defined near sides (T-test:  $p_{\text{cylinders}}<0.01$ ;  $p_{\text{spheres}}<0.02$ ) or luminance defined far sides (T-test:  $p_{\text{cylinders}}<0.03$ ;  $p_{\text{spheres}}<0.01$ ). Luminance defined far sides appear to couple slightly better than luminance defined near sides but this difference was not significant (T-test:  $p_{\text{cylinders}}=0.30$ ;  $p_{\text{spheres}}=0.11$ ). It must however be noted that luminance cues on a single surface are not very effective. Even



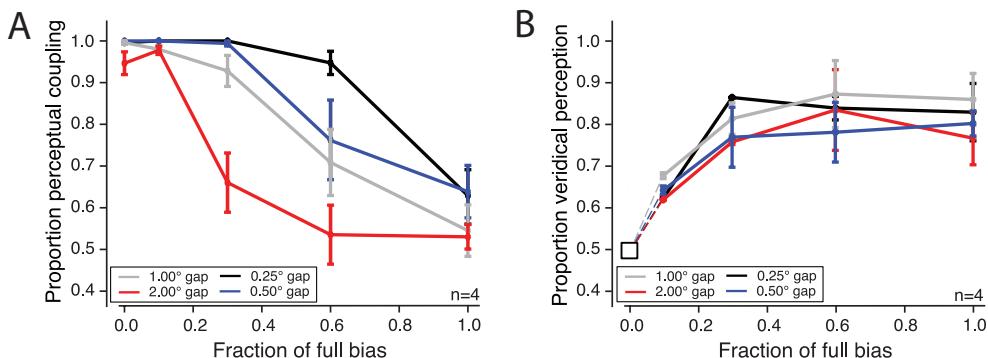
**FIGURE 5-4.** The fraction of perceptual coupling between 'halves' and ambiguous kinetic depth stimuli for five observers for cylinders (A) and spheres (B). 'Half' stimuli are defined to be the near or far sides of the full stimulus using either full luminance gradients or full disparity biases. For both types of stimuli, the only case in which the fraction of coupling is significantly larger than chance is when there is a disparity defined far side. In those cases there is also significantly more perceptual coupling than in disparity defined near sides or luminance defined far sides. Error bars represent S.E.M.

though a luminance gradient can define a convex or concave surface, the dots we used to define the concave backsides were very dim and the general tendency of observer's to perceive single surfaces as being the near side of a cylinder appears to dominate the luminance depth cue altogether.

The main conclusion of this experiment is the demonstration that perceptual coupling can occur between an ambiguous cylinder and a single surface as long as this single surface is a clearly defined cylinder backside.

### 5.5.2 Experiment 2: The spatial decay of perceptual coupling.

Our explanation of the perceptual coupling phenomenon proposes the existence of lateral connections that are responsible for information sharing between neural pools coding for spatially separated stimuli. It seems legitimate to think that the effectiveness of the information sharing mechanism will depend on the distance that needs to be bridged. In particular, one might expect that strong initial signals will be able to bridge larger distances between stimuli than weak ones. Experiment 2 tests this assumption by measuring the proportion of perceptual coupling as a function of dot luminance bias and gap-size between the cylinders. From the experiments in Appendix 2 (and previous work by Freeman & Driver, 2006; Grossmann & Dobbins, 2003), we know that perceptual coupling between a luminance biased cylinder and an ambiguous cylinder will cease to exist when the bias gets too large. Figure 5-5A demonstrates that with all gap-sizes used there is a near perfect coupling between two ambiguous cylinders and coupling at chance level with full luminance gradients. However, the moment of the drop in perceptual coupling depends not only on the strength of the luminance depth cue, but also on the distance between the cylinders (2-way ANOVA:  $F_{lum}(4,60)$



**FIGURE 5-5.** A) The influence of gap-size on perceptual coupling for four observers. The proportion of perceptual coupling is plotted against the strength of a luminance bias. The proportion of perceptual coupling decreases when luminance biases become too large. If the gap between the two cylinders increases the drop in perceptual coupling occurs at smaller luminance biases. B) The influence of luminance bias and gap-size on the proportion of trials in which the observers perceive the biased cylinder in accordance with the bias. This proportion increases fast with stronger biases but is not significantly influenced by gap-size. The open square at bias level zero is a theoretical point at chance level since there is no veridical percept here. Error bars in both plots represent S.E.M.

$= 58.59$ ,  $p_{\text{lum}} < 0.001$ ;  $F_{\text{gap}}(3,60) = 15.00$ ,  $p_{\text{gap}} < 0.001$ ). A significant interaction between gap-size and luminance bias ( $F_{\text{inter}}(12,60) = 3.19$ ,  $p_{\text{inter}} < 0.01$ ) further demonstrates that when the distance between the cylinders increases, the proportion of perceptual coupling starts to decrease at much smaller luminance biases. This suggests that spatial facilitation over larger distances needs the presence of stronger signals in the far depth plane.

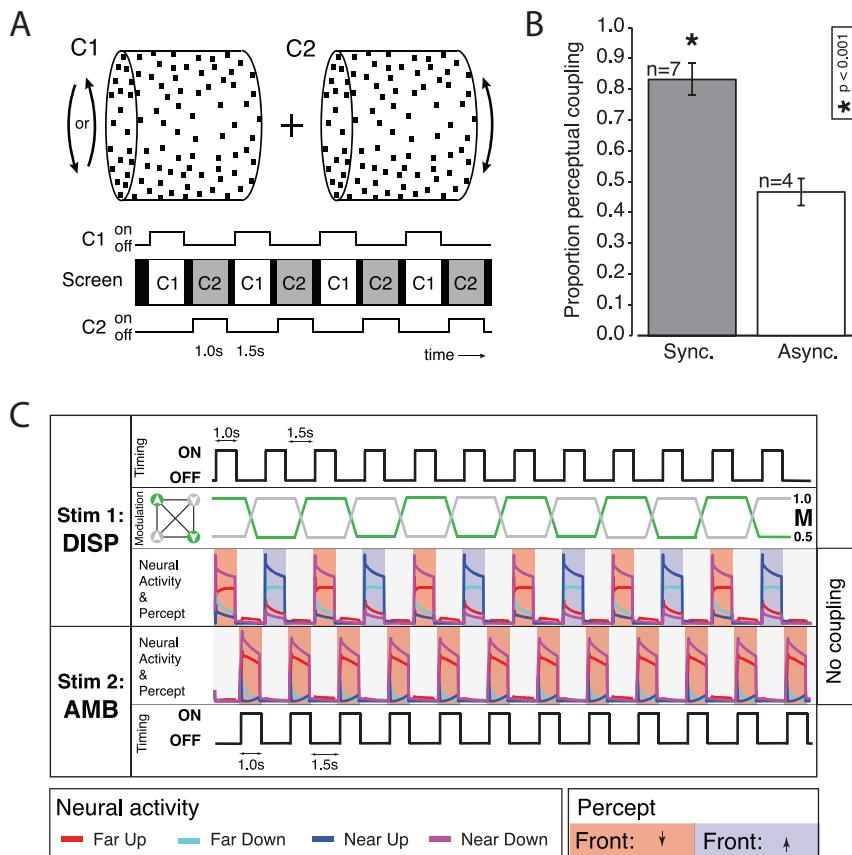
Our model predicts that a facilitatory signal from an ambiguous towards a weakly luminance biased cylinder can overcome the luminance bias (Figure 5-3D). If this phenomenon of ‘reversed coupling’ takes place, the biased cylinder will be perceived to couple with the ambiguous cylinder and rotate against its bias. The balance between the strength of the depth cue and the strength of the spatial facilitation determines whether this will happen. If the effect of facilitation indeed scales with the distance between stimuli we would thus expect that the proportion of trials in which a weakly biased cylinder is perceived veridically would be larger for smaller gap-sizes. Figure 5-5B plots the proportion of veridically perceived biased cylinders as a function of bias strength and gap-size. The effect of bias strength is highly significant (2-way ANOVA,  $p < 0.002$ ), but the effect of gap-size is not ( $p = 0.93$ ) nor is the interaction between bias and gap-size ( $p = 0.99$ ). The gap-size dependency of the ‘reversed coupling’ is however expected to be a relatively subtle effect and our rather noisy data lacks the appropriate resolution to make any strong statements about it.

### 5.5.3 Experiment 3: Asynchronous presentation.

Figure 5-6B demonstrates that while the proportion of perceptual coupling between disparity defined cylinders and ambiguous cylinders was high when they were presented simultaneously (data from Appendix 2), it is completely absent if the two cylinders are presented



with a temporal offset (all observers; T-test,  $p > 0.13$ ; group data,  $p = 0.50$ ). Here, perceptual coupling was defined as coupling between the disparity-defined cylinder and the subsequent ambiguous cylinder, since the proportion of veridical perception of the disparity defined cylinder was at ceiling level (average over observers was  $0.95 \pm 0.05$  standard deviation; not significantly different from 1.0 as indicated by a T-test,  $p = 0.47$ ). In our model, the spatial facilitation term acts on the fast h-dynamics representing local field activity (Equation 5-3) and consequently has little effect on the slower adaptation dynamics (Equation 5-4). Simulations with our model indeed reproduce the absence of perceptual coupling when the two stimuli are presented asynchronously (Figure 5-6C).



**FIGURE 5-6.** A) Temporal profile of the presentation of the two cylinders in our control experiment. The left cylinder (C1) was disambiguated by stereoscopic disparity; the right cylinder (C2) was ambiguous. Each cylinder was presented alone for 1.0 seconds separated by 1.5 seconds intervals during which the other cylinder was presented. B) Significant perceptual coupling with synchronous presentation (grey bar, data from Appendix 2) ceases to exist when the stimuli are presented asynchronously (white bar). Error bars represent S.E.M. C) Simulations with our model reproduce the lack of perceptual coupling with asynchronous presentation. The simulation was performed with the same parameters as in Figure 5-3B, only now the input to the two sets of neuronal populations was asynchronous.

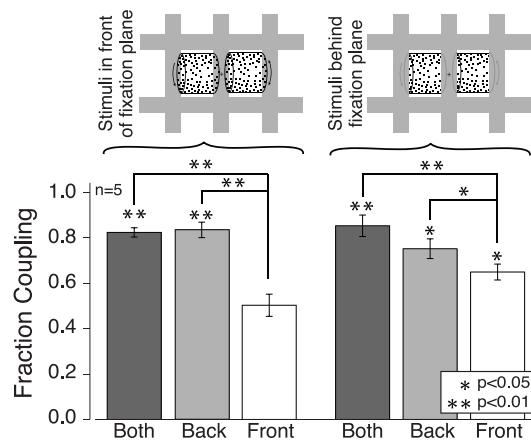


#### 5.5.4 Experiment 4: Relative vs. absolute depth.

This experiment aimed to unravel whether the distinction between near and far sides of a cylinder in perceptual coupling that is demonstrated with Experiment 1 relies on relative or absolute depth. It is important to realize that while our terminology of absolute and relative depth resembles the distinction between absolute and relative disparity (for reviews see Parker, 2007; Roe et al., 2007), they are in fact significantly different. The absolute depth of the potentially coupling surfaces is defined relative to the plane of fixation and can thus be regarded as an analog of absolute disparity, which describes the angular difference of retinal projections relative to the fovea. However, relative depth in our terminology indicates whether we are talking about a front side or backside of a cylinder and is something totally different from relative disparity, which is taken as the difference in absolute disparity between two points. A more direct analog of relative disparity would be the difference in depth between the two cylinders, but since the ambiguous stimulus is presented monocularly, it lacks an explicit location in depth and relative disparity cannot play a role. The results (Figure 5-7) demonstrate that perceptual coupling between ‘complete’ disparity-defined cylinders and ambiguous cylinders is maintained in both the near and far condition (T-test,  $p < 0.01$ ). For both these conditions the far cylinder sides alone also establish a significant fraction of coupling (T-test,  $p < 0.05$ ) that is not significantly different from the fraction that results from complete cylinders (T-test,  $p > 0.12$ ). Looking at the near sides of the disparity defined cylinders alone it becomes clear that significant coupling does not occur (T-test,  $p = 0.95$ ) when the stimuli are closer to the observer than the plane of fixation (matching the results from Experiment 1). However, when the stimuli are behind the plane of fixation the near sides can establish a significant fraction of perceptual coupling (T-test,  $p < 0.02$ ). This fraction is smaller than that for whole cylinders or far sides at the same depth location (T-test,  $p < 0.05$ ) but nevertheless present. The addition of a framework in the plane of fixation adds a minor depth cue to the display due to the partial occlusion of some of the dots at the edges of the far depth cylinders. Whereas, this manipulation greatly enhanced perceptual depth ordering, we believe that it is unlikely to have critically influenced our perceptual coupling results in any other way.

## 5.6 Discussion

The visual system uses spatial and temporal context to disambiguate local sensory information and construct a global conscious percept. If two ambiguous kinetic depth spheres or cylinders (Andersen & Bradley, 1998; Nawrot & Blake, 1989; Treue et al., 1991) are presented spatially separated but rotating about a common axis, their rotation directions couple and they switch directions simultaneously (Eby et al., 1989; Freeman & Driver, 2006; Grossmann & Dobbins, 2003). Apparently, even an ambiguous context can disambiguate a visual conflict. Studies investigating this perceptual coupling phenomenon have shown strong coupling both between multiple ambiguous stimuli and between disparity defined and ambiguous cylinders, but not between strong luminance biased and ambiguous cylinders (Freeman & Driver, 2006; Grossmann & Dobbins, 2003). This has led to the suggestion of a visibility constraint on the occurrence of perceptual coupling, stating that both sides of a context cylinder needs to be present to effectively couple rotation directions (Freeman & Driver, 2006). This visibility constraint in turn challenges the assumption that the two surfaces of a kinetic depth stimulus are



**FIGURE 5-7.** The roles of absolute and relative depth. The fraction of perceptual coupling between disparity defined ‘half’ and complete cylinders and ambiguous cylinders that were either closer to the observer than the plan of fixation (left) or further away than the plan of fixation (right). For the closer set of stimuli the results are comparable to those of experiment 1 (Figure 5-4). For the set of stimuli behind fixation the ‘near halves’ of cylinders (rightmost white bar) also cause a significant fraction of perceptual coupling. Error bars represent S.E.M.

represented in a co-dependent, mutually antagonistic way (Andersen & Bradley, 1998; Klink et al., 2008b; Li & Kingdom, 1999; Nawrot & Blake, 1991a; Treue et al., 1995). In the current study we consider an alternative explanation that is based on a general mechanism by which the brain could process partially occluded visual objects.

Our findings suggest that perceptual coupling *can* occur with single context surfaces but that its effectiveness (or lack thereof) depends on the neural mechanisms of the coupling process. The general extrapolation of spatially separated visual information into a globally consistent percept is known as spatial facilitation. Visual completion is a special case of spatial facilitation in which a single object or surface is perceived while it is only defined by spatially separated chunks of visual information. Completion is termed *modal* when illusory contours or surfaces are perceived in the foreground and amodal when it leads to the impression of an object or surface that is partially occluded or seen through an aperture (e.g. Anderson et al., 2002; Kanizsa, 1979)(Figure 5-1C). Even though there is a lively discussion about the extent to which modal and amodal facilitation share a common mechanism (e.g. Bakin et al., 2000; Hegdé et al., 2008; Murray et al., 2004; Rauschenberger et al., 2006; Weigelt et al., 2007), it is clear that they both involve the binding of spatially separated visual information. Ecological optics (Gibson, 1950) suggests that occlusion may be a more generally occurring feature than camouflage and amodal spatial binding of visual information (in far depth) should thus be more efficient than modal binding (in near depth). This idea is consistent with the finding that vernier shift discrimination is more accurate for amodally completed gratings than for modally completed ones (Anderson et al., 2002), more accurate face recognition in amodal vs. modal displays (Nakayama et al., 1989), and the demonstration of amodal, not modal,



continuation of visual motion behind an occluder (van der Smagt & Stoner, 2008). The model and experimental data that we present in this manuscript suggest that a similar amodal spatial facilitation mechanism may be responsible for the perceptual coupling and resulting disambiguation of kinetic depth stimuli.

In the light of our current results, the visibility constraint that was put forward by Freeman & Driver (2006) should be disregarded. Perceptual coupling of kinetic depth stimuli does not necessarily need two surfaces; a single far side surface suffices. In fact, the near side surfaces show very little if any coupling. When a strong luminance gradient is used to bias a kinetic depth cylinder towards a specific interpretation the amount of signal constituting the far side will be relatively small or absent, hence the failure of perceptual coupling. If an additional occluder is positioned between the observer and the cylinders perceptual coupling can also occur between near side surfaces suggesting that the near/far depth assignment results from a combination of absolute and relative depth that is particularly suitable to resolve occlusion in the visual scene. While our assumptions about the functional coupling mechanism are based on amodal visual completion it should be noted that perceptual coupling cannot be attributed to amodal surface completion (Fang & He, 2004). Whereas this explanation would be feasible for coaxial cylinders, it cannot explain why we find similar effects for coaxial spheres. The amodal information sharing is apparently occurring between pools of neurons tuned for combinations of depth and motion direction suggesting a more general mechanism by which neurons tuned to the same depth plane share sensory information. This idea is consistent with the recent finding that depth information propagates between surfaces only when these surfaces are located in the far depth plane (Georgeson et al., 2008).

Another interesting aspect of our experimental findings that is confirmed by model simulations is the existence of coupling against a stimulus bias (Appendix 2). Whereas the existence of a luminance gradient or binocular disparity is the only spatial context from which visual information can be inferred, there is additional *temporal* context in the presentation paradigm. Our use of the percept-choice paradigm not only has the advantage of being a sensitive measure to detect small imbalances in the activity of underlying neural populations (Noest et al., 2007), it is also a paradigm in which the inter-stimulus interval duration is crucial for the probability at which perception switches on consecutive trials. Because we use relatively long inter-stimulus intervals (1.5 seconds) we see an expected high level of perceptual stabilization when there is only a single stimulus (Klink et al., 2008a; Noest et al., 2007). In the two-stimulus condition with the biased stimulus stochastically changing direction there are thus two sources of contextual information leading to opposite conclusions. Whereas the *spatial* context signals percept changes, the *temporal* context signals percept stabilization. As can be seen in our results of Appendix 2, the relative strengths of the individual contexts ultimately determine conscious perception whereas perceptual coupling is high for all cases. This means that the information sharing mechanism we introduce is indeed bi-directional rather than only from the biased to the ambiguous cylinder.

The proposed connectivity between pools of neurons coding for similar sensory features at different spatial locations could be established in different ways. The most likely modes of connectivity would be 1) overlapping receptive fields of neurons in the two pools share information through their adaptation states, 2) direct single synapse connections between neurons in the two pools, or 3) an attenuating dilation of neural signal through ‘horizontal connections’ (Roelfsema, 2006) over a multitude of neurons covering the gap between stimuli (Ull-



man, 1979; van der Smagt & Stoner, 2008; Watanabe & Cole, 1995). Our use of the percept-choice paradigm allowed us to perform a specific control experiment (Experiment 3) that tested whether the information sharing mechanism occurs on the slow timescale of neuronal adaptation or on a fast timescale suggesting a direct activity-driven connection. The results demonstrate that perceptual coupling does not occur when two cylinders are presented with a temporal offset that causes them to be on the screen only during each other's interstimulus intervals. This suggests that the coupling mechanism does not occur on the slow adaptation timescale and should thus result from fast activity-driven lateral connections.

Whereas visual cortex is predominantly vertically organized in columns, horizontal connections with a length up to several millimeters have been demonstrated to connect similarly tuned clusters of neurons (Gilbert & Wiesel, 1979; Gilbert & Wiesel, 1983; Gilbert & Wiesel, 1989; Livingstone & Hubel, 1984; Malach et al., 1997; Martin & Whitteridge, 1984; Rockland & Lund, 1983). These connections are excitatory and the longer ones connect neurons with well-separated receptive fields (Ts'o et al., 1986). The number of horizontal connections decreases with increasing distance between connected clusters (Ts'o et al., 1986), which could explain why the proportion of perceptual coupling declines with increasing distance between the stimuli. Recently, lateral connections were discovered in the middle temporal area (MT) of the rhesus macaque (Ahmed et al., 2008). In MT, both depth and motion information are represented (Bradley et al., 1995; DeAngelis et al., 1998; Maunsell & Van Essen, 1983; Nadler et al., 2008) and responses are modulated by the three-dimensional structure of spatial context (Duncan et al., 2000). Lateral connections between similarly tuned clusters of neurons in MT would be an interesting candidate for our amodal spatial facilitation of kinetic depth stimuli. We are not aware of any existing studies looking into the specific distribution of lateral connections based on the depth selectivity of the neurons they are connecting, but our experiments suggest that if lateral connections are responsible for the perceptual coupling of SFM stimuli, the connections between 'far-tuned' neurons should be either stronger or more numerous than those between 'near-tuned' neurons.

The decrease in proportion of perceptual coupling with increasing distance between the stimuli is however also consistent with an attenuating dilation of neural signal over multiple cells 'covering the gap'. For orientation perception, cells in monkey primary visual cortex have been found that respond specifically to an invisible line segment only if it could be inferred from amodal completion, not when disparity information defined modal completion (Sugita, 1999). Cells in area MT or MST (middle superior temporal) could form such a bridging mechanism, either direct or via feedback from posterior parietal cortex where neural correlates of occluded motion have been demonstrated (Assad & Maunsell, 1995). In the absence of direct sensory stimulation these 'bridge-neurons' will not give rise to any percept, but their information-transporting role may cause adaptation that could perhaps be visualized using a subsequent test-stimulus on the location of the gap. A first hint that this might work can be found in a study by Fang & He (2004) that demonstrates a small (probably non-significant) adaptation effect in the non-stimulated gap between two co-rotating disparity defined cylinders (the yellow bars in their Figure 5-2B). Future experiments specifically designed to unravel the nature of the amodal information-sharing connectivity may be more successful in distinguishing between the two possible mechanisms.

Our last experiment demonstrated that the spatial facilitation mechanism is neither based purely on information about absolute depth (behind or in front of fixation), nor solely on the



relative depth of the surfaces constituting the cylinders (front side vs. backside), but rather on a mixture of the two. Whereas this seems to be an excellent functional approach to handle occlusion situations (like occlusion, spatial facilitation occurs at any depth plane that is not nearest to the observer), it complicates the physiological interpretation a little bit. The brain is known to exhibit neural substrates for both absolute and relative disparity (for reviews, see Parker, 2007; Roe et al., 2007), but the mechanisms by which these sources of depth information are combined are currently far from clear. As a result, our neural network model is likely to be a serious oversimplification of the actual process of spatial facilitation, but it provides a nice first handle in an attempt to understand how the brain uses spatially separated information in the perception of partially occluded objects. It should however be kept in mind that the proposed distinction in ‘far’ and ‘near’ tuned neurons should apparently be based on a mixture of absolute depth and depth relative to other parts of the visual scene.

In conclusion, our current findings suggest that the perceptual coupling of bistable stimuli reflects a more common mechanism by which the brain deals with occlusion. Facilitatory connections may exist between similarly tuned far depth neurons, establishing an information sharing mechanism that resolves local ambiguities by integrating spatially separated global information.

## 5.7 Acknowledgements

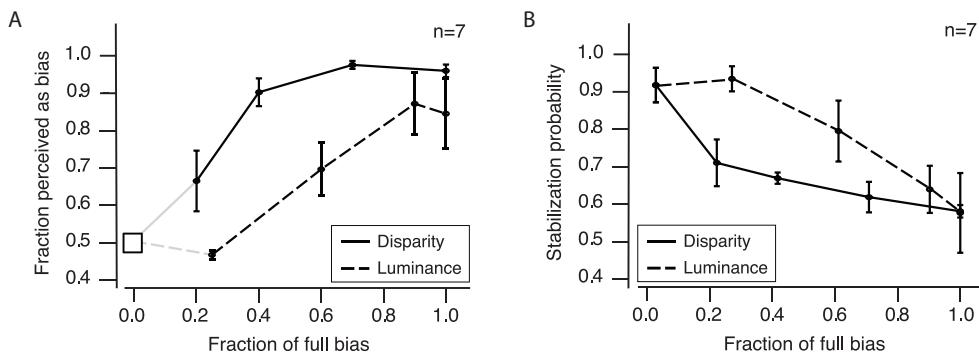
This work was supported by a VIDI grant from the Netherlands Organization for Scientific Research (NWO) and a High Potential grant from Utrecht University, both awarded to RvW. The authors thank Maarten van der Smagt for valuable discussions and comments on earlier versions of this manuscript.



## 5.7 Appendix 1

### Depth cues in a single kinetic depth cylinder

This experiment demonstrates whether there are any qualitative differences in the way that the perception of kinetic depth cylinders are influenced by either disparity or luminance defined depth cues. A single cylinder was presented at the center of the screen (see Methods for more details) and 7 observers (including 2 authors) indicated the perceived direction of the near/front surface of the cylinder by pressing a button on the keyboard. Any possible stimulus bias (disparity or luminance) had a 40% probability of changing its direction on consecutive presentations. The results are presented in Figure 5-A1 and demonstrate that the different depth cues have more or less similar qualitative effects. A quantitative comparison is difficult. Even if the two are plotted as ‘fraction of full bias’. First of all, the full bias for luminance depends on the monitor used for displaying the stimuli and secondly, it is unclear how luminance gradients would compare to ‘realistic disparity’. When stimuli are fully ambiguous (fraction of bias is zero), our experiments replicate previous findings of perceptual stabilization (Klink et al., 2008a; Leopold et al., 2002; Maier et al., 2003; Noest et al., 2007) (Figure 5-A1B). When a depth cue is introduced, it biases the stimulus towards one particular perceptual interpretation. When these depth cues are getting stronger, the stimuli are perceived consistent with the bias for a larger proportion of the trials (ANOVA:  $F_{\text{disp}}(4,30) = 26.14$ ,  $p_{\text{disp}} < 0.001$ ;  $F_{\text{lum}}(4,30) = 8.57$ ,  $p_{\text{lum}} < 0.001$ ) (Figure 5-A1A). Because the direction of the bias has an alternation probability of 40%, the proportion of perceptual stabilization decreases in accordance with the increasing veridicality (ANOVA:  $F_{\text{disp}}(4,30) = 10.57$ ,  $p_{\text{disp}} < 0.001$ ;  $F_{\text{lum}}(4,30) = 5.16$ ,  $p_{\text{lum}} < 0.003$ ) (Figure 5-A1B). Both depth cues reach high proportions of veridical perception and are thus effective determinants of perceptual interpretation.



**FIGURE 5-A1.** A) The fraction of trials that observers ( $n = 7$ ) perceived the cylinder to rotate in agreement with the bias as a function of bias strength for both disparity and luminance depth cues. The point indicated with the open square is a theoretical starting point since a stimulus cannot be perceived according to a bias if there is no bias. The effectiveness of both depth cues increases when the biases get larger and both reach high veridicality values. B) In the absence of stimulus biases we see clear perceptual stabilization. When the depth cues become stronger and observers start to perceive the stimulus in accordance with the bias more often (see A) stabilization probabilities naturally decrease since our stimulus biases changed direction with a probability of 40%. Error bars in both panels represent S.E.M.

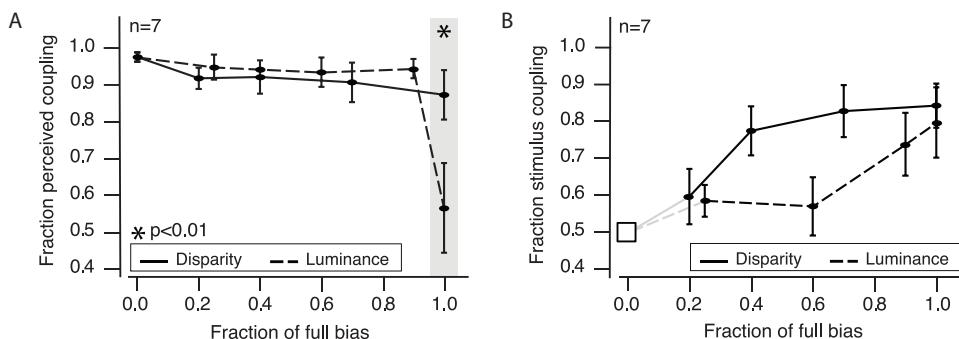


## 5.8 Appendix 2

### Perceptual coupling between kinetic depth cylinders

This experiment investigated the occurrence of perceptual coupling between two spatially separated kinetic depth cylinders rotating about a common axis. It is basically a repetition of the work of Freeman & Driver (2006) but we use a different experimental paradigm. In our percept-choice paradigm, stimuli are presented in sequences separated by short blank intervals. Freeman & Driver (2006) presented their stimuli for extended periods of 30 or 40 seconds. In the current experiment two coaxial cylinders were presented on each side of a fixation cross (Figure 5-1B). They were separated by a gap of 0.5 degrees, measured between their closest edges. The rightmost cylinder was always completely ambiguous whereas the left could have a disparity or luminance bias. Seven observers (including two authors) indicated the perceived direction of the near/front surface of both cylinders by pressing buttons on the keyboard. Any possible stimulus bias (disparity or luminance) again had a 40% probability of changing its direction on consecutive presentations.

Figure 5-A2A (next page) demonstrates the proportion of trials in which the two stimuli were perceived to rotate in the same direction as a function of the depth cue strength. Our findings confirm those of Freeman & Driver (2006). Strong coupling occurs for all values of disparity biases (solid line, no statistical differences within disparity cue strengths. ANOVA:  $F(4,30) = 0.66, p = 0.63$ ). For luminance depth cues there is also clear coupling, except for full luminance gradients (dotted line, ANOVA:  $F(4,30) = 8.10, p < 0.001$ ). For full depth cue biases the difference between luminance and disparity is highly significant (T-test,  $p < 0.01$ ) replicating previous findings by Freeman & Driver (2006). The effectiveness of the depth cues in determining perception increases when the cues get stronger (ANOVA:  $F_{\text{disp}}(4,30) = 6.19, p_{\text{disp}} < 0.001; F_{\text{lum}}(4,30) = 3.79, p_{\text{lum}} < 0.02$ ) and Figure 5-A2B demonstrates that when veridicality increases, the proportion of stimulus coupling also increases (ANOVA:  $F_{\text{disp}}(4,30) = 6.33, p_{\text{disp}} < 0.001; F_{\text{lum}}(4,30) = 3.22, p_{\text{lum}} < 0.03$ ). Stimulus coupling is defined as the fraction of the trials with perceptual coupling in which the rotation direction is consistent with the specified bias direction. Interestingly, for small depth biases the amount of perceptual coupling is very high (Figure 5-A2A) while the proportion of stimulus coupling remains relatively low (Figure 5-A2B) indicating a substantial proportion of trials in which the cylinders jointly rotated against the bias.



**FIGURE 5-A2.** **A)** The fraction of perceptual coupling as a function of bias strength for seven observers. High fractions of perceptual coupling are present for both luminance and disparity biases over almost the entire range of bias strengths. The significant difference between disparity and luminance cues occurs with full biases. Here there is still coupling between an ambiguous and a disparity defined cylinder but not between an ambiguous and a luminance defined cylinder (gray shaded area). **B)** The fraction of stimulus coupling as a function of bias strength. The fraction of stimulus coupling is the number of trials when stimuli were perceptually coupled and consistent with the bias direction divided by the total number of perceptually coupled trials. It is clear that with small biases the fraction of stimulus coupling is well below one, meaning that on a substantial number of trials the stimuli were perceptually coupled but rotated against the bias. When the bias gets stronger, the fraction of stimulus coupling also increases. Error bars in both panels represent S.E.M.s. The point indicated with an open square is a theoretical starting point in the absence of biases.

---

# Chapter 6

**Crossmodal duration perception involves perceptual grouping,  
temporal ventriloquism & variable internal clock rates**

*When you are courting a nice girl an hour seems like a second.  
When you sit on a red-hot cinder a second seems like an hour. That's relativity.*

**Albert Einstein**

**Published as**  
Klink, P.C., Montijn, J.S., & van Wezel, R.J.A. (2010). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism and variable internal clock rates. *Attention, Perception & Psychophysics*, [Epub ahead of print]



*We did not initially intend to study the perception of time. Instead, we started a project with the aim to investigate a crossmodal phenomenon that I had observed in my own perceptual experiences during the many hours of looking at ambiguously rotating sphere stimuli. I noticed that, for me, perceptual switches in rotation of such a visual stimulus were often accompanied by an illusory ‘inner sound’, consisting of rhythmic alternations between two differently pitched tones. I hypothesized that if my conscious visual perception was accompanied by these illusory sounds, than perhaps it would be possible to drive the visual perception of an ambiguous stimulus by actually playing these sounds while I was observing the stimulus. We performed many experiments to test this idea, and in the end we had to conclude that it worked very well....for me. Unfortunately, the other observers produced much more variable results and since none of them showed an effect that was comparable in strength to the one that was present in my data, we abandoned the idea altogether. However, still fascinated by the neural mechanisms of crossmodal interactions, we performed a short literature survey and came up with the idea to test whether the temporal ventriloquism effects, that were being used to describe subjective changes in temporal order, could also effect the subjective perception of event durations. This project-topic was quite a large deviation from the kind of studies that were usually carried out in our group and it was fun to read into a whole new field of literature. I later found out that when I was a subject in Raymond van Ee’s studies on the crossmodal effects of voluntary control in binocular rivalry (van Ee et al., 2009), my data again showed extraordinarily strong audiovisual interactions. Perhaps this is just a weird personal artifact, but it could also mean that there is a wide variety of moderate forms of synesthesia out there that goes largely unnoticed.*

---



## 6.1 Abstract

Here we investigate how audiovisual context affects perceived event duration with experiments where observers reported which of two stimuli they perceived as longer. Target events were visual and/or auditory and could be accompanied by non-targets in the other modality. Our results demonstrate that temporal information of irrelevant sounds is automatically used when the brain estimates visual durations, but that irrelevant visual information does not affect perceived auditory duration (Experiment 1). We further show that auditory influences on subjective visual durations only occur when the temporal characteristics of the stimuli promote perceptual grouping (Experiments 1 & 2). Placed in the context of scalar expectancy theory of time perception, our third and fourth experiments implicate that audiovisual context can both lead to changes in the rate of an internal clock and to temporal ventriloquism-like effects on perceived on- and offsets. Finally, intramodal grouping of auditory stimuli diminished any crossmodal effects, suggesting a strong preference for intramodal over crossmodal perceptual grouping (Experiment 5).

## 6.2 Introduction

Conscious perception involves the efficient integration of sensory information from different modalities. On the one hand, crossmodal integration can make perceptual experience richer and more accurate if the different modalities provide complimentary information about single objects or events. On the other hand, however, erroneous grouping of crossmodal information (e.g. grouping sources that do not belong together) can lead to distortions of conscious perception. To get around this problem, it is essential that there are efficient brain mechanisms of intra- and intermodal perceptual grouping that evaluate whether streams of sensory information should be combined into single perceptual constructs or not. While humans can be aware of some of these mechanisms, other mechanisms may play their prominent role outside of awareness (Repp & Penel, 2002). Research on the ‘unity assumption’ (i.e. the extent to which observers treat highly consistent sensory streams as belonging to a single event) has demonstrated that successful crossmodal integration of auditory and visual components in speech perception requires conscious perception of the two sensory inputs as belonging together (Vatakis & Spence, 2007). Such dependency was not found for audiovisual integration with non-speech stimuli (Vatakis & Spence, 2008). Even within single modalities subconscious perceptual grouping mechanisms play an important role, as the global perceptual organization of spatially or temporally separated ‘chunks’ of sensory information can have distinct effects on ‘local’ perception (e.g. Klink et al., 2009; Watanabe et al., 2001).

In multimodal integration, the brain typically relies more heavily on the modality that carries the most reliable information (Alais & Burr, 2004; Burr & Morrone, 2006; Ernst & Bülthoff, 2004; Recanzone, 2003; Wada et al., 2003; Walker & Scott, 1981; Welch & Warren, 1980; Witten & Knudsen, 2005). The assignment of reliability can be based on intrinsic properties of individual sensory systems, or on the signal-to-noise ratio of the available sensory input. The visual system, for example, has a higher spatial resolution than the auditory system (Witten & Knudsen, 2005). Thus, when visual and auditory information about the location of a single object in space are slightly divergent, the perceived location of the audiovisual object will be closer to the actual visual location than to the actual auditory location (Alais & Burr, 2004;



Welch & Warren, 1980; Witten & Knudsen, 2005). Such an ‘illusory’ perceived location is the basis of every successful ventriloquist performance. For the temporal aspects of perception, the auditory system is usually more reliable and thus more dominant, than the visual system (Bertelson & Aschersleben, 2003; Freeman & Driver, 2008; Getzmann, 2007; Guttman et al., 2005; Morein-Zamir et al., 2003; Repp & Penel, 2002). This is strikingly demonstrated when a single light flash is perceived as a sequence of multiple flashes when it is accompanied by a sequence of multiple auditory tones (Shams et al., 2002).

The perception of time or event duration is one specific case where conscious perception often deviates from the physical stimulus characteristics (Eagleman, 2008). Since time is a crucial component of many perceptual and cognitive mechanisms, it may be surprising that the subjective experience of the amount of time passing is distorted in many ways, such as by making saccades (Majj et al., 2009; Morrone et al., 2005; Yarrow et al., 2001) or voluntary actions (Park et al., 2003), by the emotional state of the observer (Angrilli et al., 1997), or stimulus properties such as magnitude (Xuan et al., 2007), dynamics (Kanai et al., 2006; Kanai & Watanabe, 2006) or repeated presentation (Pariyadath & Eagleman, 2008; Rose & Summers, 1995). Moreover, if temporal sensory information about duration is simultaneously present in multiple modalities, crossmodal integration can also cause distortions of subjective time perception (e.g. Chen & Yeh, 2009; van Wassenhove et al., 2008). For example, it is known that when sounds and light flashes have equal physical durations, the sounds are subjectively perceived as longer than the light flashes (Walker & Scott, 1981; Wearden et al., 1998). Furthermore, when auditory and visual stimuli of equal physical duration are presented simultaneously, the auditory system dominates the visual system and causes the durations of visual stimuli to be perceived as longer than they physically are (Burr et al., 2009; Chen & Yeh, 2009; Donovan et al., 2004; Walker & Scott, 1981).

Time perception mechanisms are classically explained with (variants of) the Scalar Expectancy Theory (SET) (Gibbon, 1977; Gibbon et al., 1984). SET proposes an internal clock mechanism that contains a pacemaker emitting pulses at a certain rate. During an event, a mode switch closes and allows for emitted pulses to be collected into an accumulator. The number of pulses in the accumulator at the end of the timed event is compared against a ‘reference time’ from memory. This comparison determines the perceived duration in a linear fashion: more accumulated pulses means longer perceptual durations. Whereas SET offers explanations for many aspects of time perception and distortion, it remains unclear how duration information from multiple modalities is integrated to allow a crossmodal estimation of event durations.

In general, the perceived duration of an event can directly be influenced by a change in pacemaker rate, a change in mode switch open/close dynamics, or distortions in memory storage and retrieval (Penney et al., 2000). Within the SET framework, the difference in perceived duration of equally long visual and auditory stimulus durations has been attributed to modality specific pacemaker rates for visual and auditory time (Wearden et al., 1998). Additionally, the dilation of subjective visual stimulus durations by simultaneously presented auditory stimuli has been explained by changes in pacemaker rate and not in mode switch latency (Chen & Yeh, 2009). Using a duration-bisection procedure, it has also been demonstrated that distortions in the memory-stage of SET can occur when a current sensory duration is compared against a previously trained reference duration that is stored in memory (Penney et al., 2000). In this paradigm, observers are trained to discriminate between a short and long dura-



tion signal (both labeled anchor durations). In a subsequent test phase they judge whether the duration of novel stimuli are closer to the short or to the long anchor duration. If both auditory and visual anchor durations have to be simultaneously kept in memory, a ‘memory-mixing’ effect occurs: the subjectively long auditory anchor duration and the subjectively short visual anchor duration mix into an intermediate reference duration that is perceived as shorter than the auditory anchor, but longer than the visual anchor of equal physical duration (Penney et al., 2000).

While some authors have attributed a difference in perceived internal clock rate to an attentional effect at the level of the mode switch (Penney et al., 2005), most have concluded that distortions of subjective time duration do not result from a change in mode-switch dynamics but rather from a change in the rate of the internal clock (Chen & Yeh, 2009; Penton-Voak et al., 1996; Wearden et al., 1998). However, since these studies all used auditory and visual stimuli with the same physical on- and offset moments it cannot be excluded that mode switch dynamics will play a more prominent role in crossmodal time perception when the on- and offsets are not the same. On the contrary, studies showing that the perceived temporal order of multiple visual stimuli can be influenced by the presence of irrelevant sounds (a phenomenon termed ‘temporal ventriloquism’) (Bertelson & Aschersleben, 2003; Getzmann, 2007; Morein-Zamir et al., 2003), suggesting that audiovisual integration may also distort the perceived on- and offset moment of visual events. One way by which temporal ventriloquism might play a role in the perceived duration of a visual event is that it shifts the subjective on- and offset of a visual event toward the on- and offset of an accompanying auditory stimulus. If these shifted subjective visual on- and offsets determine the moment at which the mode-switch closes and opens, they could very well modulate the subjective duration of a visual event without changing the rate of the internal clock. Alternatively, the mode-switch closing and opening could be determined by the physical rather than by subjective on- and offsets. In such a scenario, performance on a visual duration discrimination task should be immune to temporal ventriloquism-like effects.

The experiments presented here provide evidence for the idea that *both* the ‘rate of the internal clock’ and the perceived on- and offset of a visual target stimulus are modulated by crossmodal interactions. Below we discuss a series of human psychophysical experiments on audiovisual duration perception that exploit a two-alternative-forced-choice, prospective method of duration discrimination (i.e. observers knew they would report which of two stimuli had a longer duration). In order to investigate both the hypothesized effects of temporal ventriloquism and the previously demonstrated changes in internal clock rate, we presented auditory and visual stimuli both with and without differences in their physical on- and offsets. We start out by testing the hypothesis that an irrelevant auditory stimulus influences the perceived duration of a visual target, but that irrelevant visual stimuli do not affect the perceived duration of an auditory target (Experiment 1). While such an asymmetry has been shown with different experimental approaches (Bruns & Getzmann, 2008; Chen & Yeh, 2009), it has not yet been shown with the experimental paradigm we use throughout this study. We then continue to test the hypothesis that for any such crossmodal effect to occur the onsets and offsets of the auditory and visual stimuli need to be temporally close enough to evoke some kind of subconscious binding (Experiment 2).

The possible role of temporal ventriloquism-like effects is explored in more detail in Experiment 3, where the temporal differences between the on- and offsets of the target and non-



target stimuli in the different modalities are systematically varied. Experiment 4 sets out to determine whether the auditory dominance over visual duration discrimination is reflected in a complete shift of the time perception system from using visual temporal information to using auditory temporal information, or that some weighted average is used that relies more heavily on auditory than on visual information. Our fifth and final experiment controls for an important possible confound in all the other experiments. Any crossmodal effect on the reported perceived durations might either be due to a truly altered experience of subjective durations in the target modality caused by crossmodal interactions within the time perception system, but it could also represent a behavioral shift towards reporting perceived durations from the irrelevant non-target modality instead. Using stimulus conditions in which intra- and cross-modal grouping of stimulus elements are to be expected, we demonstrate that subconscious crossmodal grouping of auditory and visual stimuli is necessary for the crossmodal effects on duration discrimination to occur.

Ultimately, our interpretation of the results is summarized in a schematic SET-model for crossmodal duration perception (Figure 6-6). In the first stage of the model, stimulus features are perceptually grouped within and/or across modalities. The second stage incorporates a multimodal version of the SET that captures temporal ventriloquism effects in the timing of the mode-switch and accounts for additional crossmodal influences with modality dependent internal clock rates.

### 6.3 General methods

The basic experimental setup is the same for all experiments. The differences between the experiments predominantly concern the precise timing of stimuli and the kind of perceptual judgment observers were asked to report. Those specific details are described in the method sections of the individual experiments.

All stimuli were generated on a Macintosh computer running Matlab (Mathworks, Natick, MA) with the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997) and displayed on a 22" CRT monitor with a resolution 1280×1024 and a refresh rate of 100 Hz. Observers used a head and chinrest and viewed the screen from a distance of 100 cm. In all experiments, observers performed a two-alternative forced choice task; they reported which of two target stimuli they perceived to have a longer duration. The modality of the target stimuli was indicated to the observers on the screen prior to stimulus. Visual targets were white circles or squares with a diameter of ~3 degrees visual angle and an equal surface area to keep total luminance constant. The luminance of the visual targets was 70 cd/m<sup>2</sup> and they were presented on a grey background with a luminance of 12 cd/m<sup>2</sup>. Auditory targets were pure tones of 200 Hz, played to the observers through a set of AKG K512 stereo headphones at a SPL of ~64 dB (measured at one of the headphone speakers with a Temna 72-860 sound level meter). All participants had normal or corrected to normal visual acuity and no known auditory difficulties. All experiments contained randomly interleaved catch-trials in which large duration differences (400 ms) were present in the target modality, while non-targets were of equal duration. Adequate performance on catch trials is an indication that an observer is performing the tasks correctly. Poor performance on catch trials (less than 75% correct) was reason for exclusion of an observer from the data analysis. For this reason, six observers were excluded



from Experiment 3 and two from Experiment 5. The number of observers that is mentioned in the method sections of the individual experiments indicates the number of observers that performed adequately on catch-trials and whose data were included in the analysis. All observers were students or scientific staff of Utrecht University's departments of Psychology and Biology, ranging in age between 19 and 35 years.

## 6.4 Experiment 1

### Asymmetric audiovisual distortions in duration perception

This experiment investigates whether crossmodal influences between auditory and visual duration perception can be demonstrated with our experimental paradigm. If such effects are found, this experiment will further reveal whether they depend on the temporal properties of the non-target stimuli and/or the temporal relation between the target and non-target stimuli.

#### 6.4.1 Method

Ten observers (ranging in age between 21 and 30 years, 5 males and 5 females, 2 authors) participated in this experiment. They reported which of two target stimuli they perceived to have a longer duration. Prior to presentation, observers were notified whether the target stimuli would be visual or auditory. Target stimuli were always accompanied by non-target stimuli in the other modality. Before the actual experiment, all participants performed a staircase procedure to determine their individual just noticeable differences (JND) for visual and auditory stimuli with a base duration of 500 ms. In this procedure they essentially performed the same task as in the main task, i.e. comparing the duration of two stimuli, but here target stimuli were never accompanied by non-target stimuli in another modality. The staircase procedure used the Psychtoolbox' Quest algorithm (Watson & Pelli, 1983) and consisted of 25 trials converging on 82% correct, determining the minimal duration difference an observer can reliably detect at a base duration of 500 ms. The staircase was performed three times for both modalities and the averages for each modality were taken as the individual observer's JNDs. The observer-specific JNDs were then used in the main experiment. The average JND over all observers for auditory stimuli was 78.9 ms ( $\pm 8.9$  ms S.E.M.) and for visual stimuli 117.7 ms ( $\pm 8.7$  ms S.E.M.).

Stimuli in the target modality had a duration of  $500 \text{ ms} \pm \text{JND}/2$  and the order in which the long and short stimuli were presented was counterbalanced. Stimuli in the non-target modality could either both be 500 ms ( $\delta t_{\text{non-target}} = 0$ ) or 400 ms and 600 ms ( $\delta t_{\text{non-target}} = 200$  ms, see Figure 6-1A). When there was a duration difference between the non-target stimuli, the short non-target stimulus was always paired with the long target stimulus and the long non-target stimulus with the short target stimulus. The temporal midpoints of the target and non-target stimuli could either be aligned (marked 'Center-Aligned') or shifted  $\pm 250$  ms relative to each ('Center-Shifted'). We aligned stimuli by their midpoint since we expected temporal ventriloquism to play a role in the perceived on- and offsets of multimodal stimuli. Alignment by midpoints has the benefit of equal temporal deviations between the onsets and offsets of target and non-target stimuli. The interstimulus interval between target stimuli, defined as the temporal separation between their midpoints, was 1500 ms with a randomly assigned jitter



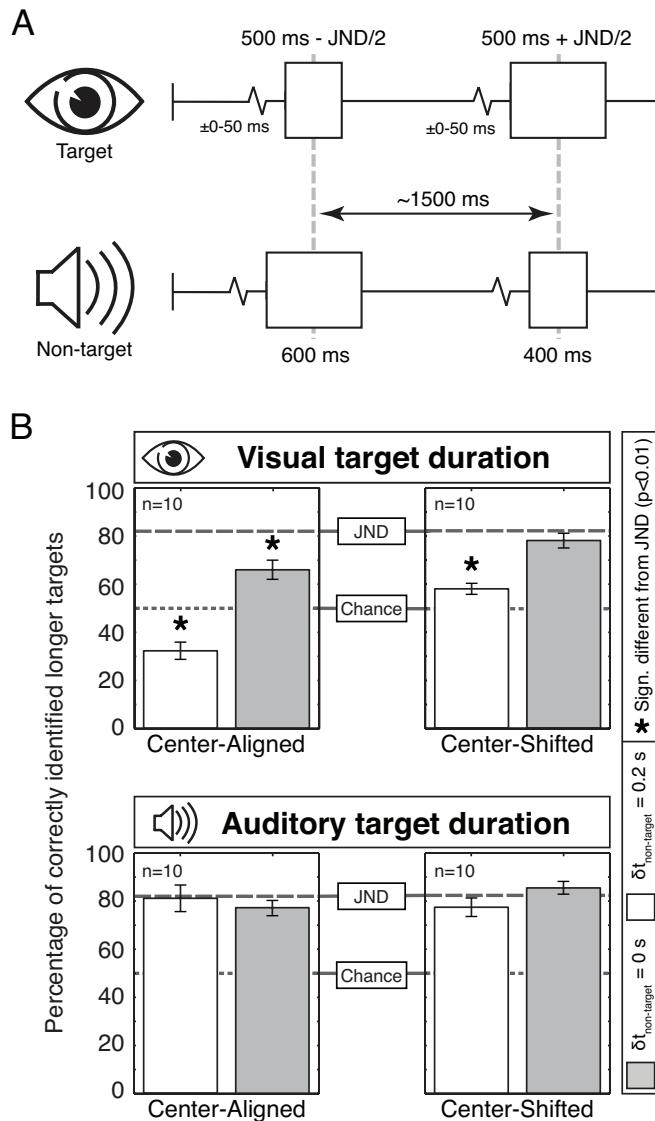
between -50 and +50 ms (Figure 6-1A). Experimental conditions were presented in blocks of 40 repetitions. Individual trials started when the observer pressed a designated key on a standard keyboard. The order of these blocks was counterbalanced.

The first five observers (including the two authors) were asked to indicate whether they perceived either the first or second stimulus to have the longest duration. Even though observers were instructed to fixate a dot on the screen during the entire duration of the experiment, this specific instruction would in principle allow them to completely ignore non-target visual stimuli by temporarily closing their eyes. None of the observers admitted to adopting such a strategy, but to avoid the possibility altogether we modified the instruction and asked a second group of five observers to report whether they perceived either (the sound accompanied with) the circle or (the sound with) the square to have a longer duration. Since the order in which the square and circle appeared was pseudorandom, this instruction forced observers to keep looking at the screen. The results from both observer-groups were highly similar and therefore combined in the group analysis.

#### 6.4.2 Results and discussion

Figure 6-1B displays the percentage of trials in which observers correctly identified the longer target stimulus for each experimental condition. If non-target stimuli have no effect on duration discrimination performance with the target stimuli, observers are expected to perform at a level of 82% correct, which was the threshold-level of the staircase procedure that determined their individual JNDs. The first thing that becomes clear from the results in Figure 6-1B is that there appears to be an asymmetry in the extent to which visual (top panel) and auditory duration discrimination performances (bottom panel) are influenced by non-target stimuli in the other modality. A Within-Subjects ANOVA (factors: Target modality visual vs. auditory, Center-Aligned vs. Center-Shifted presentation, and Non-target duration different vs. equal) confirmed that there is indeed a significant difference between the target modalities ( $F(1,37) = 119.44, p < 0.001$ ). It also returned significant differences between the cases where the non-target stimuli had the same duration (grey bars in Figure 6-1B) and the cases where the non-target stimuli had a duration difference opposite to that of the target stimuli (white bars in Figure 6-1B) ( $F(1,37) = 40.18, p < 0.001$ ).

Center-Aligned vs. Center-Shifted presentation of the target and non-target also had a significant effect on performance ( $F(1,37) = 23.17, p < 0.01$ ), but because the interaction between center alignment and target modality was significant as well ( $F(1,37) = 15.68, p < 0.001$ ), we re-analyzed the results for the two target modalities separately (Within-Subjects ANOVA with factors Center-Aligned vs. Center-Shifted presentation, and Non-target duration different vs. equal). This analysis revealed that when target stimuli are visual, there are both significant effects of Center-Aligned vs. Center-Shifted presentation ( $F(1,9) = 27.74, p < 0.001$ ) and of difference vs. no difference in auditory non-target durations ( $F(1,9) = 160.75, p < 0.001$ ). The interaction between the two was not significant but did show a trend ( $F(1,9) = 4.00, p = 0.08$ ) suggesting that the effect of auditory non-target duration differences is slightly larger when the temporal midpoints of visual targets and auditory non-targets are aligned. When the targets were auditory, neither of these contrasts reached significance (Center-Aligned/Center-Shifted:  $F(1,9) = 0.01, p = 0.92$ ; Difference/No difference in non-target:  $F(1,9) = 0.82, p = 0.39$ ). Thus, crossmodal distortions in duration perception only occur for visual targets in auditory context, not the other way around.



**FIGURE 6-1.** Set-up and results of experiment 1. **A)** An example of the experimental setup for Experiment 1. Here the targets are visual and observers report which of two visual targets (rectangles in top panel) they perceive to have a longer duration. In this example the visual targets are paired with auditory non-target stimuli that are aligned to the visual targets by their temporal ‘midpoints’ and have a duration difference in the opposite direction of the duration difference in the visual targets. **B)** The results of Experiment 1 demonstrate asymmetric crossmodal influences in duration perception. The percentage of correctly identified longer target stimuli is plotted, split by target modality. Visual targets are shown in the top panel and auditory targets in the bottom panel. Within a modality a distinction is made between cases where the target and non-target stimuli were Center-Aligned (left panels) or Center-Shifted (right panels) with respect to their temporal midpoints. Grey bars represent cases in which there was no duration difference in the non-target modality, white bars represent cases with a duration difference in the non-target modality, and error bars represent S.E.M.’s.



All different visual target conditions were individually analyzed revealing that for the Center-Aligned cases, a duration difference in the auditory non-targets reduced performance by 61.0% ( $\pm 4.4$  S.E.M.), bringing it significantly below the individual measured thresholds (T-test:  $t(9) = -13.92$ ,  $p < 0.001$ ) and even below chance ( $t(9) = -5.01$ ,  $p < 0.01$ ). This implies a strong bias to report the short visual target stimulus (paired with the long auditory non-target stimulus) as being subjectively longer than the long visual target stimulus (paired with the short auditory non-target stimulus). Thus, the presence of sounds does not merely impair performance on a visual duration discrimination task, but it actually modulates the subjective visual duration. If there is no difference in auditory non-target duration, performance on visual duration discrimination is still impaired by 20.1% ( $\pm 4.9\%$  S.E.M.), which brings it significantly below the JND-threshold ( $t(9) = -4.15$ ,  $p < 0.01$ ) but keeps it significantly higher than chance ( $t(9) = 3.90$ ,  $p < 0.01$ ). In the Center-Shifted cases, a significantly reduced performance of 29.9% ( $\pm 2.8\%$  S.E.M.) was observed when there was an auditory non-target duration difference ( $t(9) = -10.81$ ,  $p < 0.001$ ), but the 5.5% ( $\pm 3.7\%$  S.E.M.) impairment when there was no duration difference between auditory non-target stimuli was not significant ( $t(9) = -1.50$ ,  $p = 0.17$ ). In the latter case, performance was still significantly better than chance ( $t(9) = 9.15$ ,  $p < 0.001$ ) suggesting that no mentionable crossmodal effects took place. None of the individual cases for auditory targets were statistically different from the 82% threshold level (Effect-size<sub>aligned/diff</sub> =  $-1.8\% \pm 6.7\%$ ,  $t_{aligned/diff}(9) = -0.27$ ,  $p_{aligned/diff} = 0.79$ ; Effect-size<sub>aligned/no\_diff</sub> =  $-1.2\% \pm 3.1\%$ ,  $t_{aligned/no\_diff}(4) = -0.39$ ,  $p_{aligned/no\_diff} = 0.71$ ; Effect-size<sub>shifted/diff</sub> =  $-6.1\% \pm 4.7\%$ ,  $t_{shifted/diff}(9) = -1.31$ ,  $p_{shifted/diff} = 0.22$ ; Effect-size<sub>shifted/no\_diff</sub> =  $3.7\% \pm 3.3\%$ ,  $t_{shifted/no\_diff}(9) = 1.12$ ,  $p_{shifted/no\_diff} = 0.29$ ).

We conclude that visual duration discrimination performance is influenced by the presence of auditory non-target stimuli, but that the extent of impairment depends critically on both the relative on- and offsets and the duration differences between the target and non-target stimuli in both modalities. If visual target and auditory non-target stimuli are center-aligned and the auditory non-target stimuli have a duration difference opposite to that of the visual targets, performance on the visual duration discrimination task was impaired most. When the stimuli are either Center-Shifted with auditory non-target duration differences or Center-Aligned without auditory non-target duration differences, performance is less impaired. Finally, there is no significant impairment of visual duration discrimination performance when stimuli are both Center-Shifted and auditory non-target duration differences are absent.

We suspect that while the presence or absence of a duration difference in the auditory non-targets may influence the internal representation of a visual duration through cross-modal interactions, the effect of temporal alignment and its consequential difference in the on- and offsets of visual and auditory stimuli will predominantly act upon the likeliness of crossmodal binding to occur and promote the crossmodal interactions in the time perception system. The next experiment tests the hypothesis that temporal proximity of the on- and offsets of the target and non-target stimuli in the different modalities is indeed required for crossmodal effects to occur in our duration discrimination task.



## 6.5 Experiment 2

### The need for crossmodal on- and offset proximity

The asymmetric crossmodal effects demonstrated in Experiment 1 raise several questions with regard to the critical stimulus aspects that evoke the distortions of subjective stimulus duration. Do additional non-target sounds influence all subjective duration judgments or only the duration judgments for visual stimuli? Furthermore, do the differences between the center-aligned and center-shifted conditions result from a mere difference in the amount of temporal overlap between the visual targets and auditory non-targets or does the actual timing of the on- and offsets of the stimuli play the crucial role we predicted? In this second experiment we address these questions by combining auditory and visual duration judgments with contextual background sounds that, if present, start well before and end well after target presentation resulting in equal amounts of crossmodal temporal overlap, but very large crossmodal on- and offset differences.

#### 6.5.1 Method

The same ten observers that participated in Experiment 1 also performed in this experiment. They reported which of two target stimuli they perceived to have a longer duration. The target stimuli could either be visual (circles or squares) or auditory (pure tones at 200 Hz). Non-target background stimuli were always auditory (pure tones at 100 Hz) with significantly longer durations than the target stimuli. Prior to Experiment 1, all participants had performed a staircase experiment to determine their individual just noticeable differences (JND) for visual and auditory stimuli with a base duration of 500 ms. These individual JNDs were also used in this experiment. Stimuli in the target modality had a duration of  $500 \text{ ms} \pm \text{JND}/2$  and the order of the long and short target stimuli were counterbalanced. Auditory background stimuli were 2500 ms in duration and their midpoint was temporally aligned with the midpoint of the target stimuli, resulting in an on- and offset difference of minimally 1000 ms -  $\text{JND}/4$ , which should be more than enough to prevent audiovisual integration (Jaekl & Harris, 2007). Auditory background sounds, if present, could be played in conjunction with either the short target stimulus only, the long target stimulus only, or both target stimuli (Figure 6-2A). The interstimulus interval between target stimuli, defined as the temporal separation between their midpoints, was 2500 ms with a randomly assigned jitter between -50 and +50 ms. Observers were familiarized with the stimuli and task before the experiment started and all of them indicated that auditory target stimuli could be easily distinguished from the auditory background sound. Each combination of target/non-target stimuli was presented 40 times, resulting in 320 trials that were distributed over 4 blocks of trials in a counterbalanced way to allow observers to have short breaks. Individual trials started when observers pressed a designated key on a standard keyboard.

#### 6.5.2 Results and discussion

Figure 6-2B plots the percentage of correctly identified longer targets for each experimental condition. None of the experimental conditions appear to have any effect on the observers' performance. A statistical analysis of the data (within-subjects ANOVA) confirms that there is no significant difference between visual targets (left panel) and auditory targets (right



panel)( $F(1,27) = 0.54$ ,  $p = 0.48$ ), nor is there an effect of non-target stimulus timing ( $F(3,27) = 1.50$ ,  $p = 0.24$ ). The only two distinctions between the auditory non-target stimuli that did have an effect in Experiment 1 and the auditory background stimuli that had no effect in Experiment 2 are their frequency and the difference in on- and offset timing between target stimuli and non-target sounds. It seems highly unlikely that auditory influences on perceived visual duration would be crucially different for sounds of 100 Hz and sounds of 200 Hz, or that an observer's capability of blocking the auditory influence with attention would depend on such a minor frequency difference. Comparing the results from Experiment 1 and Experiment 2, we conclude that an auditory non-target only affects the perceived duration of a visual target when the on- and offsets of the target and non-target stimuli are close enough in time. Such temporal proximity of target and non-target on- and offsets might merely allow the subconscious binding of auditory and visual stimuli, thereby promoting crossmodal distortions of duration perception. Alternatively, a temporal ventriloquism-like effect on the perceived on- and offsets of the target stimuli may also play a role in the construction of perceived event durations. This possibility will be explored in the next experiment.

## 6.6 Experiment 3

### Auditory distortions of visual duration: A parametric approach

The first two experiments demonstrated that duration judgments about a visual target stimulus are distorted by the presence of auditory non-target stimuli with on- and offsets that are close in time to the on- and offsets of the visual target stimulus. It remains unclear, however, whether these distortions occurred because the sounds increased the perceived duration of the short visual stimulus, decreased the perceived duration of the long visual stimulus, or both. Furthermore, it is unclear whether the temporal proximity of these on- and offsets merely promote crossmodal interactions, or whether the temporal difference in on- and offset has a more systematic effect (as we would expect from temporal ventriloquism). Our third experiment employs a parametric method to investigate the influence of auditory stimuli on visual duration judgments. Two visual stimuli of equal duration had to be compared, while one of the two was accompanied with an auditory non-target stimulus that could either have a shorter or a longer duration than the visual stimulus it was paired with.

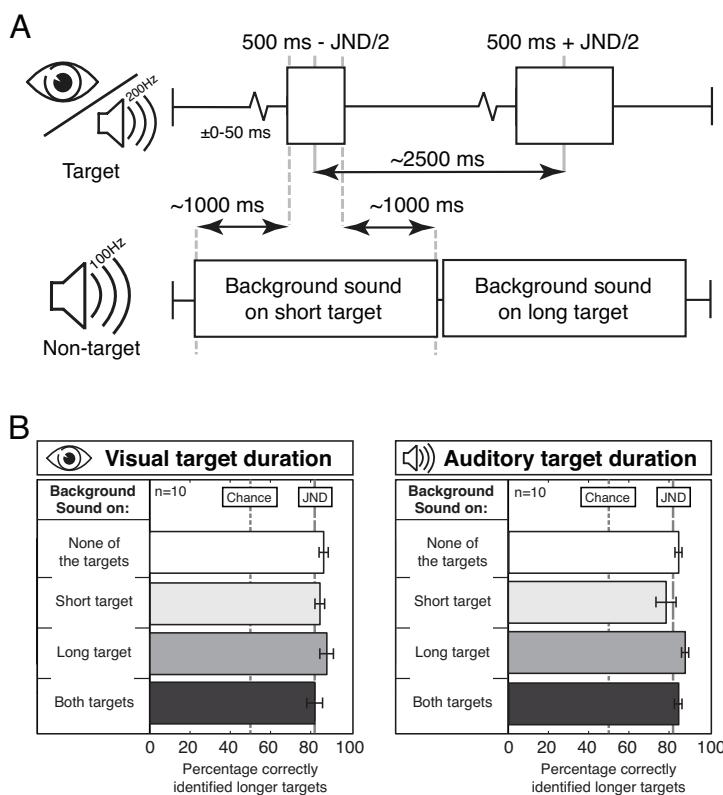
#### 6.6.1 Method

Twelve observers (ranging in age between 19 and 35 years, 7 males and 5 females) participated in this experiment. Two of these observers also participated in Experiments 1 and 2. They reported which of two visual target stimuli they perceived to have a longer duration. One of the target stimuli (which could be the first or second, randomly assigned and counterbalanced) was paired with an auditory non-target stimulus. The visual target stimuli both had a duration of 500 ms while the auditory non-target stimuli had pseudo-randomly assigned durations ranging from 150 ms to 850 ms in 50 ms steps (Figure 6-3A). This resulted in duration differences between the visual target stimuli and auditory non-target stimuli ranging from -350 ms to +350 ms. Target and non-target stimuli were temporally aligned by the midpoint of their duration. The interstimulus interval between target stimuli, defined as the temporal separa-



tion between their midpoints, was 1500 ms with a randomly assigned jitter between -50 and +50 ms. Each stimulus combination was repeated sixteen times in pseudo-random, counter-balanced order (yielding a total of 240 trials for each observer). Individual trials started when observers pressed a designated key on a standard keyboard.

Since we did not inform our observers that most of the visual stimuli had equal physical durations (physical duration differences were only present in catch trials), we subjected them to an extensive de-briefing procedure after the experiment. All observers reported that they considered duration judgments on some trials to be a lot easier than on other trials, but they also claimed that even on the relatively hard trials they usually had a reasonable idea of which stimulus had the longest duration. None of the observers reported to have been aware of the fact that almost all visual stimuli had no actual duration difference.



**FIGURE 6-2.** Set-up and results of Experiment 2. **A)** Observers reported which of two target stimuli they perceived to have a longer duration. Target stimuli could either be visual circles or squares or auditory tones (200 Hz) and the duration difference between the targets matched the individually determined just noticeable differences (JND) for each observer. Target stimuli could be accompanied by long background sounds (100 Hz) that could either be paired with the short target stimulus, the long target stimulus, or both. **B)** The results are plotted as the percentage correctly identified longer targets for each experimental condition. No significant effects were found of either target modality or non-target condition. Error bars represent S.E.M's.



### 6.6.2 Results and discussion

The results of Experiment 3 are shown in Figure 6-3B. The percentage of trials in which the visual target *with* an auditory non-target was perceived as longer than the visual target *without* an auditory non-target stimulus is plotted against the difference in duration between an auditory non-target and a visual target. A first thing to notice is that the presence of auditory non-targets can result in both longer and shorter perceived visual durations, depending on the relative duration of the non-target sounds. This effect of non-target duration is significant ( $F(14,165) = 14.27, p < 0.001$ ). If visual targets are paired with auditory non-targets of equal or longer physical duration, the perceived duration of this visual target is increased (Figure 6-3B)(T-tests on individual points:  $t(11)$  between 2.44 and 7.15, all p-values  $< 0.04$ ). Interestingly, even when the auditory non-target stimulus with which the visual target is paired is of exactly the same duration, the visual target is still perceived to have a longer duration than the visual-only stimulus of equal duration on 66.8% ( $\pm 4.7\%$  S.E.M.) of the trials, which is significantly above chance ( $t(11) = 3.59, p < 0.01$ ). When, however, the auditory non-targets are between 100 and 350 ms shorter than the visual targets they are paired with, the visual targets are significantly more often perceived to be shorter than they physically are (Ranging from  $59.5 \pm 3.2\%$  to  $75.0 \pm 3.8\%$  of the time,  $t(11)$  between -5.52 and -2.54, all p-values  $< 0.03$ ).

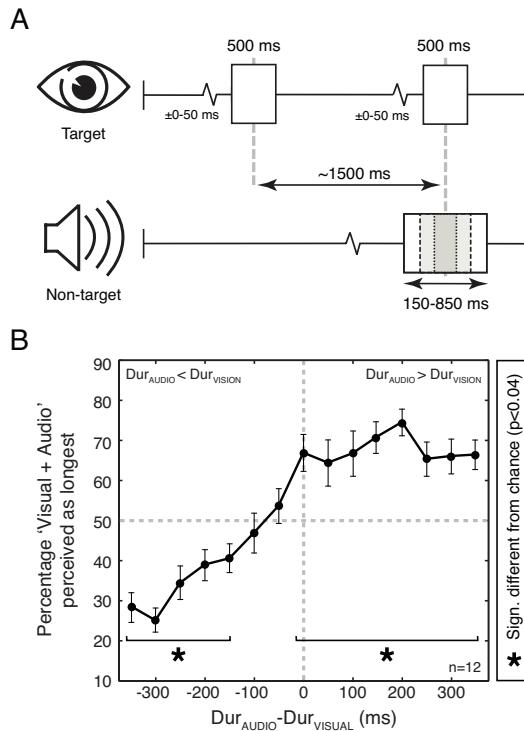
From the perspective of Scalar Expectancy Theory, a change in perceived duration can occur either by a change in pacemaker rate, a change in duration that the mode-switch is closed and pulses are fed into the accumulator, or by distortions in the translation of information from the accumulator stage to the reference memory. Previous studies have demonstrated that when an auditory and a visual stimulus have the same physical duration, the auditory stimulus is perceived to be longer than the visual stimulus (Penney et al., 2000; Walker & Scott, 1981; Wearden et al., 1998). This effect has been attributed to an ‘auditory pacemaker rate’ that is faster rate than the ‘visual pacemaker rate’ (Penney et al., 2000; Wearden et al., 1998). When auditory and visual stimuli are grouped crossmodally, there could be an ‘audiovisual’ pacemaker rate that is faster than the visual pacemaker rate, yet slower than the auditory pacemaker rate.

Whereas this explanation holds well for a perceived dilation of visual duration, it cannot explain the observed perceived shortening of visual durations when visual targets are paired with significantly shorter auditory non-targets. To be able to account for this effect we have to incorporate an audiovisual-integration-driven, temporal ventriloquism-like change in mode switch timing. In temporal-order-judgment studies, typically used to investigate temporal ventriloquism, it has been shown that the temporal order discrimination performance for two subsequently presented visual stimuli greatly improves when the first visual stimulus is preceded by an auditory tone and the second visual stimulus followed by another tone (Bertelson & Aschersleben, 2003; Getzmann, 2007; Morein-Zamir et al., 2003). The predominant explanation for this effect is that the temporal onsets of the auditory stimuli capture the onsets of the visual stimuli, thereby effectively shifting their perceived temporal position further apart. A similar thing could happen in our experiments where the actual on- and offsets of visual targets may be involuntarily captured by the on- and offsets of the auditory non-targets and perceptually shift towards them. In SET, this would result in an altered closing time of the mode switch. An alternative, more trivial explanation for our results might be that our observers strategically switched to reporting differences between the auditory non-target duration and the visual target duration when they were unable to reach a decision about a differ-



ence between visual target durations. This seems unlikely since, upon de-briefing, observers reported not to have been aware of the fact that the vast majority of visual targets actually had equal durations. If the abovementioned strategy would still play a role in observer's reports, it would thus not be a consciously initiated strategy, but rather a subconscious neural process of which an observer is not aware (Repp & Penel, 2002). This may in fact be just another way of suggesting that the brain's time perception mechanism is subjected to crossmodal influences that change its functional 'strategies'.

When visual targets were paired with an auditory non-target of equal duration (point 0 on the x-axis in Figure 6-3B), they were usually perceived to have a longer duration than their visual-only companion targets. Since temporal ventriloquism-like effects are unlikely to play a significant role here (on- and offsets moments are the same for both modalities), the cause of this distortion may be suspected to lie in an altered rate of the internal clock, or in SET terminology: the rate of the pacemaker. The idea that auditory distracters may influence the rate of the pacemaker for visual duration judgments has been proposed before (Chen & Yeh,



**FIGURE 6-3.** Set-up and results of Experiment 3. **A)** Observers reported which of two visual target stimuli they perceived to have a longer duration. One of the two target stimuli was accompanied by an auditory non-target stimulus with a variable duration (this was counterbalanced between the first and second stimulus). Visual stimuli always had a duration of 500 ms, while durations of the auditory non-target stimuli ranged from 350 ms (150 ms shorter than the visual target) to 850 ms (150 ms longer than the visual target). **B)** The results of the experiment are plotted as the percentage of trials in which the visual target stimulus with an auditory non-target was perceived to have the longest duration of the two visual stimuli. The thick black line plots the averaged data of 12 observers. Error bars represent S.E.M.



2009; Penton-Voak et al., 1996; Wearden et al., 1998), but it remains unclear whether the time perception system just switches from a slow visual to a fast auditory pacemaker, or whether there could be something like an audiovisual pacemaker running at an intermediate rate. These possibilities are tested in Experiment 4.

## 6.7 Experiment 4

### The relative rates of the pacemaker

From a Scalar Timing Theory point of view, the finding that the perceived duration of a visual target stimulus increases in the presence of an auditory non-target stimulus of equal physical duration (Experiment 3) could be attributed to an increased pulse rate of a central amodal pacemaker (Chen & Yeh, 2009; Penton-Voak et al., 1996; Wearden et al., 1998), a difference in the intrinsic rates of independent modality-specific pacemakers (Mauk & Buonomano, 2004; van Wassenhove et al., 2008), or to modality-specific accumulator dynamics dealing with pulses from a central amodal pacemaker (Rousseau & Rousseau, 1996). Since our current experiments cannot explicitly distinguish between these possibilities, we will discuss our data using a more general terminology of modality-specific pacemaker rates rather than attribute any effect to actual modality-specific pacemakers or accumulators. In this fourth experiment we set out to unravel whether any differences can be observed between supposedly ‘pure visual’, ‘pure auditory’ and ‘audiovisual’ pacemaker rates.

#### 6.7.1 Method

Eleven observers (ranging in age between 21 and 28 years, 7 males and 4 females, 2 authors) participated in this experiment. Three of these observers (including the authors) also participated in Experiments 1 and 2, while three others also participated in Experiment 3. The observers were presented with a visual and an auditory target stimulus and they reported which of the two they perceived to have a longer duration (probing the auditory vs. visual pacemaker rate). The duration of the visual stimulus was always 500 ms, while the duration of the auditory stimulus varied from 400 to 600 ms in steps of 50 ms. Each pair of stimuli was presented 40 times in pseudorandom order and individual trials started when observers pressed a designated key on a standard keyboard. The order of the visual and auditory stimuli within a single trial was pseudo-randomly chosen to prevent fixed order effects (Grondin & McAuley, 2009).

Three of the above mentioned observers (one author) and eight additional observers (ranging in age between 20 and 28 years, 2 males and 3 females) also performed two additional conditions in which they compared 1) the durations of a pure auditory target stimulus and a visual target stimulus that was paired with an auditory non-target stimulus of equal physical duration (probing the auditory vs. audiovisual pacemaker rate), and 2) the durations of a pure visual target stimulus and a visual target stimulus that was paired with an auditory non-target stimulus of equal physical duration (probing the visual vs. audiovisual pacemaker rate). The unimodal stimuli were always 500 ms, while the crossmodal stimulus pairings varied in duration from 400 ms to 600 ms in steps of 50 ms. Each stimulus pair was presented 40 times in pseudorandom order. The order of the stimuli within a trial was pseudorandom as well, and individual trials started when observers pressed a designated key on a standard keyboard.



### 6.7.2 Results and discussion

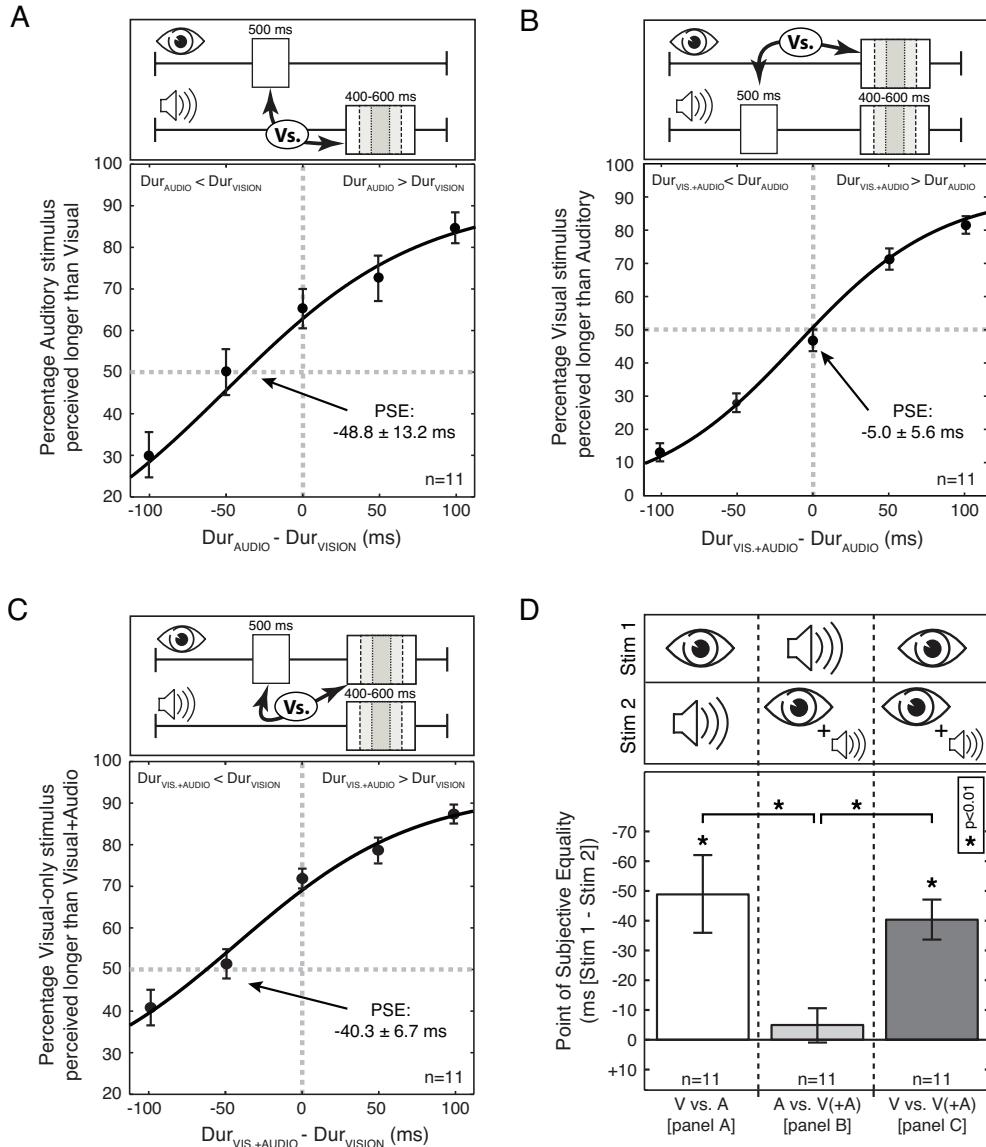
The results of Experiment 4 are displayed in Figure 6-4. We fitted the data of individual observers to a Weibull function using the psignifit 2.5.6 toolbox for matlab (see <http://bootstrap-software.org/psignifit/>) that implements a maximum-likelihood method (Wichmann & Hill, 2001) to estimate the point of subjective equality (PSE). At the PSE, observers are equally likely to label either of two stimuli as longer, indicating that their subjective durations can be regarded as equal. The group-averaged psychometric curves are plotted in Figures 6-4A, 6-4B and 6-4C as thick black lines. Figure 6-4A demonstrates a significant effect of duration difference on the percentage of 'auditory longer' responses ( $F(4,50) = 17.62, p < 0.001$ ). When auditory and visual stimuli have the same physical duration, the auditory stimulus is perceived to have a longer duration in 65.3% ( $\pm 4.8\%$  S.E.M.) of the trials. This percentage is significantly above chance ( $t(10) = 3.20, p < 0.01$ ). The average PSE was -48.8 ms ( $\pm 13.2$  ms S.E.M.), which is significantly different from zero ( $t(10) = -3.70, p < 0.01$ ), indicating that on average observers would perceive our visual and auditory targets as having equal duration when the visual stimulus was in fact about 50 ms longer.

Figure 6-4B plots how the perceived duration of a visual target stimulus paired with an equally long non-target sound compares to the perceived duration of purely auditory target stimuli. Data-points represent average data of 11 observers. There is a significant effect of duration difference on the percentage of 'visual longer' responses ( $F(4,50) = 69.63, p < 0.001$ ). When the audiovisual stimulus is of equal physical duration as the purely auditory stimulus, observers do not significantly perceive any of the two as longer ( $t(10) = -1.07, p = 0.31$ ). This notion is confirmed by the fact that the average PSE is not significantly different from zero ( $-5.0 \pm 5.6$  ms S.E.M.,  $t(10) = -0.88, p = 0.40$ ).

In Figure 6-4C, purely visual target durations are compared with the duration of visual target stimuli that are paired with equally long non-target sounds. There is again a significant effect of duration difference on the percentage of 'visual-only longer' responses ( $F(4,50) = 37.03, p < 0.001$ ). Also, when the visual target stimuli have the same physical duration, the one paired with the sound is perceived to have a longer duration in 64.8% ( $\pm 3.0\%$  S.E.M.) of the trials. This percentage is significantly above chance ( $t(10) = 5.00, p < 0.01$ ). The average PSE was -40.3 ms ( $\pm 6.7$  ms S.E.M.), which is significantly different from zero ( $t(10) = -5.98, p < 0.01$ ), indicating that on average observers would perceive the targets as having equal duration when the visual stimulus paired with the sound was in fact about 40 ms longer.

To directly compare the supposed effects of purely auditory, purely visual and audiovisual pacemaker rates we plotted the average PSE's for the different conditions in Figure 6-4D. The difference in PSE for comparing pure visual target durations to pure auditory target durations (white bar) and comparing pure visual target durations to visual target durations paired with a sound (dark grey bar) is not significant ( $F(1,20) = 0.33, p = 0.57$ ), while both these conditions do significantly differ in PSE from the condition in which pure auditory target durations and visual target durations paired with sounds are compared (light grey bar;  $F(1,20) = 9.33, p < 0.01$  and  $F(1,20) = 16.17, p < 0.01$  respectively). From the perspective of SET, these results suggest that the time perception system either automatically switches to the auditory pacemaker rate when sounds are present or that there is an additional audiovisual pacemaker rate that is highly similar to the auditory pacemaker rate.

It is possible that in all experimental conditions where observers were asked to use the duration of a visual target stimulus for comparison in a duration discrimination task and, at





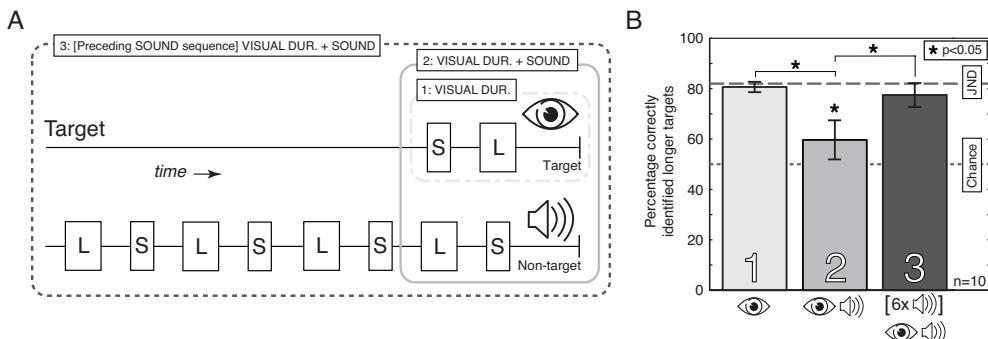
◀ **FIGURE 6-4.** The results of Experiment 4. A) The duration of a sound is compared with that of a visual stimulus. The group-averaged psychometric curve (thick black line) is shifted to the left indicating that when a sound and a visual stimulus are of equal physical duration, the sound is significantly more often perceived to have a longer duration than the visual stimulus. Data-points represent the average data of eleven observers (error bars are S.E.M.). B) The duration of a target sound is compared with that of a visual target stimulus that is paired with a non-target sound of equal physical duration as the visual stimulus. When the two targets are of equal physical duration observers perform at chance level. Data-points represent the average data of eleven observers (error bars are S.E.M.) and the thick black line is the group-averaged psychometric function. C) The duration of a visual target stimulus is compared with that of a second visual target stimulus paired with a non-target sound of equal physical duration. The group-averaged psychometric curve (thick black line) is shifted to the left indicating that a visual stimulus is perceived to have a longer duration when it's paired with a sound of equal duration. Data-points represent the average data of eleven observers (error bars are S.E.M.). D) A comparison of the shifts in PSE for the experiments presented in panel A to C. Significant deviations from zero are observed for Visual vs. Auditory targets (white bar) and Visual vs. Visual targets with Auditory non-targets (dark grey bar), but not for Auditory vs. Visual targets with Auditory non-targets (light grey bar). Error bars indicate S.E.M.

the same time, ignore the auditory non-target that it is paired with, observers in fact subconsciously switched to using the duration of the auditory non-target (Repp & Penel, 2002). In order to distinguish such an explanation from the more tentative hypothesis that the observed changes in duration discrimination performance are due to subconscious crossmodal grouping and its consequential influences on the brain's time perception system, we performed a fifth experiment. In that experiment we manipulated the likeliness of intramodal and crossmodal stimulus grouping to investigate how this affects the previously demonstrated crossmodal distortions in duration discrimination.

## 6.8 Experiment 5

### Intramodal grouping prevents crossmodal duration effects

While the results of all previous experiments strongly suggest that the crossmodal grouping of auditory and visual stimuli is an essential prerequisite for the occurrence of crossmodal interactions in duration perception, we cannot exclude an alternative hypothesis according to which observers exclusively use the duration of auditory non-targets in their duration discrimination tasks. This fifth experiment examines if auditory influences on visual duration perception persist if we disturb the supposed crossmodal grouping by allowing the auditory stimuli to be grouped intramodally rather than crossmodally. This manipulation should have no effect on a behavioral switch towards using the auditory non-targets in the duration comparison. It should, however, affect changes in perceived duration based on crossmodal interactions of grouped stimuli in the time perception system. The fact that intramodal perceptual grouping reduces or abolishes crossmodal effects has been shown with other experimental paradigms before (Bruns & Getzmann, 2008; Keetels et al., 2007; Lyons et al., 2006; Penton-Voak et al., 1996; Sanabria et al., 2005; Sanabria et al., 2004; Vroomen & de Gelder, 2000), but if we want to use the argument of subconscious grouping in the context of our own findings, we need to demonstrate that it is also true for the duration discrimination paradigm we used in our experiments.



**FIGURE 6-5.** Set-up and results of Experiment 5. A) The experiment comprised three conditions: 1) A purely visual duration discrimination task (L = Long, S = Short), 2) A visual duration discrimination task with the visual targets paired with auditory non-targets that had a duration difference with opposite sign compared to the duration difference of the visual targets, and 3) The same condition as in 2 but this time the comparison was preceded by a stream of auditory non-targets that were rhythmically consistent with the two non-target sounds in the task. B) Visual discrimination performance (bar marked 1) is at the predetermined level for individual JNDs. In the presence of non-target sounds performance drops significantly (bar marked 2), but when a stream of non-target sounds preceded the task, performance goes back up (bar marked 3). Error bars indicate S.E.M.s.

### 6.8.1 Method

Ten observers (ranging in age between 20 and 28 years, 7 males and 3 females, 1 author) participated in this experiment. Two of these observers also participated in Experiments 1, 2 and 4; one participated in Experiments 3 and 4; three in Experiment 1 and 2, and four observers did not partake in any of the other experiments. Prior to the main experiment, observers were subjected to a staircase procedure in order to determine their discrimination thresholds for visual durations. The staircase procedure used the Psychtoolbox' Quest algorithm and consisted of 25 trials, converging on 82% correct. It was performed three times and the averages of the three obtained threshold values were taken as the individual observer's just noticeable differences (JNDs). These JNDs (average  $135.6 \text{ ms} \pm 9.8 \text{ ms}$  S.E.M.) were then used in the main experiment where observers were asked to judge which of two visual stimuli with a duration difference equal to their individually determined JND had a longer duration. Three conditions were tested. In the first condition, the visual target stimuli were the only stimuli presented. In the second condition the visual target stimuli were paired with auditory non-target stimuli that had a difference in duration opposite to that of the visual stimuli (as in Experiment 1). The third condition was similar to the second, but now the critical stimulus presentations were preceded by three unimodal repetitions of the non-target sound stimuli (Figure 6-5A). The visual target stimuli had durations of  $500 \text{ ms} \pm \text{JND}/2$  and the order in which the long and short stimuli were presented were counterbalanced. Auditory non-target stimuli had durations of 400 ms (paired with the long visual stimulus) and 600 ms (paired with the short visual stimulus). The visual target stimuli and auditory non-target stimuli were aligned by their temporal midpoint and the interstimulus interval between target stimuli was 1500 ms. In the condition with the preceding sounds, the pair of short and long tones was played three times with the tones in the same order as they would eventually have, when



paired with the visual targets. In addition, these pairs were presented with fixed interstimulus intervals of 1500 ms (between the midpoints) to create the vivid experience of a consistent auditory stream. Each experimental condition was repeated 20 times in pseudorandom order.

### 6.8.2 Results and discussion

The results of this experiment are plotted in Figure 6-5B. When observers discriminate durations of two purely visual targets (Figure 6-5A) they perform at the same level that was used to determine their individual JNDs (light grey bar marked with 1 in Figure 6-5B) ( $80.7\% \pm 1.9\%$  correct,  $t(9) = -0.72$ ,  $p = 0.49$ ). This is not surprising since they are essentially performing the same task as in the preceding staircase procedure. When the visual targets are paired with auditory non-targets having opposite duration differences (Figure 6-5A), they are performing the same task as the observers of Experiment 1. As was found in Experiment 1, performance on identifying the longer visual stimulus is significantly impaired by the presence of auditory non-targets (middle bar marked with 2 in Figure 5B) ( $59.7\% \pm 7.8\%$  correct,  $t(9) = -2.88$ ,  $p < 0.02$ ).

However, if this condition is preceded by a stream of irrelevant auditory non-targets, performance goes back up and is indistinguishable from the 82% threshold level ( $77.5\% \pm 4.7\%$  correct,  $t(9) = -0.96$ ,  $p = 0.36$ ). We therefore conclude that a subconscious intramodal grouping of auditory non-targets into a consistent auditory stream prevents the subconscious crossmodal binding that is necessary for the crossmodal influence of auditory non-targets on the discrimination performance of visual target durations. These results clearly advocate the idea that the distortions of visual duration discrimination performance by irrelevant auditory stimuli presented in this study are based on interactions between crossmodally grouped stimuli within the time perception system, instead of a mere behavioral switch towards reporting the durations of auditory non-targets rather than the durations of visual targets.

## 6.9 General discussion

Adequate estimation of event durations is critical for both behavioral and cognitive performance, but how does the brain estimate event durations? Perceiving the duration of an event is in a sense a classic cue combination problem (e.g. Landy et al., 1995). In order to be as accurate as possible, the brain will rely on all available relevant cues and weigh their influences based on their relative reliabilities, determined by a multitude of factors such as signal to noise ratio and intrinsic resolution. But how does the brain ‘know’ whether different cues provide information about a single perceptual objective and should thus be combined? This question is particularly interesting when the different cues come from sensory modalities (Driver & Spence, 1998; Driver & Spence, 2000; Kanai et al., 2007b; Sugita & Suzuki, 2003; Vroomen et al., 2004; Wallace et al., 2004). In a series of five experiments we explored the interactions between vision and audition in the perception of event duration. Since the auditory system is thought to represent time more reliably than the visual system, we would expect the brain to recruit auditory temporal information when it needs to resolve visual temporal problems (like in a visual duration discrimination task). Visual information on the other hand, should not be used when resolving auditory temporal tasks. This expectation is confirmed by our experiments that provide clues about when and how these crossmodal effects may occur.



Experiment 1 confirms the asymmetric nature of crossmodal influences on duration discrimination with simultaneous auditory and visual sensory input. Performance on a visual duration discrimination task was significantly impaired by the presence of auditory non-targets that either had an opposite duration difference compared to the visual stimuli they were paired with, or no duration difference at all. Oppositely, visual non-targets did not influence performance on an auditory duration discrimination task at all. This asymmetry in crossmodal audiovisual influences confirms previous results (Bruns & Getzmann, 2008; Chen & Yeh, 2009; but the opposite has also been shown: van Wassenhove et al., 2008) and is most likely caused either by asymmetries in involuntary crossmodal grouping or by asymmetries in modality specific temporal reliabilities. The former suggests that irrelevant sounds are automatically grouped with relevant visual targets whereas irrelevant visual stimuli are not automatically grouped with relevant sounds. If this is the case, it could very well be due to asymmetries in the modality specific reliability of temporal information. Since the auditory system has a much higher temporal resolution than the visual system, it is not improbable that the brain would by default employ available auditory information when it estimates visual durations (and not the other way around). This becomes even more likely if we realize that crossmodal duration distortions only occur when stimulus onsets are in close temporal proximity of each other. When, in Experiment 1, auditory non-targets were slightly shifted in time relative to the visual targets they were paired with, their influence on visual duration discrimination performance significantly decreased.

The results of Experiment 2 add further support to this idea. When the same visual duration discrimination task as in Experiment 1 is performed while visual targets are paired with auditory non-targets that start well before and end well after the visual targets are presented, performance is not influenced by the sounds. Consequently, the mere presence of sound is not enough to evoke changes in subjective visual durations. If the onsets and offsets of the auditory and visual stimuli are distinctly different this could be interpreted as a no-go signal for crossmodal binding.

Whereas the crossmodal difference in on- and offset may function as an important prerequisite for the mere occurrence of crossmodal effects (i.e. when it is within the range of on- and offset asynchronies where crossmodal effects do occur), the actual size of this difference in on- and offset may be important. In a related phenomenon termed temporal ventriloquism, the performance on a visual temporal order judgment (TOJ) task is influenced by the presence of irrelevant auditory stimuli (Bertelson & Aschersleben, 2003; Getzmann, 2007; Morein-Zamir et al., 2003). If a first non-target sound is played before a first visual target and a second non-target sound is played after a second visual target, observers are able to detect much smaller temporal differences in the onset of the two visual targets than when these two non-target sounds are played in between two visual targets. The prevailing explanation of temporal ventriloquism suggests that visual onsets are ‘pulled’ towards the auditory onsets, thereby creating a perceived audiovisual onset that is a weighted average of the visual and auditory onsets but leans more towards the auditory than towards the visual onset moment. In a similar way the on- and offsets of auditory non-targets may shift the perceived on- and offset of visual targets and influence the subjective visual target duration.

Theories of duration perception often boil down to variations on Scalar Expectancy, or Scalar Timing Theory (SET: Gibbon, 1977; Gibbon et al., 1984). Basically, this theory states that in order to perceive durations a neural mode switch closes, allowing pulses from a pace-



maker to be collected in an accumulator. After reopening the switch the number of pulses in the accumulator is compared against a reference memory to establish a perceived duration. From an SET point of view, the temporal ventriloquism-like explanation of altered subjective duration could be interpreted as an altered duration of mode-switch closure (closed upon perceived onset, opened upon perceived offset). Alternatively, changes in subjective duration could also be caused by a change in SET's pacemaker rate or the translation of the accumulator state to the reference memory ('memory-mixing', Penney et al., 2000).

Conventional approaches to disentangle effects of the mode switch and the pacemaker rate focus on the latency of the switch by calculating the *intercept* and *slope* of a linear regression through the relationship between a range of physical and perceived base durations (Penton-Voak et al., 1996; Wearden et al., 1998). This analysis method generally demonstrates that audiovisual duration distortions result from a change in pacemaker rate and not from changes in switch latency (Chen & Yeh, 2009; Penton-Voak et al., 1996; Wearden et al., 1998). Furthermore, it has been repeatedly shown that auditory events are generally perceived to have a longer duration than visual events of equal physical duration (but see Boltz, 2005 for the absence of this effect if naturalistic stimuli are used; , or Grondin, 2003 for a review of the specific circumstances for intermodal effects on timing deviations ; Penney et al., 2000; Walker & Scott, 1981; Wearden et al., 1998). Studies using visual and auditory targets of equal physical duration generally reach the conclusion that the auditory pacemaker rate is faster than the visual pacemaker rate and that the auditory distortions of perceived visual durations should be attributed to these changes in the rate of the internal clock (Chen & Yeh, 2009; Penton-Voak et al., 1996; Wearden et al., 1998).

In our approach, we cannot calculate intercepts and slopes because we use a single base duration. However, our results from Experiment 3 do suggest that the temporal ventriloquism-like effects on mode switch latency and the previously demonstrated variable internal clock rates may both play a role. When we paired a visual target with an auditory non-target whose duration ranged from shorter, via equally long, to longer than the visual target, we observed that when this visual target was accompanied by an equally long or longer sound it was perceived as longer than an equally long visual target without a sound. This confirms previous findings and suggests that the rate of the internal clock indeed plays a role. However, we also observed a significant shortening of the subjective visual duration when the visual target was paired with a shorter auditory non-target. Such shortening cannot be explained by a pacemaker rate account, but it does fit the prediction of temporal ventriloquism-like effects on the mode-switch latency.

The collective results of the first three experiments raise two important questions. First, does the time perception system switch between a visual and an auditory pacemaker rate or does it dynamically scale its pacemaker rate in the context of crossmodal sensory evidence? This question was addressed in Experiment 4 where we compared the relative pacemaker rates derived from duration discrimination experiments in which observers compared visual target durations against auditory target durations, unimodal visual target durations against auditory target durations paired with visual non-targets, and unimodal auditory target durations against auditory target durations paired with visual non-targets. The results confirm the idea that both the auditory pacemaker rate and the hypothesized audiovisual pacemaker rate are higher than the visual pacemaker rate (Penton-Voak et al., 1996; Wearden et al., 1998), but they also suggest that if an audiovisual pacemaker rate would even exist, it cannot be distin-



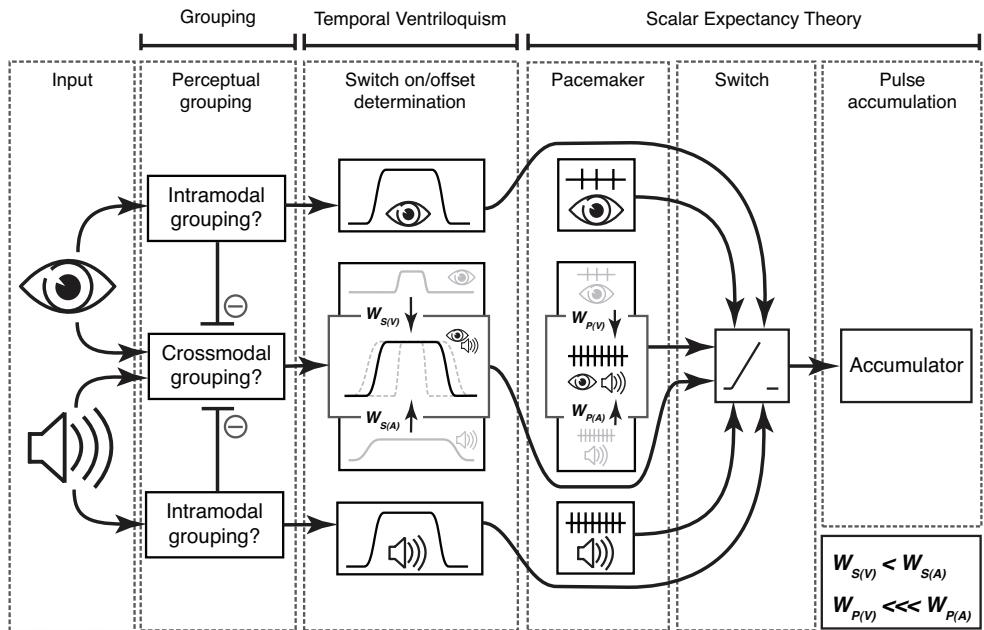
guished from a purely auditory pacemaker rate.

The second question that arises from our experiments concerns the nature of the demonstrated auditory influences on visual duration discrimination performance. Are the effects truly caused by crossmodal interactions in the time perception system or are observers (subconsciously) switching from using visual target durations to using auditory non-target durations in the visual duration discrimination task? Knowing that a global temporal context might influence perceived interval duration (Jones & McAuley, 2005), we created experimental conditions in which we manipulated the likeliness of crossmodal grouping of auditory and visual stimuli by introducing a global context that promoted intramodal grouping of auditory stimulus elements instead (Experiment 5). Such a manipulation should not affect performance if observers simply used auditory instead of visual duration in their task, but we expected it to have strong effects if the demonstrated crossmodal influences on duration discrimination depended on crossmodal grouping of auditory and visual stimuli. Our results demonstrate that the crossmodal effects of auditory stimuli on visual duration discrimination performance are abolished when the auditory stimuli can be intramodally grouped with a preceding stream of sounds and confirm previous studies that demonstrated a precedence of intramodal over crossmodal grouping that prevents crossmodal influences (see also Bruns & Getzmann, 2008; Keetels et al., 2007; Lyons et al., 2006; Penton-Voak et al., 1996; Sanabria et al., 2005; Sanabria et al., 2004; Vroomen & de Gelder, 2000).

When all results are taken into account, the general rule emerges that auditory temporal information is recruited to improve the accuracy of visual duration judgments, but only if there is sufficient reason for the brain to assume that the information of the two senses are about the same event. In none of our experiments was there an *a priori* reason why pure tones and visual squares or circles should be considered as independent information sources of a single event, but their temporal co-occurrence was apparently enough to evoke subconscious crossmodal binding.

We can schematically depict our interpretation of the results in a functional crossmodal SET-model (Figure 6-6). Such a scheme consists of an intra- and crossmodal grouping stage where stimuli are perceptually grouped, followed by temporal ventriloquism-like effects on the control of the mode switch, and a pacemaker that running at weighted crossmodal pulse rate that, under audiovisual conditions, is heavily (if not completely) dominated by the ‘auditory rate of time’.

An alternative explanation that seems to fit our data equally well at first sight incorporates the process of ‘memory-mixing’ instead of crossmodal pacemaker rates as a source for changes in the subjective rate of the internal clock (Penney et al., 2000). The memory-mixing hypothesis is usually tested using a duration-bisection paradigm in which target stimuli are judged to be closer in duration to either a short or a long trained reference duration that is kept in memory. If the remembered reference durations for auditory and visual targets are similar, distortions occur that indicate that these reference durations become a mixture of a fast-rate auditory and a slow-rate visual pulse-count. Consequently, stimuli of equal physical duration as the reference memory are perceived as longer when they are auditory, and as shorter when they are visual (Penney et al., 2000). In our paradigm, observers are asked to make shorter-longer judgments between sequentially presented stimuli, without having to keep a trained, standard reference in long-term memory. Memory-mixing may however still play a role in the immediate retrieval of interval duration information for perceptual judgments. In this



**FIGURE 6-6.** A functional scalar expectancy model of audiovisual duration perception. Visual and auditory sensory input first enters the brain as separate sources of information. It is then processed by a perceptual grouping stage that determines if intra- or crossmodal will occur. If crossmodal grouping does not occur, a unimodally defined mode switch is closed to allow pulses from a pacemaker with a unimodally defined rate to be collected in an accumulator. If the auditory and visual inputs are however grouped crossmodally, the period that the mode switch is closed is determined by the weighted contribution ( $W$ ) of both modalities (temporal ventriloquism) with a stronger emphasis on the auditory information ( $WS(A) > WS(V)$ ). The rate at which pulses are then collected in the accumulator is again a weighted average of the unimodal auditory and visual pulse rates. Here the dominance of audition is even stronger (in Experiment 5, the audiovisual pacemaker rate was not statistically different from the auditory pacemaker rate) resulting in weight factor asymmetry  $WP(A) \ggg WP(V)$ .



explanation, the crossmodal effects would not happen during the encoding of target duration, where the visual and auditory pacemaker each feed their pulses into a unimodal accumulator, but rather upon retrieval of the encoded target duration. The short-term unimodal memories could become mixed, resulting in distortions of the subjective crossmodal interval duration.

In conclusion, we demonstrated that the brain automatically uses temporal information from irrelevant sounds to judge durations of visual events, provided that the temporal characteristics of the two sensory streams of information are such that crossmodal binding is feasible. The distortions of visual duration perception through the crossmodal influence of audition is caused both by the perceived onset and offset of the visual stimuli (a temporal ventriloquism-like effect for interval duration) and by the integrated activity of a functional pacemaker during this period.

Interesting objectives for future studies include investigations of the perceptual grouping process (what are the critical criterions for intra- and crossmodal grouping?), the apparent asymmetry in crossmodal influences (will lower auditory signal-to-noise ratios allow visual influences on auditory duration perception?), the possible role of memory-mixing, and a search for activity shifts in pacemaker-like neural substrates under different uni- and multimodal conditions (Buhusi & Meck, 2005). Despite a long history in time perception research there is obviously still a lot of effort needed before we might begin to understand how the brain accomplishes the seemingly effortless perception of the temporal aspects of our multimodal surroundings.

## 6.10 Acknowledgements

This research was supported by a VIDI grant from the Netherlands Organization for Scientific Research (NWO) and a High Potential grant from Utrecht University, both awarded to RvW. We thank Chris Paffen for helpful comments and suggestions.

# **Chapter 7**

## **Experience-driven plasticity in binocular vision**

*It is the potential for plasticity of the relatively stereotyped units of the nervous system that endows each of us with our individuality.*

**Eric Kandel**

**Published as**

Klink, P.C., Brascamp, J.W., Blake, R., & van Wezel, R.J.A. (2010). Experience-driven plasticity in binocular vision. *Current Biology*, 20(16), 1464-1469

**Highlighted in**

Wilson, H.R. (2010) Binocular rivalry: Neurons unwire when they can't simultaneously fire. *Current Biology*, 20(17), R715-717



*The story behind this chapter goes all the way back to the Functional Neurobiology Christmas lunch in 2008. Jan Brascamp had recently defended his PhD thesis and would soon move to Nashville to start working as a postdoctoral fellow with Randolph Blake in the Vision Research Center of Vanderbilt University. We agreed that I would visit his new lab ‘sometime’ in the future to do a project together. The opportunity for a visit was there when we both attended the Vision Science Society meeting in Naples (USA) in May of 2009. We figured that the conference would give us more than enough inspiration to come up with an idea for a project to work on in Nashville afterwards. However, all of a sudden (or so it seemed) we were drinking coffee in Atlanta while we were waiting for our connecting flight to Nashville and we realized that we had not yet decided what project we were going to start working on the next day. Upon our arrival in Nashville we met up with some of Jan’s colleagues at one of the local bars and every now and then we talked about a project idea. I believe it was Jan that came up with the idea of doing an adaptation binocular rivalry experiment involving the prolonged wearing of eye-patches. This idea was promptly dubbed ‘the pirate experiment’ and stimulated by the mockery and general laughter that we received from whoever we explained the idea to, we started programming experimental code the next day. The first results were so promising that we decided to try an experiment in which we would patch one of our own eyes for a continuous period of 24 hours. I can recommend this procedure to anyone that is interested in some really psychedelic perceptual experiences and not afraid to bump into the occasional lamppost. The shocked reactions from some of the researchers in Wilson Hall that saw us walking around with one of our eyes patched were hilarious. No, we did not get in a fight, and we did not have any nasty eye infection either. It was just a rather unconventional experiment. I continued the project when I got back to Utrecht, but the ongoing digital discussions among Jan, Randolph, Richard and myself drove the eventual set of experiments that are now described in this chapter. While the title of the manuscript ultimately became ‘Experience-driven plasticity in binocular vision’, the project folder on my computer has always remained labeled ‘The Nashville Pirate Experiment’.*

---



## 7.1 Abstract

Experience-driven neuronal plasticity allows the brain to adapt its functional connectivity to recent sensory input. Here we use binocular rivalry (Alais & Blake, 2005), an experimental paradigm where conflicting images are presented to the individual eyes, to demonstrate plasticity in the neuronal mechanisms that convert visual information from two separated retinas into single perceptual experiences. Perception during binocular rivalry tended to initially consist of alternations between exclusive representations of monocularly defined images, but upon prolonged exposure, mixture percepts became more prevalent. The completeness of suppression, reflected in the incidence of mixture percepts, plausibly reflects the strength of inhibition that likely plays a role in binocular rivalry (Hollins, 1980). Recovery of exclusivity was possible, but required highly specific binocular stimulation. Documenting the prerequisites for these observed changes in perceptual exclusivity, our experiments suggest experience-driven plasticity at interocular inhibitory synapses, driven by the (lack of) correlated activity of neurons representing the conflicting stimuli. This form of plasticity is consistent with a previously proposed, but largely untested, anti-Hebbian learning mechanism for inhibitory synapses in vision (Barlow, 1990; Barlow & Földiák, 1989). Our results implicate experience-driven plasticity as one governing principle in the neuronal organization of binocular vision.

## 7.2 Results

Perceptual advantages of binocular vision including stereopsis and enhanced contrast sensitivity through binocular summation require integration of initially separated monocular streams of information. Mechanisms responsible for binocular integration are shaped by activity-dependent neural development, both prenatally when ocular dominance columns are first established and for several years postnatally, when binocular mechanisms are refined based on visual experience (see Katz & Crowley, 2002). Whereas the neuronal components subserving binocular integration may not change much after this critical period, the computational mechanisms, likely reflected in synaptic connectivity and efficacy, may be continuously recalibrated in response to modified sensory experience. This ongoing neuronal fine-tuning might in fact be the reason why some strabismus patients that have not adequately developed stereopsis during childhood can still acquire stereoscopic depth vision later in life through extensive visual therapy (for anecdotal evidence see Barry, 2009).

Exposure to binocular rivalry stimuli (Alais & Blake, 2005) creates a well-controlled modified sensory context deviating from the system's 'standard' in the sense that the brain receives incompatible, non-matching inputs instead of matching ones. Under such conditions, binocular integration fails and, instead, observers tend to alternately perceive the monocular images. This perceptual cycling is commonly believed to arise from neural processes that include mutual inhibition between neuronal representations of the two images (Alais & Blake, 2005; Klink et al., 2008b; Wilson, 2007). During smaller fractions of the time viewing rivalry, observers also transiently perceive various mixtures of both monocular images (Brascamp et al., 2006; Hollins, 1980; Yang et al., 1992), the most common being transparent superimpositions of both images and patchwork-like zones of local monocular dominance termed 'piecemeal' (Figure 7-1A). Mixtures suggest that even during rivalry periods of partial binocular integration occur. The absolute predominance of different mixture percepts depends on stimulus fea-



tures including size (Blake et al., 1992), spatial frequency (Hollins, 1980; O’Shea et al., 1997; Yang et al., 1992) and global context (Kovács et al., 1996), and the incidence of these lapses in perceptual exclusivity plausibly depends on the strength of mutual inhibition (Hollins, 1980), a notion supported by simulations with existing binocular rivalry models (Noest et al., 2007; Wilson et al., 2001) [Supplemental Information 1, Figure 7-S1].

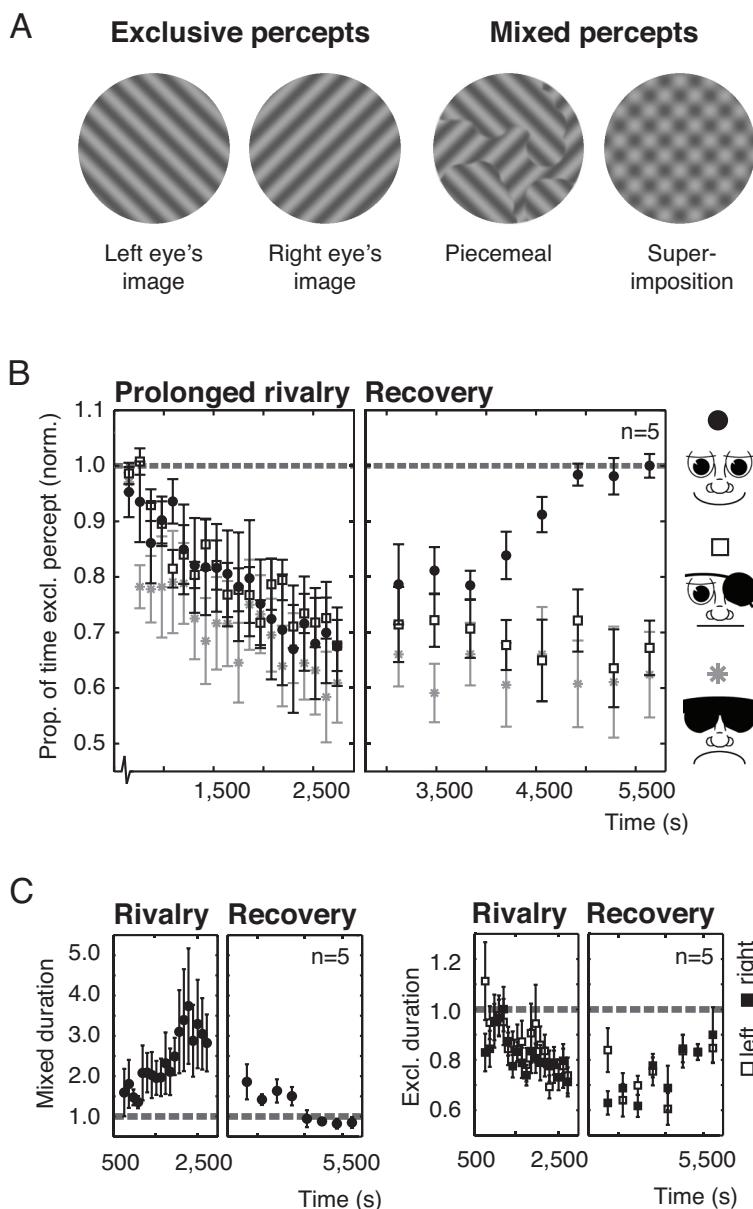
To test whether binocular integration is indeed a plastic mechanism that adapts to sensory experience, we presented the eyes with incompatible images for prolonged periods of time, sometimes interspersed with non-rival stimuli, while observers continuously reported whether they perceived either one of the exclusive monocular images or a mixture.

### 7.2.1 *Experiment 1: Perceptual exclusivity and binocular rivalry*

Observers viewed rival stimuli for prolonged durations while tracking periods of exclusive dominance and mixed percepts (Figure 7-1A). The same rival images were presented to each eye throughout the experiment. If initial perceptual exclusivity in binocular rivalry were due to the ‘unnatural’ sensory context of dissimilar images in the two eyes causing strong mutual inhibition and preventing binocular integration, we would expect exclusivity to progressively decrease while experience with the modified sensory context accumulates. As expected from earlier results (Hollins & Hudnell, 1980), our observations confirm this prediction (Figure 7-1B), demonstrating a substantial decrease in exclusivity over 35 minutes of rivalry (Figure 7-1B, Spearman rank correlation,  $R = -0.46$ ,  $p << 0.001$ ). Data points represent averaged data from 100 s rivalry trials, separated by 10 s rests. Data for individual observers was normalized by baseline values, determined in four rivalry trials (100 s rest) directly preceding the experiment.

The idea that the altered exclusivity in our experiment reflects experience-driven plasticity yields a second, more counterintuitive prediction: Exclusivity should not passively recover to baseline after having dropped during rivalry, but instead should require correspondence of visual signals from both eyes. In the second part of our experiment, immediately following the first, recovery of exclusivity was investigated with periods of exposure to various conditions of monocular or binocular stimulation. In one condition, observers walked around the laboratory with both eyes open. The matching, natural visual input to both eyes should cause a recalibration of the binocular mechanisms and re-strengthen the inhibition putatively weakened during rivalry. Because re-strengthening should be evidenced by increased perceptual exclusivity, the periods of free viewing were interspersed with brief rivalry trials. In a second condition free viewing was replaced by episodes without visual stimulation that should leave exclusivity unaltered. A third condition contained periods of monocular stimulation where one eye was patched during free viewing.

Significant increases in the proportion of exclusive dominance indeed occurred when two eyes received matched stimulation during free viewing (Figure 7-1B, solid circles, Spearman rank,  $R = 0.75$ ,  $p << 0.001$ ), both because mixed percept epochs became briefer and exclusive percepts became longer (Figure 7-1C, Spearman rank,  $R_{\text{mix}} = -0.68$ ,  $p_{\text{mix}} << 0.001$ ;  $R_{\text{excl}} = 0.45$ ,  $p_{\text{excl}} < 0.01$ ). Consistent with our prediction, no such recovery was observed throughout 48 minutes without visual stimulation (Figure 7-1B, gray asterisks; Spearman rank,  $R_{\text{no\_stim}} = -0.01$ ,  $p_{\text{no\_stim}} = 0.97$ ). Recovery was also entirely absent in the third, monocular stimulation condition (Figure 7-1B, open squares; Spearman rank,  $R_{\text{patched}} = -0.08$ ,  $p_{\text{patched}} = 0.61$ ) implying



**FIGURE 7-1.** Dynamics of perceptual exclusivity. **A)** Perceptual experiences during binocular rivalry. Exclusive percepts correspond entirely to one eye's stimulus. Mixed percepts resemble patch-like (piecemeal) or transparent superimpositions of the stimuli. **B)** The average proportion of normalized exclusivity for five observers against time. In the recovery stage, observers experienced either normal binocular vision (solid black circles), monocular vision only (open squares), or no visual stimulation at all (gray asterisks). **C)** The average epoch durations for mixed (left panel, solid circles), exclusive 'left', and 'right' percepts (right panel, open and black squares) of the two-eyes-stimulated condition. Dashed lines represent baseline levels. All error bars, s.e.m.



that binocular correspondence is essential for recalibration.

To further examine the failure of recovery with monocular stimulation, the two first authors subjected themselves to an extended period of continuous eye-patch wearing for 24 hours. Remarkably, decreased exclusivity levels barely recovered during this day of patching, yet recovery started immediately after both eyes received matched stimulation during free viewing [Supplemental Information 2, Figure 7-S2]. The longevity of decreased exclusivity in the absence of binocular input is reminiscent of the enduring time-course of contingent adaptation effects (e.g. Vul et al., 2008) and perhaps ‘storage’ of non-contingent after-effects (Thompson & Movshon, 1978; van de Grind et al., 2004; Verstraten et al., 1994). The slow decay of adaptation in all these cases could have a common cause: neurons encoding a specific adapting stimulus may retain their adapted state so as long as they are shielded from novel sensory experience, thereby precluding recalibration (Thompson & Movshon, 1978; van de Grind et al., 2004; Vul et al., 2008).

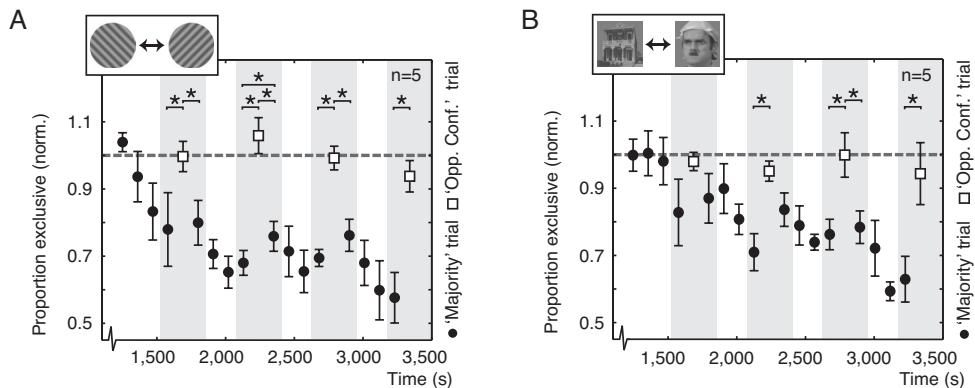
The results of these first experiments support experience-driven plasticity in the connectivity between neuronal representations involved in binocular rivalry by implying both the weakening and re-strengthening of inhibition in the anticipated conditions. While the necessity of binocular stimulation is clear, several remaining questions regarding the exact prerequisites for the observed changes in exclusivity prompted the following experiments.

### 7.2.2 Experiment 2: Decrease of perceptual exclusivity

To establish the prerequisites of decreasing exclusivity, two variations of our first experiment were performed in which we temporarily inverted the stimulus-eye configuration on every fifth trial (‘*opposite-configuration trials*’) so that the same monocular stimuli were presented to the opposite eyes. While this manipulation leaves the global competition between binocular stimulus representations unaffected, it does activate different monocular neurons on those specific opposite-configuration trials. Figure 7-2A demonstrates the results using the same stimuli as in Experiment 1. The opposite-configuration trials (white squares) have significantly higher levels of exclusivity than their temporal neighbors (Figure 7-2A, paired t-test,  $p < 0.05$ ). Additionally, exclusivity decreases only for the eye-stimulus configuration used in the majority of trials (Spearman rank,  $R_{\text{majority}} = -0.48$ ,  $p_{\text{majority}} << 0.001$ ;  $R_{\text{opp-conf}} = -0.22$ ,  $p_{\text{opp-conf}} = 0.36$ ).

Whereas superimposition mixture percepts may be readily understood in terms of weak inhibitory gain, the occurrence of piecemeal mixtures more likely reflects weak inhibitory spatial coherence or weak excitatory lateral connectivity (Alais & Melcher, 2007; Kovács et al., 1996). We repeated the experiment using images of a house and a face as rival targets to establish whether changes of exclusivity also occur with more complex images for which spatial coherence is particularly strong. Again, exclusivity decreased for the major eye-stimulus configuration, but not for the opposite-configuration trials (Figure 7-2B, Spearman rank:  $R_{\text{majority}} = -0.65$ ,  $p_{\text{majority}} << 0.001$ ;  $R_{\text{opp-conf}} = -0.15$ ,  $p_{\text{opp-conf}} = 0.54$ ). An additional control experiment designed to disentangle the relative contributions of superimposition and piecemeal percepts further suggested that decreases in exclusivity are predominantly caused by increases in the incidence of superimposition [Supplementary Information 3, Figure 7-S3].

The opposite-configuration results support the idea that inhibitory connections involved in experience-driven plasticity are at least partially interocular, promoting inhibition between



**FIGURE 7-2.** Prerequisites for decreasing exclusivity. A) The average proportion of exclusive grating percepts over time for five observers. The eye-stimulus configuration was the same for most trials ('Majority trials'; solid black circles), but was switched on some interleaved trials ('Opposite configuration trials'; open squares). B) Similar to A, but here the monocular images were complex pictures of a house and a face, not orthogonal sinusoidal gratings. Dashed lines represent baseline exclusivity; Statistics asterisks,  $p < 0.05$ ; All error bars, s.e.m.

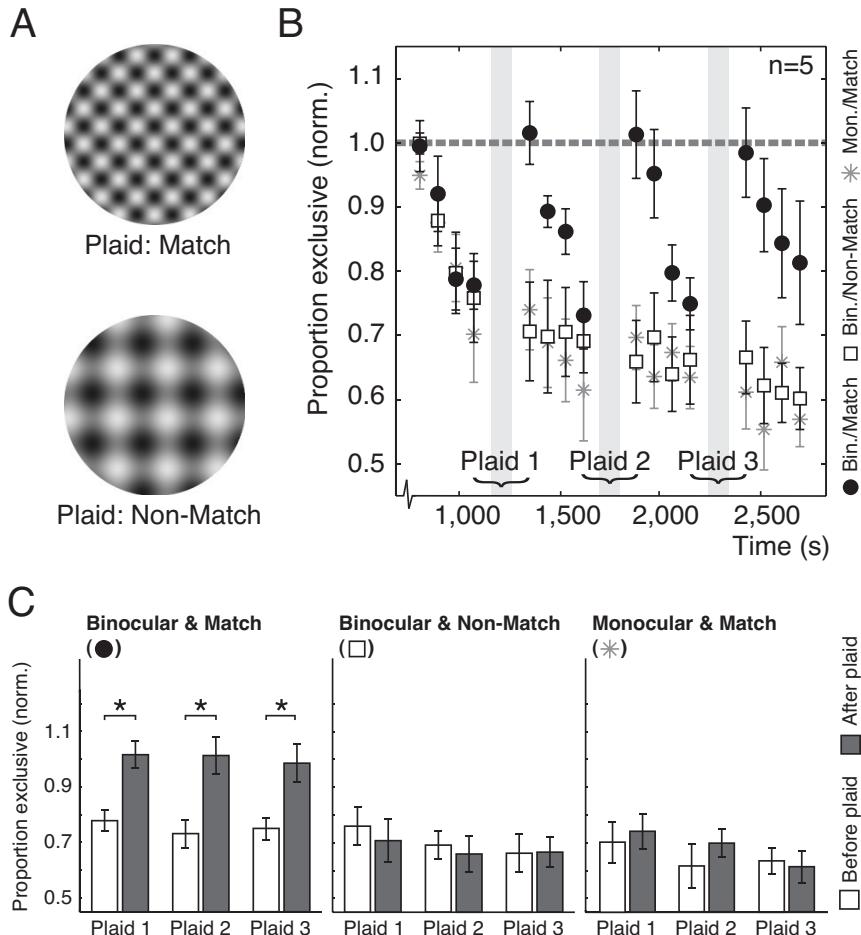
representations of rivaling images comprising eye-of-origin information. If eye-of-origin information were not involved, the stimuli on the majority of trials and the opposite-configuration trials should be equivalent and yield equivalent results. The eye-specificity is consistent with current thinking about binocular rivalry as a hierarchical process involving multiple stages of visual processing (Blake & Logothetis, 2002; Wilson, 2003).

### 7.2.3 Experiment 3: Recovery of perceptual exclusivity

We next investigated the requirements for re-strengthening of inhibition more closely. When binocular free viewing in Experiment 1 caused recovery, both eyes received matching naturalistic input containing a broad range of orientations and spatial frequencies, presumably including those of our rivalry targets. To investigate whether recovery merely requires binocular correspondence or specific binocular correspondence of the rivaling stimulus features, we performed experiments in which we replaced our rivalry gratings with a high contrast plaid stimulus on every fifth trial. "Matching" plaids with the same spatial frequency and orientations as the rivaling gratings (Figure 7-3A) were presented to two eyes simultaneously (Figure 7-3B, black circles) or one eye at a time, alternating between eyes every few seconds (Figure 7-3B, gray asterisks). Plaids with different spatial frequency and orientations (Figure 7-3A) were also presented to two eyes simultaneously (Figure 7-3B, white squares). Figure 7-3C compares the exclusivity levels between trials directly preceding and following plaid presentations. Only binocularly presented, matching plaids evoked a significant recovery of exclusivity (Paired t-tests,  $p_{\text{Bin}/\text{Match}} << 0.001$ ;  $p_{\text{Mon}/\text{Match}} = 0.10$ ;  $p_{\text{Bin}/\text{NoMatch}} = 0.35$ ) supporting the hypothesis that re-strengthening of inhibition only occurs during coinciding activity of eye-specific, orientation- and spatial frequency-tuned neurons. It also argues against an alternative hypothesis of reduced exclusivity through contrast adaptation. In principle, such adap-



tation might reduce exclusivity by lowering the activity of suppressing neurons. However, during presentation of matching plaid stimuli, when the same stimulus features are present as during rivalry, contrast adaptation should continue causing exclusivity to further reduce, not recover like we observed.

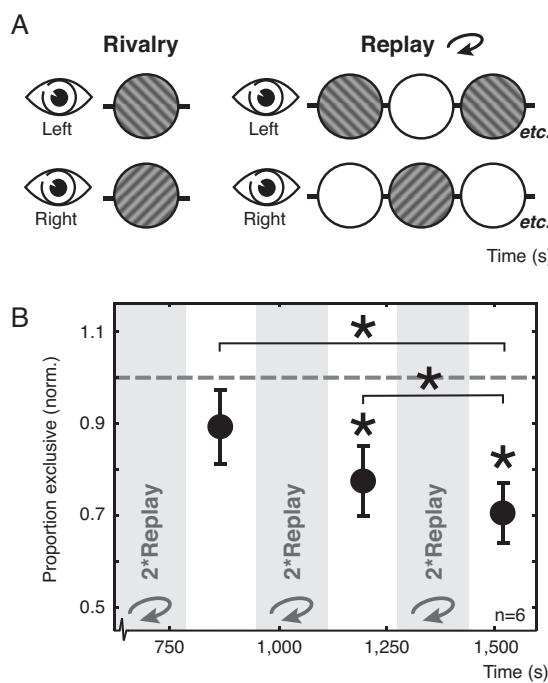


**FIGURE 7-3.** Prerequisites for recovery of decreased exclusivity. **A)** The plaid stimuli that were interleaved with rivalry trials. Matching plaids had the same components as the rivaling gratings while non-matching plaids' components had different spatial frequencies and orientations. **B)** The average proportion exclusivity over time for five observers. Rivalry trials were interleaved with plaid presentations (gray areas). Matching plaids were presented to two eyes (solid black circles) or one eye (gray asterisks) simultaneously. Non-matching plaids were always presented to two eyes (open squares). The dashed line represents baseline exclusivity. **C)** Exclusivity, compared between rivalry trials that directly preceded (white bars) and followed (gray bars) plaid presentation. Statistics asterisks,  $p < 0.05$ ; All error bars, s.e.m.



#### 7.2.4 Experiment 4: Replay-Rivalry

Our results suggest that prolonged binocular rivalry weakens interocular inhibition through recalibration of binocular integration mechanisms in response to cumulative experience with non-fusible input. If such experience-driven binocular plasticity is a generic property of visual perception, the choice for rivalry stimuli should not be essential. Monocular, non-rivaling, stimulation might also weaken inhibition if it activates neurons corresponding to one eye without simultaneously activating their counterparts belonging to the other eye. We tested this prediction using the reported percept durations of baseline rivalry trials to create ‘replay-trials’ where individual eyes were alternately presented with their corresponding monocular images. This manipulation provides the required activity patterns without evoking rivalry (Figure 7-4A). Observers viewed three sets of two monocular replay-trials followed by a single rivalry trial to measure exclusivity. The decreasing exclusivity following replay-trials depicted in Figure 7-4B (t-tests,  $p < 0.05$ ) favors an interpretation where experience-driven plasticity is not restricted to rivalry, but serves as a generic principle of binocular vision.



**FIGURE 7-4.** The effect of monocular replay-rivalry. **A)** While during rivalry both eyes are simultaneously stimulated with conflicting images, replay-rivalry consists of alternating monocular stimulations with a temporal structure based on individual perceptual reports during real rivalry. **B)** The average proportion of exclusivity in three rivalry trials that were each preceded by two replay trials (gray areas). The dashed line represents baseline exclusivity; Statistics asterisks,  $p < 0.05$ ; Error bars, s.e.m.



## 7.3 Discussion

Experience literally changes our view of the world. Neuronal processes converting retinal images to conscious perception constantly adapt to changing sensory contexts. Our results demonstrate that upon prolonged exposure to binocular rivalry stimuli, the nature of the accompanying perceptual experience progressively changes. Where observers initially perceive mostly alternations between exclusive representations of monocular images, mixtures of the two images (Hollins & Hudnall, 1980) become more prevalent over time. Building upon the idea that binocular rivalry involves inhibition between neuronal populations representing competing images (Alais & Blake, 2005; Klink et al., 2008b; Wilson, 2007), we suggest that the rise in predominance of mixed percepts is caused by weakening of inhibitory efficacy (Hollins, 1980).

### 7.3.1 Anti-Hebbian plasticity

A theoretical framework for inhibitory plasticity in vision has been constructed around so-called ‘anti-Hebbian’ inhibitory synapses (Barlow, 1990; Barlow & Földiák, 1989). Hebbian synapses are well known as a neuronal principle for experience-driven plasticity. The basic idea is that when a presynaptic excitatory neuron participates in successfully activating a postsynaptic neuron, their synaptic bond is strengthened and the correlation between their response patterns increases. While there is abundant biological evidence for Hebbian learning in synapses mediating excitatory interactions (Caporale & Dan, 2008; Sejnowski, 2003), the related principle for inhibitory connections has received far less attention. Extending Hebb’s postulate, Barlow and Földiák have proposed that inhibitory interactions are similarly strengthened and weakened by coinciding pre- and postsynaptic activity or a lack thereof (Barlow & Földiák, 1989). Since such a plasticity-scheme decorrelates pre- and postsynaptic activity, it is sometimes dubbed ‘anti-Hebbian’ (Földiák, 1990)(a term also used for several other decorrelating synaptic mechanisms (Nelson, 2004)). Anti-Hebbian plasticity is inherent in several models of unsupervised neuronal learning (e.g. Deco & Obradovic, 1995; Földiák, 1990) and an indirect route via inhibitory interneurons has been physiologically demonstrated in several species and brain structures (Bell et al., 1997; Lamsa et al., 2007; Tzounopoulos et al., 2004; Yazaki-Sugiyama et al., 2009). However, plasticity rules for direct inhibitory synapses appear to be more variable (Caporale & Dan, 2008) and while such anti-Hebbian learning has been suggested in the context of contingent visual after-effects (Barlow, 1990; Barlow & Földiák, 1989), pattern-adaptation (Carandini et al., 1997) and center-surround suppression (Falconbridge & Badcock, 2006), direct behavioral evidence is sparse.

Our binocular rivalry results are consistent with anti-Hebbian learning mechanisms for interocular inhibition in binocular vision. Assuming that perceptual dominance of one rival image indicates successful suppression of the competing neuronal representation, dominance may entail activity in presynaptic neurons representing the dominant image without equivalent activity in the postsynaptic neurons encoding the (suppressed) opposite image. These are exactly the conditions for which anti-Hebbian weakening of inhibitory efficacy was postulated, explaining why initially high perceptual exclusivity should progressively decrease with viewing time. Furthermore, the anti-Hebbian principle predicts that (re-)strengthening of inhibition would require simultaneous activation of the same neurons involved in rivalry. This can arguably be achieved by presenting binocularly corresponding stimuli with features



to which those specific neurons are tuned. Our experiments demonstrate both the predicted drop in perceptual exclusivity and the expected dependence of recovery on stimulus features.

### 7.3.2 Plasticity and rivalry

Previously demonstrated changes in perceptual experience with prolonged or repeated rivalry include short-term slowing of perceptual switch-rates during single binocular rivalry trials (Suzuki & Grabowecky, 2007; van Ee, 2005) and long-term speeding of switch rates when sessions are repeated over many days (Suzuki & Grabowecky, 2007). While short-term effects were explained by contrast adaptation build-up (Suzuki & Grabowecky, 2007; van Ee, 2005), long-term effects were attributed to plasticity in neuronal responses and/or connectivity within multiple brain areas (Suzuki & Grabowecky, 2007). Because none of the abovementioned studies included the dynamics of mixture percepts in their binocular rivalry evaluation, it is difficult to unify the changes in switch-rate with our changes of binocular integration. However, one emerging notion is that the adult visual system seems more plastic than previously realized and future studies of binocular rivalry need to appreciate that exposure to rival stimuli may cause plastic changes in the very neuronal mechanisms targeted for study.

The many similarities and differences in the dynamics of binocular rivalry and other forms of perceptual rivalry (Blake & Logothetis, 2002; Brascamp et al., 2005; Klink et al., 2008a; Klink et al., 2008b; Pearson & Clifford, 2005; van Ee, 2005) have promoted the idea that different types of rivalry, while perhaps resolved at different processing stages, may share common computational components in their rivalry-resolving mechanisms (Klink et al., 2008b). Since mutual inhibition is conceivably one of those components (Klink et al., 2008b), it would be interesting to know whether plasticity of inhibitory efficacy also influences other forms of rivalry. The reduced exclusivity observed in our study proved to be specific to eye-stimulus configuration, locating the proposed plasticity mechanism at a stage of binocular rivalry processing that includes eye-of-origin information. (Blake & Logothetis, 2002; Wilson, 2003). Still, this does not entirely preclude the possibility of inhibitory plasticity in other forms of rivalry or at other processing levels. Furthermore, it implies that plastic interocular inhibition may be a general mechanism of binocular vision, raising the intriguing question what might happen if exposure to rival stimulation were prolonged for hours or days, impractical though it might be to find out.

### 7.3.3 Conclusions

Our findings suggest experience-driven (anti-Hebbian) plasticity as one governing principle in the neuronal organization of binocular vision. It is tempting to envision this mechanism as a means for interocular gain-control during binocular combination. It could balance monocular signals so that binocular contrast and surface lightness are not noticeably different from monocular viewing (Ding & Sperling, 2006). On this view, our binocular rivalry experiments reveal the operation of such inhibitory mechanism and its dynamic modification. The experience-driven plasticity we demonstrate may provide important clues towards solving the longstanding question of how rivalry and stereopsis can emerge from a single neuronal organization of binocular vision (Andrews & Purves, 1997; Blake et al., 1991; Ding & Sperling, 2006; Grossberg et al., 2008; Livingstone, 1996).



## 7.4 Experimental procedure

Observers viewed stimuli through a mirror stereoscope in a quiet, darkened room. Rival stimuli were surrounded by an alignment-ring to facilitate binocular fusion. Observers continuously reported perceptual experience by pressing buttons on a keyboard. One of two buttons was held while observers exclusively perceived the corresponding monocular stimulus. Both buttons were released when mixtures were perceived. The basic experimental paradigm consisted of a baseline determination followed by two stages differing in the timing of stimulus presentation. During baseline determination, individual observers' levels of exclusivity were established with stimulus presentations lasting 100 s, separated by 100 s rests during which observers viewed the alignment frame only. During the first part of the actual experiment, stimulus presentations also lasted 100 s, but rests were reduced to 10 s. In Experiment 1, a second part comprised stimulus presentations of 60 s and rests of 300 s. These long rests consisted of 240 s of predefined visual input (depending on the condition) and 60 s of uniform field adaptation during which observer's viewed a gray screen. For all rivalry trials we calculated proportions of exclusivity as the sum of all exclusive percept durations divided by the total trial duration. These proportions were normalized by the average baseline proportion of exclusivity for each observer. A more detailed description of the Experimental Procedure is available as Supplemental Experimental Procedure.

## 7.5 Acknowledgements

We thank our observers for their perseverance during these long, demanding experiments, Helmut Kröger for sharing his code for our model simulations, and Sidney Lehky, Martin Lankheet, Tomas Knapen and André Noest for commenting on earlier versions of this manuscript. This work was supported by a VIDI (RvW) and a Rubicon grant from the Netherlands Organisation for Scientific Research (JB), an Utrecht University High Potential grant (RvW), NIH grant EY13358 (RB), and a grant (R32-10142) from the World Class University program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (RB).



## 7.6 Supplemental experimental procedure

### 7.6.1 Stimuli

Observers viewed stimuli through a mirror stereoscope to allow visual stimulation of each eye independently. Experiments were run on a dedicated Mac Pro using Matlab (the Mathworks, Natick) and the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli were presented on a gamma corrected 22" CRT monitor that was placed at a viewing distance of 81.3 cm and ran at a resolution of 1280×1024 pixels with a refresh rate of 100 Hz. All stimuli were  $1.5 \times 1.5^\circ$  in size and were presented on a uniform gray background. Both stimuli and background had an average luminance of 12 cd/m<sup>2</sup>. Stimuli were surrounded by a dichoptically presented alignment ring (50 cd/m<sup>2</sup>) to promote binocular fusion. Basic stimuli were sine-wave gratings with a spatial frequency of 5 cycles/°, orientations of +45° and -45° (for the two eyes respectively) and a luminance contrast of 30%. When we used house-face stimuli, their average luminance was also ~12 cd/m<sup>2</sup> with RMS contrast of 50%. Plaid stimuli were obtained by taking the pixel-based average luminance of two superimposed gratings with 75% contrast. In the 'matching-plaid' condition the component gratings had a spatial frequency of 5 cycles/deg and orientations of +45° and -45°. In the 'non-matching-plaid' condition the component gratings had a spatial frequency of 2.5 cycles/° and orientations of 0° and 90°.

### 7.6.2 Procedure

Observers reported perception using two buttons on a standard keyboard. Either of the two keys was held when an observer perceived one particular monocular stimulus as exclusively dominant. Both keys were released if they perceived a mixture of the two monocular stimuli. For every rivalry trial we recorded the percept durations for the three possible percepts. From those perceptual epochs we calculated the proportion exclusivity as the sum of mixed-percept durations divided by the trial duration. There is a large variability in absolute proportions of baseline exclusivity between observers (e.g. in Experiment 1, baseline values varied between 0.52 and 0.92, with a mean of  $0.74 \pm 0.03$  s.e.m.). To correct for this variability, we normalized the exclusivity levels for individual observers on all trials by dividing it by their average proportion of exclusivity on baseline trials.

### 7.6.3 Experiment specifics

Experiment 1 consisted of two parts preceded by a baseline measurement. During the baseline measurement, we presented stimuli for 100 s separated by 100 s interstimulus interval (ISI) when only the alignment ring was presented on the screen. This procedure was repeated 4 times and the average performance on these trials was taken as the starting-level of exclusivity for individual observers. The first part (dubbed the 'Prolonged rivalry stage') of the actual experiment consisted of 19 rivalry trials that also lasted 100 s, but were separated by shorter ISIs of 10 s. During the last part of the experiment (dubbed the 'Recovery stage'), 8 trials of 60 s duration were each preceded by an ISI of 5 minutes. Of those 5 minutes, the first 4 minutes were different for three different experimental conditions: 1) Observers walked around the hall with both eyes open, thus receiving 'naturalistic' binocular visual input. 2) Observers walked around the hall wearing an eye-patch that blocked out visual input to one eye. The eye that was covered by the eye-patch was alternated on consecutive ISIs so that each eye



received the same total amount of visual input. 3) Observers remained in the experimental room wearing an eye-mask that covered both their eyes to prevent visual input. A sound fragment warned them that the experiment needed to be continued. The last minute of the ISI was the same for all conditions and consistent of a period of light adaptation during which observers viewed the uniform gray background only.

Experiment 2 contained three variations on a common theme. There were always 6 baseline trials (ISI 100 s) and 20 experimental trials (ISI 10 s). There was no second experimental stage with long ISI's. The stimulus-eye configuration was switched on every 2<sup>nd</sup> (baseline) or 5<sup>th</sup> (actual experiment) trial, so that the same stimuli were presented to the opposite eyes. In the first variation, stimuli were gratings. The second variation used images of a house and a face.

In Experiment 3 there were 4 baseline trials, followed by 16 experimental trials. The distinctive element for three experimental conditions was the presentation of a plaid stimulus on every 5th experimental trial. 'Matching plaids' (same spatial frequency and orientation as monocular gratings) could be presented either monocularly (switching eye every 10 s to balance total eye stimulation) or binocularly. 'Non-matching plaids' were always presented binocularly.

Experiment 4 contained 3 baseline trials and 9 experimental trials. Every 3rd experimental was a rivalry trial, while the others were 'replay-rivalry' trials. To create replay conditions, all epochs of exclusive perception obtained during baseline trials were saved in an eye-specific list. On every replay trial, presentation durations were taken from a random permutation of this list of percept durations and images were presented to the correct eye in alternating fashion while observers reported perception.

#### 7.6.4 Observers

All observers had normal or corrected to normal visual acuity and some experience with psychophysical binocular rivalry experiments. They were selected for having a reasonably fast perceptual switch-rate (in the order of one switch every few seconds) and no extreme eye-dominance. After the procedure was explained to them and examples were shown, they gave informed consent and confirmed that they understood the instructions. Five observers participated in experiments 1 to 3. Six observers participated in experiment 4. In experiments 1 and 2, two observers were authors (CK & JB), while in the other experiments only one author (CK) participated. All observers, except the authors, were naive with respect to the goal of the experiment.

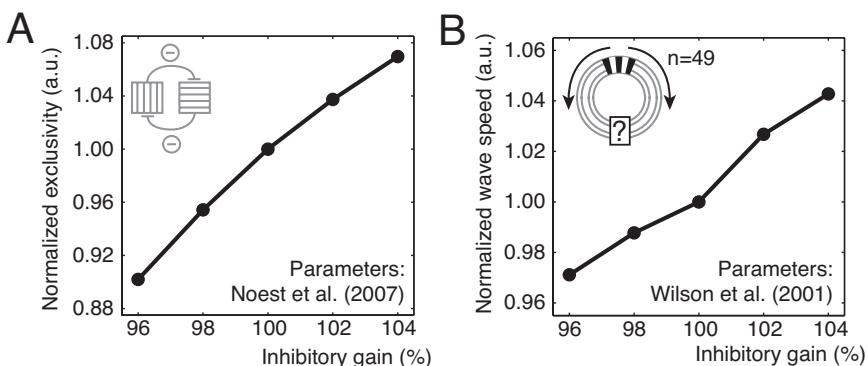


## 7.7 Supplemental information 1: Modeling data

In our experience-driven learning account, the strength of inhibition in binocular rivalry is dynamic and dependent on recent sensory input. It would be interesting to see whether manipulations of inhibitory parameters in existing models of binocular rivalry affect the model predictions in any way that is consistent with our experimental data. A division between non-spatial and spatial models of binocular rivalry may provide insight in the occurrence of superimposition and piecemeal percepts respectively.

Non-spatial models generally assume that the most active of two rivaling monocular populations determines perception. The difference in activity between those populations may be regarded as an indication for the level of exclusivity of perception, i.e. if one population is very active while the other is silent, perception is likely to be exclusive, but if both populations are active and the difference between their activity levels is relatively small perception could be more mixed. We chose two prominent rivalry models from recent literature to test the influence of inhibition on perceptual exclusivity: A non-spatial rivalry model published by Noest et al. (2007) and a spatial model published by Wilson et al. (2001). Using all parameter values that were mentioned in their original publications, we systematically manipulated the strength of inhibition (non-spatial model) or the spatial profile of inhibition (spatial model).

For the non-spatial model we defined exclusivity as the average difference in activity between the two simulated rivalry populations divided by their mean activity level. Simulation results for an altered inhibitory gain were compared to simulation results for the original parameter-set (inhibitory gain is 100%). Consistent with our experimental data, increasing the gain of inhibition resulted in higher levels of exclusivity, whereas decreasing the gain caused lower levels of exclusivity (Supplemental Figure 7-S1A). Wilson et al.'s spatial binocular rivalry model describes the dynamics of traveling waves of perceptual dominance (Lee et al., 2005). The transition from one percept to the other starts at a random point on the



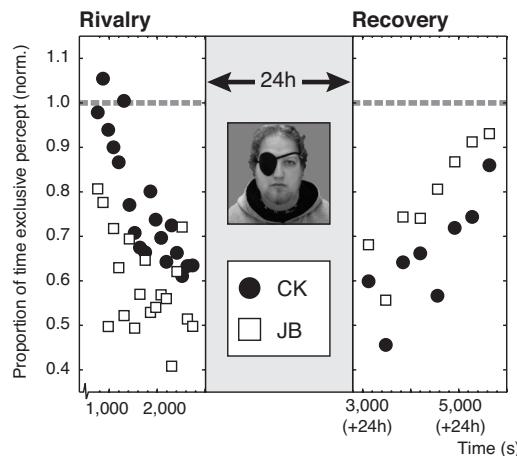
**FIGURE 7-S1.** The strength of inhibition in existing rivalry models. **A)** Simulations with a non-spatial rivalry model (Noest et al., 2007) in which we manipulated the strength of inhibition, while all other parameters were kept the same as in the original publication. **B)** Simulations with a traveling wave model of binocular rivalry (Wilson et al., 2001) in which we manipulated the spatial influence of inhibition, while all other parameters were kept the same as in the original publication.



perceived stimulus and then spreads in a wave-like fashion across it. A perceptual transition is complete when the wave has engulfed the stimulus entirely, so the speed at which this wave travels is an indication for the duration of this type of transition between two exclusive percepts. Faster waves cause shorter transitions and thus higher levels of exclusivity. We manipulated the spatial extent of inhibition and compared our results to the model prediction with the original parameter-set. Simulations of a dominance wave traveling over a ring of 49 neurons (mimicking Wilson's spatial layout) confirm previous findings (Knapen et al., 2007b) that demonstrated an increase in wave speed due to increased spatial influence of interocular inhibition (Supplemental Figure 7-S1B). Since faster traveling waves result in higher proportions of exclusivity, these findings are also very well compatible with our interpretation of the current experimental data.

## 7.8 Supplemental information 2: The 24-hour experiment

Experiment 1 already demonstrated that monocular visual input during repeated five-minute episodes was incapable of restoring the proportion of exclusivity back to baseline values (Figure 7-1B). This might mean that binocular input is needed, but an alternative explanation would be that the episodes of monocular stimulation were simply too short. To evaluate this possibility, we extended the episode of monocular input to 24 hours (minus approximately 7 hours of sleep). The two first authors (CK & JB) performed a variation of experiment 1 in which they started wearing an eye-patch immediately after the prolonged rivalry viewing in the first part of the experimental paradigm and finished the last part of the experiment, featuring rivalry trials separated by binocular free viewing, 24 hours later. During these 24 hours



**FIGURE 7-S2.** The 24 hours eye-patch experiment. The proportion of exclusive percepts over time for two individual observers (CK, solid black circles; JB, open squares). The experiment was similar to Experiment 1, but after the first stage, observers wore an eye-patch for 24 hours (gray area) before finishing the experiment. During the final part of the experiment, both eyes received 'naturalistic' binocular input during the ISI's (see also Figure 7-1B, solid black circles). The dashed line represents baseline exclusivity.



they continuously wore the eye-patch. After the 24 hours with the eye-patch the final stage of the experiment was started with 15 minutes of uniform field adaptation to allow the previously covered eye to adapt to the luminance of the presentation screen. The decrease of exclusivity during the first part of the experiment had been highly significant (Supplemental Figure 7-S2, Spearman rank correlation,  $R = -0.51$ ,  $p < 0.001$ ), matching the results of experiment 1. Remarkably, the decreased proportions of exclusivity hardly changed over the 24 hours of wearing an eye-patch, but immediately started to increase when the last experimental stage commenced with episodes of ‘naturalistic’ binocular input (Supplemental Figure 7-S2, Spearman rank correlation,  $R = 0.79$ ,  $p < 0.001$ ).

## 7.9 Supplemental information 3: Superimposition vs. piecemeal

In our measurements different types of mixture percepts are lumped together, but their individual predominance may rely on different aspects of the suppressive inhibition. Superimposition percepts are plausibly understood in terms of weakened mutual inhibition, but piecemeal percepts may require a somewhat different explanation. The local spatial zones of monocular image dominance in piecemeal percepts plausibly reflect a weakening of the spatial coherence of inhibitory connections or the extent of the excitatory lateral connectivity responsible for perceptual grouping (Alais & Melcher, 2007; Kovács et al., 1996; Lee & Blake, 2004).

The absolute predominance of superimposition and piecemeal percepts strongly depends on stimulus features (Blake, 2001), such as stimulus size (Blake et al., 1992), spatial frequency (Hollins, 1980; O’Shea et al., 1997; Yang et al., 1992) and global context (Kovács et al., 1996). We performed a control experiment to establish whether the reduction of the proportion exclusive percepts that we see with our specific binocular rivalry stimuli is due to changes in the predominance of superimposition percepts, piecemeal percepts, or both. Four observers (one author) participated in the control experiment. They had also participated in the main experiments but apart from the author, were naïve to the purpose of this control experiment. Observers viewed 15 binocular rivalry trials that each lasted 60 seconds. During these trials they kept a button pressed down whenever they experienced exclusive dominance of either monocular image. After every trial they reported whether the mixture percepts in that trial were predominantly piecemeal or superimposition percepts by setting a slider that was presented on the screen and ranged from 100% piecemeal to 100% superimposition in 60 steps. All observers performed this control experiment both with the grating stimuli and the house and face images.

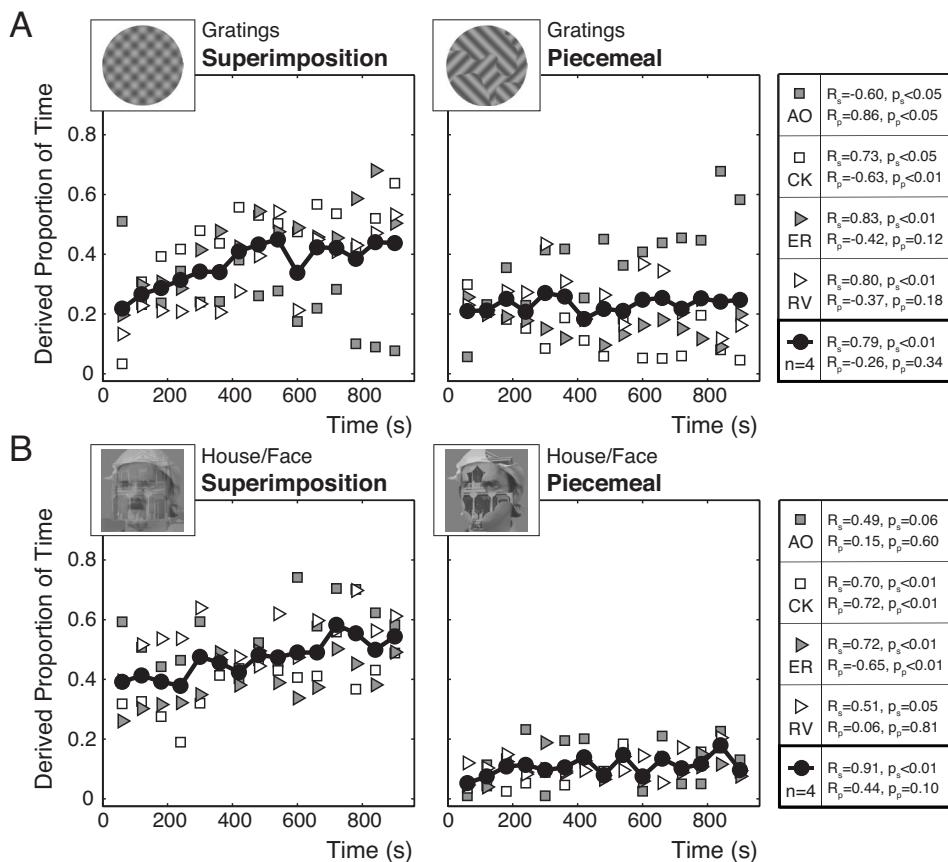
We analyzed whether there were any time-dependent changes in the proportion exclusive percepts and the observers’ rating of the mixture percepts. Since we knew both the proportion of time an observer experienced mixture percepts on every trial and the observer’s rating of the relative proportions of the mixture percepts that should be attributed to superimposition and piecemeal, we were able to derive an indirect measure of the proportions of both mixture percept types for every trial by multiplying these values. All observers experienced a significant decrease in the proportion exclusive percepts over time for both stimulus types (Spearman rank correlations,  $p < 0.05$  for all observers). For house/face images, observers experienced much more superimposition percepts than piecemeal percepts ( $81.4\% \pm 8.1\%$



standard deviation of the mixture percepts should be attributed to superimposition on average) and this dominance of superimposition remained constant over time (Spearman rank correlation,  $R = -0.23$ ,  $p = 0.42$ ). As can be seen in Figure 7-S3B, the derived proportion of superimposition significantly increased over time for all observers (individual symbols, squares and triangles) and the group average (thick lines and circles), while the proportion of piecemeal did not change at all (and even decreased a little for one observer). Statistical values of a Spearman rank correlation analysis are displayed in the legends. This selective increase in superimposition without increases in piecemeal suggests that for these house/face stimuli the drop in exclusivity is most likely resulting from a decrease in inhibitory gain.

The results from the grating stimuli tell a slightly different story. The overall dominance of superimposition was much weaker ( $60.6\% \pm 22.3\%$  standard deviation over all trials) and more variable over observers. For three of the four observers both the dominance of superimposition over piecemeal percepts and the derived proportion of superimposition increased over time while the derived proportion of piecemeal remained constant (Figure 7-S3A). One observer (AO, gray squares), however, showed the opposite pattern with piecemeal percepts increasing both in dominance and derived proportion over time, while the derived proportion of superimposition percepts significantly decreased.

The results of this control experiment imply a dominant role for superimposition percepts (likely related to inhibitory efficacy) in the demonstrated changes of perceptual exclusivity upon prolonged binocular rivalry viewing. While this may be true for the specific stimuli used here, the known dependence of superimposition and piecemeal percepts on stimulus features as well as the deviating results of one of our observers (AO) suggest that piecemeal percepts (plausibly related to spatial coherence of inhibitory connectivity) may also be important for some observers, or under different circumstances.



**FIGURE 7-S3.** Superimposition vs. piecemeal. From the overall proportion of exclusive percepts and the observers indication of the relative dominance of superimposition over piecemeal we calculated the derived proportion of time that superimposition and piecemeal percepts were dominant. Results with both grating stimuli A) and house/face images B) are plotted for individual observers (squares and triangles) and the group average (circles and thick line). The results of Spearman rank analyses of the data, testing whether the plotted proportions change over time, are displayed in the legends.



# Chapter 8

## Appendix

### Some spikes are more informative than others

*All animals are equal, but some are more equal than others.*

**George Orwell**

#### Published as

Klink, P.C. (2008). Some spikes are more informative than others.  
*Journal of Neuroscience*, 28(19), 4844-4845

#### A Journal Club contribution, based on

Masse, N.Y. & Cook, E.P. (2008). The effect of middle temporal spike phase on sensory encoding and correlates with behavior during a motion-detection task.  
*Journal of Neuroscience*, 28(6), 1343-1355



*This appendix chapter takes a slight deviation from the general theme of this thesis. It does not describe any experiments we performed ourselves, but instead highlights a recent finding by others. This finding is however very relevant for any neurophysiological study of sensory perception since it addresses one of neuroscience's fundamental questions: How is information transferred among neurons? The lack of conclusive answers to this question presents the neurophysiologist with an unavoidable data analysis problem. If we want to decode information from neural data, it would be very helpful to know how this information was encoded in the first place. The paper that is discussed in this chapter demonstrates that not every action potential in a neuronal response is equally informative about the corresponding perception and/or behavior. This essentially means that neuronal information is inhomogeneously distributed, not only over cells, but also over specific action potentials within a single cell's activity pattern. This notion has inspired me to analyze the neurophysiological data in Chapter 4 not only in terms of rates, but also in terms of spike timing. My interest in the mechanisms of neural encoding and decoding has further grown during recent neurophysiological recordings from the catfish electroreception system that I performed together with Martin Lankheet. The report on the findings of that study, addressing the interplay between information encoding and spike generation mechanisms, is not included in this thesis, but will hopefully appear in the neuroscience literature soon.*

---



What is the language of the brain when it links perception and behavior? Studies in the past have suggested that there are subsets of ‘most informative’ neurons for perception and behavior (for a review: Parker & Newsome, 1998) and recent findings have revealed that perception and behavior can both be induced with the stimulation of only a small number of neurons (Houweling & Brecht, 2008; Huber et al., 2008). Well-known examples of selective subsets of neurons are found in the direction-, disparity- and speed-tuning properties of cortical motion-sensitive neurons in middle temporal area (MT) of the monkey brain (for a review: Born & Bradley, 2005). As a popular cortical area to investigate the link between perception and action, neuronal activity in area MT has been studied using different analytic measures such as spike-rates, local field potentials, oscillations and response latencies, but the information density of the neuronal activity is usually analyzed between, rather than within single neurons. The known dissociation between more and less informative neurons, however, raises the question whether the amount of information carried by the neural activity within a single neuron could also be nonuniformly distributed. A recent study by Masse and Cook (2008), published in the *Journal of Neuroscience*, addresses this question and demonstrates that for a motion detection task, some spikes generated by neurons in MT are indeed more informative about the stimulus content than others. Importantly, the more informative spikes are also more closely related to the monkey’s response behavior. These novel findings suggest a relation between spikes, stimuli and behavior that varies on a relatively short timescale within a single neuron.

Masse and Cook (2008) used datasets from several studies in which the activity of single neurons in area MT was recorded while monkeys detected the onset of coherent motion in a dynamic random dot pattern. Importantly, the motion stimulus was updated relatively slowly (every 27 ms) causing the neurons to discharge in an oscillatory manner with a frequency that matched the stimulus updates. Based on this oscillatory firing pattern, the authors subdivided the neuronal responses based on whether they occurred during the rising or falling phase of the oscillation (Figure 2 in Masse & Cook, 2008). Using spike-triggered averages (the average stimulus content preceding a spike), the authors demonstrated that spikes during one of the phases (often the rising phase) were more informative about the presence of motion in a neuron’s preferred direction than spikes during the other phase. To link these findings to behavior, correlations were calculated between the spike-triggered average and two measures based on receiver-operant characteristics (ROC): neurometric value and detect probability. These measures indicate to what extent an ideal observer would be able to distinguish coherent and random motion (neurometric value) or different behavioral responses (detect probability) based solely on the neuronal data.

Correlations between neurometric value and detect probability revealed that the spikes that were most informative about the presence of motion in the neuron’s preferred direction were also more informative about the coherence of motion and more predictive of the monkey’s behavioral response. Interestingly, the activity during the most informative encoding phase was correlated with the animal’s correctly reporting the onset of coherent motion, whereas activity during the weaker encoding phase was correlated with the animal’s failing to detect coherent motion onset (Figure 6 in Masse & Cook, 2008).

Models of perceptual decision making typically involve the accumulation of spike-rate-encoded sensory evidence towards a decision moment (for a review: Gold & Shadlen, 2007). This mechanism assumes that all spikes from a single neuron are similarly related to the stim-



ulus content and the observer's behavioral response. It is unclear how the neural activity that is correlated with failed behavioral responses should be incorporated in such a mechanism. If the integration of sensory evidence towards a perceptual decision would occur upstream from MT, this issue could be resolved with a selection mechanism that processes both positively and negatively correlated spikes in an appropriate manner. Unfortunately, an equally problematic issue directly follows from this solution, since it remains unclear how such an upstream decision area would 'know' whether spikes are positively or negatively correlated with the stimulus content. Searching for the answer to this question in the oscillations of activity might have been promising if the strong encoding phase was always either the rising or the falling phase of the oscillatory neuronal response, but this varies between cells (Figure 5 in Masse & Cook, 2008). This inter-neuronal variation makes the selection mechanism for the most informative spikes even more complex, strongly suggesting that there must indeed be a later neural stage that combines information from multiple neurons.

The oscillatory neural response in the current study is a direct result of a slowly refreshed stimulus and thus to some extent artificial. Generally, the temporal precision of a neural code tremendously increases if spikes are phase-locked to some oscillatory process because it allows a continuous latency code (Butts et al., 2007), but these oscillatory processes do not necessarily have to be this artificial, or even oscillatory. With more natural stimuli, spikes could just as well be phase-locked to the stimulus, to spikes from other neurons, or to internal oscillations of the local field potential (LFP). A relation between the LFP and spikes further has the advantage of offering a direct relation between the input (LFP) and output signals (spikes) of a cortical region.

The information carried in precise spike times is many times larger than that in spike rates, which are typically integrated and averaged over intervals of tens of milliseconds. The information content increases even more when spikes are part of a fixed pattern of relative spike times originating from several synchronized neurons (Tiesinga et al., 2008). Masse and Cook (2008) used classical spike counts in more or less arbitrarily chosen time intervals to conclude that some spikes are more informative than others. A more precise analysis of spike times from synchronously recorded neurons might provide insights into the mechanism by which the information asynchrony arises and tell us more about the brain's decision mechanisms. It is, however, very well possible that the brain has ways of selecting the most informative neurons and best encoding spikes (or spike patterns) that go beyond the detection mechanisms of our current analysis techniques.

Apart from the discussed problems with the current interpretation of Masse and Cook's results, the authors have convincingly demonstrated that even within a single neuron some spikes are more informative than others and apparently the brain relies mainly on these more informative spikes to shape behavior. Unfortunately, it remains unclear how this information asynchrony arises or how the brain integrates the information from multiple neurons. It is up to future research to unravel general rules regarding the context that gives spikes their information value and the neural mechanisms on which the brain bases perceptual decisions and behavior. Whereas our current understanding of the brain's language is sufficient to participate in interesting dialogues, there is clearly still a tremendous amount of neural grammar and semantics to be learned.



## 8.1 Acknowledgements

The author thanks Richard van Wezel, Martin Lankheet and Bert van den Berg for constructive comments on this manuscript.



# **Chapter 9**

---

## **Summary & General discussion**

*It would be so nice if something made sense for a change.*

**Alice, in Walt Disney's adaptation of  
Lewis Carroll's *Alice's Adventures in Wonderland***

**Parts of this *General discussion* are scheduled to appear as invited review article**  
Klink, P.C., van Wezel, R.J.A., & van Ee, R. (In preparation). United we sense, divided  
we fail: Neural mechanisms of contextual inference in multistable perception

*Special issue proposal, submitted to Phil Trans R Soc Lond B,  
provisionally titled "Multistability in perception: Binding sensory modalities"*



## 9.1 Summary & General discussion

The visual system resolves perceptual ambiguities through inference. Ambiguous visual stimuli are sometimes considered to only exist in the laboratory and bear very little relation with vision in the real world. However, while visual ambiguities in natural vision may be relatively scarce at the level of complete patterns or objects, sensory information on a more local scale (both spatial and temporal) is in fact hardly ever unambiguous. Understanding the resolution of ambiguous sensory input on a global scale, might facilitate the unraveling of the, possibly similar, disambiguating mechanisms that operate on a more local scale. The use of contextual influences to study the mechanisms of conscious visual perception allows us to take the ambiguous stimulus approach beyond the status of ‘a window on consciousness’. Extending this common metaphor, clever implementations of context in ambiguous vision experiments let the vision scientist reach through the window on consciousness and actually interfere with the mechanisms that promote conscious visual perception. As any young parent, or any man that has ever played with a new electronic gadget without reading the manual, can tell you: interfering with a process is by far the fastest way to learn how things work. Especially when there is no manual.

In this thesis, several research projects illustrate how a broad range of contexts can provide valuable insights in the neural basis of vision. The main findings that are extensively presented in the different research chapters of this thesis are summarized below, followed by a brief discussion of other recent findings and their relation to the work in this thesis. Finally, a few suggestions will be made for future research that may take the field of vision science into new and exciting directions.

### 9.1.1 *Chapter 2: General rivalry mechanisms*

In Chapter 2 of this thesis, a unified view is presented that aims to find a common functional neuronal basis for binocular and perceptual rivalry. The typical perceptual dynamics during rivalry are hypothesized to be an emergent phenomenon of the basic functional organization of the visual system. Since it seems unlikely that the brain contains a specific ‘rivalry-resolving-module’, it would not be very effective to search for similarities and differences between binocular rivalry and perceptual rivalry in precise neural substrates or anatomical brain structures. One will certainly find differences, but it can hardly be a surprise that different stimuli evoke different response patterns. It also falls well within expectation that similarities will be found. Both cases of rivalry involve visual processing and in both cases there is a similar perceptual problem that needs to be resolved. To compare the two types of rivalry in a more meaningful way, we should focus on the functional aspects that inspired the question of common neural machinery in the first place: both types of rivalry present the brain with sensory information that is equally suitable for multiple perceptual interpretations, yet the brain generates only one percept at a time. In looking for similarities between binocular and perceptual rivalry, it would thus seem most promising to focus on the functional mechanisms and neuronal computations that resolve the visual ambiguities.

For binocular rivalry this has been done excessively over the past decades (Baker, 2010). In theoretical and computational work on the neuronal basis of binocular rivalry, the seminal work of Pim Levelt plays a prominent role (Levelt, 1965). Revolving around four propositions,



this work describes the intricate relationship between stimulus contrast and the perceptual dynamics during binocular rivalry viewing. These propositions were adapted to a more recent set of experimental findings (Brascamp et al., 2006) and rephrased in terms of stimulus bias to make them applicable to perceptual rivalry. The idea is that altering the contrast of one of the two monocular images in binocular rivalry is essentially a manipulation of the strength of one of the two rivaling agents in the visual competition, which is in essence an introduction of a stimulus bias. If binocular rivalry between two images with equal contrast would be regarded as ‘balanced rivalry’, the introduction of any stimulus strength offset between the two monocular images would result in ‘biased rivalry’. Under such interpretation, it is relatively straightforward to apply Levelt’s proposition to perceptual rivalry. All that is needed is a method of biasing a bistable stimulus towards one of the two perceptual interpretations without completely negating the ambiguity.

Using a bistable, rotating structure-from-motion (SFM) sphere stimulus, composed of white dots on a black background a depth order bias was introduced by manipulating the luminance of the dots. Surprisingly, all four classic propositions turned out to be just as applicable to perceptual rivalry as they are to binocular rivalry. However, this result is not interpreted as evidence for a generic neuronal rivalry resolver. Existing evidence clearly demonstrates that it depends on the precise stimuli and task which parts of the brain are involved in rivalry. However, within these different neuronal circuits, the rivalry may be resolved through similar rules of neuronal computation. Such an interpretation basically means that the relevant neuronal operations involved in the different types of rivalry may be generalized. This opens the door for more general modeling approaches and promotes the fusion of two parallel lines of research that both aim to understand how conscious percepts arise from complex neuronal interactions. Most importantly, it stresses the urgency of a functional view on visual rivalry.

### 9.1.2 *Chapter 3: Temporal context and voluntary control*

*Temporal context.* The research in Chapter 3 uses both binocular rivalry and perceptual rivalry (SFM cylinders) to demonstrate the roles of temporal context and attention on the conscious perception of ambiguous visual stimuli. Methodologically, it demonstrates the strength of combining psychophysical experiments with computational models, based on existing neurophysiological knowledge. A previously proposed, low-level neural model explaining the percept-choice dynamics of repeatedly presented ambiguous stimuli (Noest et al., 2007) was extended in a biologically plausible way to interpret a new set of data.

The original model describes a single rivalry resolving stage of neural processing and predicts that the perceptual dynamics for intermittently presented ambiguous stimuli crucially depend both on the duration of stimulus presentations and on durations of stimulus removal. Our experiments partially confirm this prediction by demonstrating that short interruptions result in a high probability of perceiving alternative stimulus interpretations on subsequent trials, and long interruptions result in a high probability of perceiving the same interpretation over and over again. While the effects of stimulus presentation duration were much weaker than those of interruption duration, the difference in presentation duration effects between the types of rivalry do suggests that the visual competition in binocular rivalry is resolved at an earlier stage of cortical processing than the competition involved in the perceptual rivalry



of bistable structure-from-motion.

*Attentional context.* The effects of endogenous attention were tested in similar experiments as the effects of temporal context. Also described in Chapter 3, these experiments demonstrate that the clear effects of stimulus timing persist when observers were asked to voluntarily control their perception during the intermittent presentation of ambiguous stimuli. However, voluntary control did have a limited, yet striking, effect on perception. These effects were consistent with a gain control mechanism at very early stages of visual processing. Such a gain control mechanism resembles the effective contrast-enhancing effects known from many studies on visual attention (Reynolds & Chelazzi, 2004).

### 9.1.3 Chapter 4: Neurophysiology of temporal context

A neurophysiological study, described in Chapter 4, further investigated the temporal effects of interrupted, sequential stimulus presentations at the level of the single neuron. Responses were recorded from motion-sensitive single units in two monkeys, while they were viewing sequences of ambiguous and non-ambiguous motion stimuli with a variable blank duration between stimulus presentations. For all stimuli, a striking stabilization of response patterns was found. This stabilization critically depended on the blank duration between stimulus presentations. Both the spike-count variability and the variability in spike timing decreased with increasing blank durations. This effect was shown to exist both for ambiguous and non-ambiguous stimuli, which indicates that response stabilization may be a generic neural mechanism that takes place whenever neurons are repeatedly stimulated. The hallmark characteristic of multiple possible perceptual interpretations with ambiguous stimuli could magnify this effect at the perceptual level. Local field potential data, simultaneously recorded with the action potentials, provided an indirect measure of the activity patterns in the local cortical networks that surrounded the single neuron from which the action potentials were recorded. Blank duration dependent power increases in the high Gamma range suggest that response stabilization in single neurons is driven by increased coherence within the local cortical network.

### 9.1.4 Chapter 5: Spatial context

The effects of spatial context are addressed in Chapter 5. It is first established how luminance and disparity cues individually stabilize the perceptual interpretation of an ambiguous, rotating SFM cylinder. In an additional set of experiments, two spatially separated SFM cylinders were presented to our observers simultaneously. One of these cylinders was completely ambiguous (lacking all explicit depth cues), while the other one was more or less disambiguated by the addition of either luminance or disparity depth cues. Under some specific conditions, depth information transferred from one cylinder to the other, resulting in a coupling of perceived rotation direction between the two cylinders. Mapping out the exact prerequisites for the spatial transfer of depth cues from one stimulus to the other, a functional organization is revealed that suggests lateral connectivity within the neural network that infers the structure of the rotating cylinders from the motion of the dot patterns. While lateral connectivity in visual cortex is a well-known anatomical principle, the specific results of these experiments



suggest that the extent of horizontal connectivity might depend on the depth tuning of the involved neurons. Neurons that selectively respond to visual motion located in depth behind the plain of fixation (further away from the observer) are implied to be more strongly horizontally connected than neurons that selectively respond to similar motion located in front of the plain of fixation (closer to the observer). Such an asymmetric connectivity structure makes sense from an ecological optics perspective. Partial occlusion, a very common visual phenomenon, causes visual objects in the background to be 'split-up' into multiple chunks of visual information on the retina. In dealing with occlusion situations, it might help the visual system to have the necessary hardware to 'bind' these chunks of information and re-compose a single object representation rather than represent the individual chunks of visual information separately.

#### 9.1.5 *Chapter 6: Crossmodal context*

Crossmodal integration is another contextual influence that may alter visual perception. It is addressed in Chapter 6 of this thesis. Perception is classically studied within a single modality. However, in the real world, information about events or objects is often present in more than one modality. This thesis-chapter describes how the perceived duration of an event is influenced by the simultaneous presence of both auditory and visual temporal information. Our results suggest that while auditory temporal information is automatically and involuntarily recruited for the perception of visual durations, the opposite is not true. Integrating theories of cue combination, perceptual grouping, temporal ventriloquism and scalar timing in a single framework of crossmodal event duration perception, we offer an explanation for our results that could serve as a general mechanism for crossmodal binding in the brain. In short, it proposes that intramodal grouping takes precedence over crossmodal grouping, that the occurrence of crossmodal grouping crucially depends on the spatiotemporal proximity of the stimulus features in the different modalities, and that upon crossmodal integration, the eventually perceived features are a weighted average of the information in the involved modalities with a higher weight on the modality that usually provides the most reliable information about this feature.

#### 9.1.6 *Chapter 7: Plasticity*

Many of the research chapters in this thesis arrive upon the conclusion that the demonstrated effects could be attributed to the interactive connectivity between neurons. In Chapter 7, we demonstrate that, at least for some of these connections, the gain (or efficacy) of this functional connectivity does not have a fixed value. Instead, the visual system constantly recalibrates its connectivity based on the current and recent sensory experience. In this chapter, the experiments are centered on a feature of binocular rivalry that is often either considered irrelevant, or ignored altogether. Binocular rivalry is generally characterized as a situation without binocular fusion with perception alternating between exclusive representations of the two monocular images. However, during a limited fraction of the time, binocular fusion does (partially) occur and observers report perceiving mixtures of the two monocular images. If the incidence of mixture percepts is carefully tracked over a prolonged period of binocular



rivalry viewing, it is revealed that the proportion of mixture percepts steadily increases over time. A generally accepted mechanistic view of the neural underpinnings of binocular rivalry involves two populations of neurons that are each selectively coding for one of the two competing images. Both the perceptual exclusivity and the perceptual alternations in binocular rivalry are thought to arise from a mixture of cross-inhibition between the two populations and the self-adaptation within the populations. The observed increase in the proportion of mixture percepts with prolonged rivalry viewing can be readily interpreted as a decrease in the strength of the cross-inhibitory connections.

Interestingly, additional experiments demonstrated that the strength of cross-inhibition and its resulting proportion of mixture percepts do not recover passively. Instead, it requires specific binocularly matching visual input that supposedly drives the same neurons that are also involved in the binocular rivalry. These specific input requirements suggest the presence of a synaptic learning mechanism that might be categorized as ‘anti-Hebbian’. The Hebbian synapse is a well-known neural principle that is often summarized as ‘neurons that fire together, wire together’. It means that if the activity of one neuron successfully contributes to the evoked activity of a second neuron, their synaptic bond strengthens, thereby increasing the correlation in their activity patterns. The ‘anti-Hebbian’ synapse that is proposed in our experiments employs a similar rule, but because the involved synapse is inhibitory, an increased synaptic bond leads to a decorrelation of activity patterns through stronger inhibition. Demonstrating this previously unknown form of experience-driven plasticity in the neuronal organization of adult binocular vision, our findings might help to understand how binocular fusion and binocular rivalry can arise from a single neuronal organization of binocular vision.

## 9.2 Reaching through a window on consciousness

If there is one thing that can be concluded from the studies described in this thesis, it is that ambiguous visual stimuli are a very appropriate tool to study the neural mechanisms of conscious vision. The impressive amount of existing work on visual rivalry is still rapidly growing, but it is time to reconsider the way in which visual ambiguities are studied. It is important to make a clear distinction between studying the visual rivalry phenomenon in itself and studying vision by using perceptual rivalry as a tool. While the former approach may result in an increasingly detailed characterization of some specific types of rivalry, the latter seems more promising in revealing basic operational mechanisms of brain functioning. It seems unlikely that the brain has built-in machinery that is specifically dedicated to the resolution of ambiguous visual stimuli, other than the same neural machinery that is also involved in normal vision. Many aspects of the perceptual dynamics of visual rivalry will therefore directly reflect the basic properties of the observer’s visual system.

Personal biases for a specific perceptual interpretation of an ambiguous stimulus, for instance, are very commonly encountered in visual rivalry work. Studies using TMS to stimulate the cortex during sequential stimulation of ambiguous stimuli have demonstrated that personal biases and the phenomenon of perceptual stabilization are independent (Brascamp et al., 2010). While the perceptual stabilization effect may depend on generic adaptation effects involved in all neuronal processing, the personal biases may result from inhomogeneities in the sensitivity of individual neurons or brain areas. Such a view is supported by a widespread



cortical signature of personal biases (Raemaekers et al., 2008) and a correlation between personal bias and actual brain structure (Kanai et al., 2010) that may be genetically determined (Miller et al., 2010). These findings are more informative about the inhomogeneous nature of the cortex than that they provide any intrinsically new insights in visual rivalry dynamics.

The spatial location where a transition in binocular rivalry from one perceptual interpretation to the other initiates, seems closely related to these cortical inhomogeneities. Perceptual alternations do not occur simultaneously for the full area of visual field occupied by the stimulus, but instead start at one location and spread throughout the stimulus resembling a traveling wave of dominance (Lee et al., 2005; Wilson et al., 2001). The location where this wave starts can be manipulated by locally increasing the contrast of the suppressed stimulus (Paffen et al., 2008), indicating that binocular rivalry could be used as a tool to map the cortical sensitivity throughout the visual field. Using the dynamics of these traveling waves in binocular rivalry it has also been hypothesized that the spatial organization of lateral connectivity in early visual cortex may be selectively tuned by higher-level feedback (Knapen et al., 2007b).

A whole series of experiments on the perceptual stabilization phenomenon, evoked by interrupted presentation of ambiguous stimuli, has also revealed a range of more basic neuronal aspects of visual processing. One important finding is that neuronal adaptation in the cortex occurs on multiple timescales in parallel (Brascamp et al., 2009; Brascamp et al., 2008). Adaptation alters the neuronal response sensitivity and affects the response magnitude to future stimulus presentation. When observers are pre-exposed to one of the monocular images that is later used in binocular rivalry, it depends on the duration of this pre-exposure and the strength of the stimuli how the pre-exposure affects initial dominance during rivalry (Brascamp et al., 2007). While short exposure to weak stimuli primes the visual system, causing the same monocular image to be initially dominant upon rivalry viewing (flash facilitation), longer exposure with a strong stimulus causes the opposite effect resulting in the initial dominance of the other image (flash suppression). Such a gradual transition from facilitative to inhibitory adaptation effects may have an important, yet undisclosed, function in neuronal processing. A functional correlation of sensitivity between neurons that represent adjacent regions of space has been demonstrated by the spatial transfer of perceptual stabilization that gets weaker with increasing distance between the probed spatial locations (Knapen et al., 2009). In accordance with the neurophysiological findings in Chapter 4 of this thesis, such spatial decay suggests an involvement of the local cortical network in the resolution and stabilization of local visual ambiguities.

The extent to which visual rivalry occurs depends on the amount of conflict that is embedded in the visual stimulus (Brouwer et al., 2009; Knapen et al., 2007a; van Ee et al., 2002). A natural continuum from fully unambiguous to completely balanced, ambiguous visual input is consistent with the idea that all these stimuli are being processed by basically the same neural mechanisms and implies that the functional properties of visual rivalry might be most informatively targeted by focusing on one of its most fundamental properties: the occurrence and resolution of a perceptual conflict.

Surprisingly, the occurrence of a perceptual conflict and the corresponding perceptual alternations do not always require the simultaneous presence of competing stimulus interpretations. Instead, rivalry is instigated whenever the stimuli are presented in close enough temporal proximity (less than approximately 350 ms apart)(van Boxtel et al., 2008). These



findings again illustrate that the perceptual dynamics during rivalry viewing is more likely a by-product of the normal visual processing mechanisms, than a separate autonomous rivalry mechanism. If perceptual alternations do occur, their temporal dynamics are highly comparable between different types of rivalry (Brascamp et al., 2005; van Ee, 2005) implying basic neuronal properties to be at the basis of this phenomenon.

The fact that perceptual processing of different stimulus types will involve both qualitatively and quantitatively different neuronal populations explains much of the demonstrated differences between binocular rivalry and perceptual rivalry. The susceptibility to attentional influences or voluntary control is one of the differences between binocular rivalry and perceptual rivalry that is often stressed as a crucial divergence (Klink et al., 2008a; Meng & Tong, 2004; van Ee et al., 2005). If observers are asked to voluntarily control their perception during rivalry, they may be able to do so up to a certain level, but they are usually better in controlling perception during perceptual rivalry than during binocular rivalry. If we would interpret voluntary control as the conscious attentional strategy of which it bears all the characteristics (Klink et al., 2008a), this dissociation may not be surprising. Attention is thought to influence neuronal processing through a straightforward gain control mechanism (Reynolds & Chelazzi, 2004). Since the effects of attention are stronger in higher visual cortical areas (Maunsell & Cook, 2002), the effects of voluntary control will naturally also be stronger for stimuli that are processed at a later moment in the hierarchical chain of cortical processing. Provided that binocular rivalry takes at least partially place between very low-level, eye specific neuronal representations, attention will only have a limited influence on perception. More general, the potential success of voluntary control will thus always depend on the stimulus characteristics involved in the rivalry (Brouwer & van Ee, 2006).

Expanding this line of reasoning, it makes sense that the efficacy of voluntary control over binocular rivalry perception can be significantly improved by adding congruent sensory information in another modality (van Ee et al., 2009). Even though there may be a slight cost of attentional resources involved in the crossmodal perceptual grouping, the benefits of having a higher-level multimodal percept, the gain of which may be more strongly scaled through voluntary control, will far outweigh these attentional investments. An experimental condition in which observers are asked to passively view a rival stimulus is often included in studies on the role of attention in visual rivalry. Whereas such condition is usually labeled as a ‘no attention’ condition, it must in fact still involve a limited amount of attention because observers are willfully watching a stimulus. Evidence for this hypothesis comes from studies that demonstrate that pulling attention away from the rivalry stimulus has an effect on perceptual dynamics that resembles a gain decrease on the activity of the involved visual neurons (Alais et al., 2010; Paffen et al., 2006). Apparently, attention is not something that is either present or not. Instead, it is always present in the form of a gain control mechanism, the influence of which is difficult to notice when it is ‘stuck in neutral’ (i.e. a gain factor of one)(Lee & Maunsell, 2009; Reynolds & Heeger, 2009).

Returning to rivalry, there is one specific topic that will have to be investigated better in the future. Current studies on perceptual rivalry have the tendency to look at periods of absolute dominance of one of the two perceptual interpretations, while the actual switch between the two percepts is often neglected. However, the switch appears to be the moment when the brain reorganizes its mind, or more mechanistically stated, the perceptual state embedded in the relevant neuronal populations gets altered. Studies on the transitions between exclusive



dominance percepts (Brascamp et al., 2006) and the occurrence of mixture percepts without absolute dominance (Blake et al., 1992; Hollins, 1980; Hollins & Hudnall, 1980; Klink et al., 2010; Kovács et al., 1996; O'Shea et al., 1997; Yang et al., 1992) are still relatively sparse, but they can potentially reveal fundamental principles of the neuronal organization involved in visual rivalry (as we show in Chapter 7 of this thesis).

Altogether, rivalry should no longer be considered as an interesting, yet exceptional, perceptual phenomenon, but rather be more extensively used as a tool to study the neuronal foundations of vision in general. Especially when this tool is placed in a broad range of carefully designed contexts might we be able to reach through the window on consciousness and start learning about the neural mechanisms that shape conscious visual perception by interacting with them directly.

### 9.3 Studying inference at the network level

Recent technical advances have given neuroscience an enormous boost into the 21st century. Within the wide range of techniques that are currently available to study brain functioning, each individual approach has its own strengths and limitations. While the choice of experimental technique will essentially always depend on the precise research questions and hypothesis at hand, neural processing at the level of the neural network has become a lot more accessible through recent technological advances. At the same time, these are precisely the levels where many of the effects demonstrated in studies with more classical techniques were hypothesized to emerge. The research in this thesis does not form an exception to this idea. For instance, we have hypothesized that spatial interactions are mediated by lateral connections; that perceptual stabilization could be based on a neuronal response stabilization driven by an increased involvement of the local cortical network; and that plasticity in binocular vision arises from perceptual (synaptic) learning within a specific network of neuronal populations.

Such network properties may now be studied in human and nonhuman primates with high-field fMRI, two-photon microscopy and large intracranial electrode arrays. Whereas initial fMRI scanners imaged the brain using a magnetic field strength of 1.5 Tesla which results in a spatial resolution of lower than  $2 \times 2 \times 2$  mm voxels, modern high-field scanners employ a 7 Tesla magnetic field. Such a strong magnetic field can result in voxels of approximately  $0.5 \times 0.5 \times 3$  mm, small enough to image individual ocular dominance and orientation columns (Yacoub et al., 2008; Yacoub et al., 2007). While this technique is still heavily being developed, it would be a large step forward if the dynamics of human brain activity at the level of cortical columns could be recorded with such non-invasive techniques.

In animal models, two-photon imaging provides a novel extension on conventional confocal optical imaging techniques, allowing crisp spatiotemporal recording of complete populations of neurons. Combined with, for instance, calcium concentration sensitive dyes introduced into a specific subset of neurons, the activity levels of tens or hundreds of neurons can be measured with a temporal resolution of milliseconds (Svoboda & Yasuda, 2006). Besides its high temporal resolution, this technique also provides a very high spatial resolution, allowing the imaging of precise morphological structures and changes therein, such as those occurring in dendritic spines when the brain learns through experience (Keck et al., 2008). Another



advantage of two-photon imaging over conventional confocal microscopy is its capacity to image a lot deeper into the cortical tissue, making it possible to reconstruct the three-dimensional functional and anatomical structure of the cortical networks.

In line with the suggestion that interfering with a process is a good approach to learn about its functional mechanisms, some neural stimulation techniques are expected to provide valuable insights about brain functioning in the near future. The non-invasive application of TMS has already been growing in popularity for several years, but some much more precise invasive stimulation techniques have also been introduced recently. Optogenetic stimulation offers a spectacular potential for the precise manipulation of neural functions *in vivo*. In short, optogenetics involves transgenic laboratory animals that have light-sensitive ion channels in the membrane of a subset of their neurons. By introducing a glass fiber into the brain and stimulating these ion channels with the proper light stimulus, an experimenter can essentially switch this class of neurons on and off, while simultaneously recording either the animal's behavior or the activity in another set of neurons (Miller, 2006; Zhang et al., 2010). While *in vivo* optogenetics, combined with the simultaneous recording of network activity at a high spatiotemporal resolution, sounds like the Holy Grail of interactive systems neuroscience, the technique is still highly experimental. However, the first results of optogenetics studies slowly start to appear in the literature and many laboratories around the world have embraced its potential and work hard on improving many different aspects of the technique. It seems to be merely a matter of time before optogenetics will be part of the neuroscientist's standard experimental toolbox.

## 9.4 Vision in the laboratory

The important role of contextual inference in the perception of ambiguous visual stimuli is a recurring theme throughout this thesis. However, when one studies vision in the laboratory, a typical stimulus will not be a realistic representation of the patterns that usually hit the retina during 'naturalistic' vision in the real world. Instead, an isolated aspect of this visual scene will be translated into a highly abstract visual stimulus, stripped from most natural context and presented against a diffuse, uniform background in a dark and quiet room. This deviation from the real world situation has the disadvantage that one could potentially miss crucial contextual determinants of basic mechanisms of visual perception, simply because this specific contextual information was removed from the experimental stimulus. This may be a particularly important concern for ambiguous visual stimuli, since these stimuli derive their defining characteristic, the perceptual ambiguity, from exactly such lack of disambiguating context. At the same time, the isolated features approach is probably the only realistic approach in studying basic visual processing.

Compared to typical laboratory stimuli, naturalistic stimuli contain an enormous number of additional features and factors that one has to take into account when trying to understand an experimentally obtained effect. Even in an ideal situation, when a researcher would have full experimental control over all these individual parameters, the number of possible sources and interactions that could potentially contribute to any effect in the data will far outreach the amount of time and resources available to study that specific effect. In contrast, there would be absolutely no harm in trying to evaluate the basic knowledge obtained with abstract



experimental stimuli in a more realistic setting. Such an approach would extend the scope of experiments from the laboratory setting to a more realistic situation without introducing the limitation of the enormous amount of additional variables.

Along these lines, the field of natural image statistics may reveal fundamental information about the parametric range in which certain visual features occur in our natural visual surroundings (Geisler, 2008). Visual stimuli that are generated in the laboratory may cover enormous ranges of luminance, contrast, spatial frequency, speed, color, etc. However, based on principles of evolution, we would hypothesize that neurons will have evolved to be most sensitive to the feature-ranges that they encounter the most. Testing neural mechanisms with stimuli whose features fall well outside these ranges may yield interpretable results, but it remains questionable whether these results will be representative for the way in which the brain works in a more naturalistic environment. While often overlooked, it seems relevant for most vision science experiments to include an evaluation of the relation between the laboratory situation and the visual system's natural habitat.

## 9.5 Learning the language of neurons

In many theoretical explanations of visual processing mechanisms, including those on the contextual resolution of visual ambiguities, information transfer is assumed to take place between individual neurons or neuronal populations. When one studies the neural underpinnings of some particular brain function experimentally, it is non-trivial to know how information is transferred between neurons. 'Brain activity' is an umbrella term that covers many different potential neuronal languages. Single unit neurophysiology works in terms of instantaneous spike rates, precise spike times or local field potentials that are thought to reflect the cumulative synaptic potential of a large number of local neurons; fMRI measures blood flow or blood oxygenation levels; two-photon imaging with calcium-sensitive dyes records intracellular calcium concentration. Focusing on perception alone, there is evidence for information transfer in spike counts, spike times, interspike intervals, spiking coherence, low frequency oscillation. It is likely that the brain uses all these specific coding strategies simultaneously. Furthermore, comparable to human language, the distribution of information in neuronal 'language' is not homogeneous. Some cells may be more informative in providing information about a particular event (Parker & Newsome, 1998) and even within a single cell some spikes are more informative than others (Masse & Cook, 2008)(See the Appendix Chapter 8 in this thesis for a more extensive discussion). The deciphering of neuronal language is a massive challenge. Luckily there are numerous skilled translators working together in an attempt to break the coding scheme and translate the ever-increasing record of neural communication.

## 9.6 In conclusion

The experiments in this thesis describe how the visual system seemingly effortlessly uses context to resolve local ambiguities and create coherent and stable conscious visual percepts. The neural mechanisms that facilitate such contextual inference are hypothesized to involve information transfer in relatively local cortical networks. The connectivity within these networks



is suggested to be self-organizing and reasonably dynamic, allowing the brain to adapt its functional processing networks to novel experiences. Focusing future research efforts on the neural dynamics within local cortical networks and the nature of interneuronal information transfer in visual processing will likely result in a much more detailed understanding of the complex operational principles that enable us to, quite literally, see things in context.

---

# Samenvatting

**Summary in Dutch**





Het onderzoek dat is beschreven in dit proefschrift houdt zich bezig met de brede vraag naar de rol van contextuele informatie in het tot stand komen van onze bewuste visuele waarnemingen. Met contextuele informatie wordt alle informatie bedoeld die niet direct in een visueel patroon of object zelf aanwezig is, maar die wel informatie kan bevatten over de perceptuele interpretatie ervan. Omdat ieder visueel neuron slechts reageert op de aanwezigheid van stimuli in een relatief klein stukje van de visuele ruimte, moet het brein informatie integreren over zowel de tijd als de ruimte. Daarnaast kan ook de aanwezigheid van informatie in een andere modaliteit dan de visuele aanknopingspunten bieden voor de interpretatie van visuele patronen. Het begrijpen van de neuronale mechanismen van context integratie op het niveau van complete visuele patronen en objecten kan helpen in het leren begrijpen van meer algemene neuronale integratiemechanismen die op grote schaal in het brein voorkomen. Hieronder volgen samenvattingen van de afzonderlijke onderzoekshoofdstukken die in dit proefschrift zijn opgenomen.

### *Hoofdstuk 2: De algemene geldigheid van Levelt's stellingen onthult universele computationale mechanismen in visuele rivaliteit*

De visuele perceptie literatuur staat bol van vergelijkingen tussen binoculaire rivaliteit en perceptuele rivaliteit. Bij binoculaire rivaliteit ontstaat een perceptueel conflict doordat radicaal verschillende beelden worden getoond aan de afzonderlijke ogen. Het brein voegt deze beelden niet samen. In plaats daarvan wisselt onze perceptuele interpretatie automatisch iedere paar seconden tussen de twee beelden. Bij perceptuele rivaliteit gebeurt iets soortgelyks, maar nu doordat een plaatje dat wordt aangeboden aan beide ogen meerdere, even waarschijnlijke, interpretaties kent. Een bekend voorbeeld hiervan is de Necker kubus waarbij een tweedimensionale afbeelding van een kubus kan worden waargenomen in twee driedimensionale oriëntaties. In dit hoofdstuk gebruiken we een tweedimensionale projectie van een transparante roterende bol met stippen op de oppervlakte. Omdat deze stimulus geen informatie bevat over de diepte-ordening van de stippen is de draairichting ambigue. We laten zien dat de neurale mechanismen die er voor zorgen dat slechts een enkele perceptuele interpretatie tegelijk kan worden waargenomen, vanuit een computationeel perspectief, universeel zijn en zowel voor perceptuele rivaliteit als binoculaire rivaliteit gelden. We doen dit door een set stellingen te generaliseren, die in de jaren zestig door Pim Levelt zijn geformuleerd in de specifieke context van binoculaire rivaliteit. De originele stellingen beschrijven de relatie tussen het contrast van de monoculaire beelden in binoculaire rivaliteit en de perceptuele dynamica die hier het gevolg van zijn. Door een herdefiniëring in termen van stimulus bias kunnen deze stellingen ook eenvoudig worden toegepast op perceptuele rivaliteit. De voorspellingen die dat oplevert worden in een serie experimenten allemaal bevestigd.

### *Hoofdstuk 3: Vroege interacties tussen neuronale adaptatie en vrijwillige sturing bepalen perceptuele keuzes tijdens bistabiele visuele perceptie*

In dit hoofdstuk wordt gebruik gemaakt van sequenties korte vertoningen van visuele stimuli die ofwel binoculair, ofwel perceptueel rivaliseren. Voor zowel binoculaire rivaliteit als per-



ceptuele rivaliteit zijn de presentatielijnen en de duur van de tussenliggende perioden zonder stimuli systematisch gemanipuleerd. De resultaten laten zien hoe de keuze van het brein voor een van beide percepten afhangt van de temporele eigenschappen van de vertoning. De duur van de stimulusonderbrekingen bleek cruciaal te zijn. Bij lange onderbrekingen ( $> 1$  s) zagen de proefpersonen vaak dezelfde perceptuele interpretatie tijdens opeenvolgende vertoningen, terwijl ze bij korte onderbrekingen ( $< 0.5$  s) juist veel vaker wisselende interpretaties waarnamen. In de effecten van presentatielijnduur, die overigens veel zwakker waren, vonden we bovendien aanwijzingen voor neuronale adaptatieverschijnselen die plaats moeten vinden voordat de rivaliteit in het brein wordt opgelost. Deze verschijnselen waren voor binoculaire rivaliteit veel zwakker dan voor perceptuele rivaliteit wat suggereert dat binoculaire rivaliteit op een veel vroeger niveau van visuele informatieverwerking plaatsvindt dan perceptuele rivaliteit.

In een tweede set experimenten onderzoeken we de onderliggende mechanismen van de vrijwillige sturing die proefpersonen kunnen uitoefenen tijdens het kijken naar ambigue visuele stimuli. Proefpersonen kunnen deels zelf bepalen welke van de twee perceptuele interpretaties dominant wordt bij iedere stimulusvertoning. Onze experimenten laten zien dat de effecten van deze vrijwillige sturing vergelijkbaar zijn met de effecten van verhoogde attentie voor één van de twee rivaliserende stimuli. Computationeel kunnen deze effecten geïnterpreteerd worden als een versterkingsstap die een bias voor één van de twee perceptuele interpretaties kan realiseren in een vroeg stadium van visuele informatieverwerking voordat de rivaliteit wordt opgelost en het brein “een percept kiest”.

#### *Hoofdstuk 4: Onderbroken stimulus presentatie stabiliseert neuronale responsies in hersengebied MT van makaken*

De effecten van onderbrekingsduur tijdens de onderbroken vertoning van ambigue visuele stimuli zijn verder onderzocht op het niveau van activiteitspatronen van individuele neuronen. Verschillende stimuli werden aangeboden in sequenties van ongeveer tachtig herhalingen waarbij de onderbrekingsduur systematisch gevarieerd werd tussen 250 en 2000 ms. Een eerste gevolg van de herhaalde presentatie is dat een gestimuleerd neuron adapteert aan de stimulus en minder sterk reageert op de aanwezigheid van deze stimulus. Als er tussen de stimuluspresentaties meer tijd is om te herstellen van deze adaptatie, is deze reductie in responsie-amplitude minder groot. De meer interessante vindingen liggen echter op het vlak van variabiliteit in de activiteit. Als een stimulus wordt aangezet gaat de variabiliteit naar beneden, maar de mate waarin dat gebeurt hangt af van de onderbrekingsduur. Lange onderbrekingen zorgen voor een lagere variabiliteit dan korte onderbrekingsduren. Dit geldt zowel voor de variabiliteit over de hele sequentie van stimuluspresentaties als voor de fluctuaties in responsies op opeenvolgende stimulus presentaties. Ook de temporele structuur van de serie actiepotentialen die door het neuron wordt gegenereerd in reactie op een stimulus is stabieler als de onderbrekingen langer zijn. Deze stabilisatie-effecten in de activiteit van enkele cellen gingen gepaard met een verhoging van de power in de hoge Gamma frequentieband van het locale veldpotentiaal. Dergelijk powerverhogingen worden doorgaans geïnterpreteerd als een versterking van de coherentie in de activiteit van het lokale corticale netwerk. Het lijkt er derhalve op dat het lokale corticale netwerk de activiteitspatronen van enkele cellen actief stabiliseert om de signaal-ruis verhouding te verhogen. Dit proces heeft waarschijnlijk tijd



nodig, waardoor langere onderbrekingsduren tot een grotere mate van stabiliteit leiden. Een dergelijk algemeen neuronaal principe zou, in het geval van ambigue stimuli, een belangrijke rol kunnen spelen in het stabiliseren van perceptie over herhaalde stimulus presentaties.

#### *Hoofdstuk 5. Aan occlusie gerelateerde laterale verbindingen stabiliseren kinetische diepte stimuli door middel van perceptuele koppeling*

Ook spatiële context kan de perceptie van ambigue stimuli stabiliseren. In dit hoofdstuk werken we met zogenaamde 'kinetische diepte stimuli' waarbij de driedimensionale structuur van de stimulus wordt gegenereerd door bewegingsinformatie. In dit geval was er wederom sprake van een roterende cilinder met een ambigue draairichting. In een eerste stap hebben we uitgezocht hoe dergelijke stimuli gedisambigueerd kunnen worden door het parametrisch toevoegen van diepte informatie via binoculaire dispariteit of een luminantie gradiënt. In de hierop volgende experimenten werden steeds twee cilinders tegelijk op het scherm getoond, waarvan er eentje volledig ambigue was en de ander een bepaalde mate van diepte informatie kon bevatten. Onder specifieke omstandigheden werd deze diepte informatie overgedragen van de ene naar de andere cilinderstimulus. De totale set experimenten suggereert dat deze informatie overdracht plaatsvindt via de horizontale verbindingen in de visuele cortex tussen groepen cellen die specifiek gevoelig zijn voor dezelfde eigenschappen van een stimulus die zich achter het fixatiepunt bevindt. Een dergelijk mechanisme van informatie-groepering in de achtergrond van een visuele scène kan het visueel systeem goed gebruiken bij het interpreteren van voorwerpen die zich achter andere voorwerpen bevinden en daar deels door worden afgedekt.

#### *Hoofdstuk 6. Crossmodale perceptie van tijdsduur omvat perceptuele groepering, temporeel ventriloquisme en variabele snelheden van een interne klok*

Tijd is een relatief begrip en dit geldt al helemaal wanneer sensorische informatie over de duur van een gebeurtenis aanwezig is in verschillende modaliteiten. In dit hoofdstuk wordt de perceptie van de duur van gebeurtenissen onderzocht in een situatie waar temporele informatie aanwezig is in zowel de visuele als de auditieve modaliteit. Onze resultaten laten zien dat het brein auditieve temporele informatie automatisch betrekt bij het waarnemen van de duur van visuele gebeurtenissen, maar dat andersom visuele informatie niet wordt gebruikt voor het waarnemen van auditieve duren. Om de vindingen van een serie experimenten te verklaringen combineren we theorieën van perceptuele groepering, cue combinatie, temporeel ventriloquisme en een intern klok-mechanisme in een algemeen raamwerk van crossmodale tijdspercepcie. Hierin heeft de intramodale groepering van stimuli voorrang op crossmodale groepering. Daarnaast hangt het crossmodale groeperen cruciaal af van het verschil in spatiotemporale stimuluseigenschappen tussen de modaliteiten. Als crossmodale groepering optreedt zal de uiteindelijk waargenomen duur een gewogen gemiddelde zijn van de duur-informatie in de verschillende modaliteiten. Hierin zal de informatie in de modaliteit die doorgaans de meest betrouwbare informatie geeft zwaarder wegen. Voor temporele informatie is dat de auditieve modaliteit.



## *Hoofdstuk 7. Door ervaring veroorzaakte plasticiteit in binoculair zien*

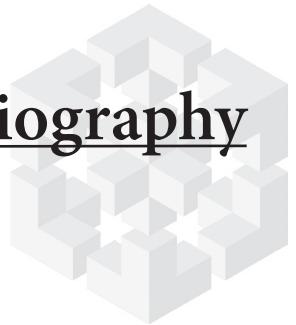
In veel van de hoofdstukken in dit proefschrift wordt de verklaring voor de gevonden effecten gezocht in de verbindingen tussen neuronen. In dit hoofdstuk laten we zien dat voor sommige van deze verbindingen de sterkte niet constant is, maar afhangt van de recente activiteitspatronen, waardoor het visueel systeem zichzelf continue kan recalibreren aan de recente sensorische geschiedenis. De experimenten in dit hoofdstuk zijn gebaseerd op een specifieke eigenschap van binoculaire rivaliteit die vaak als irrelevant wordt beschouwd of zelfs eenvoudigweg genegeerd. Binoculaire rivaliteit wordt vaak gekarakteriseerd als een situatie zonder binoculaire fusie, waarin perceptie fluctueert tussen twee exclusieve representaties van de monoculaire beelden. Er zijn echter ook perioden waarbij binoculaire fusie wel (deels) optreedt en proefpersonen een mengsel waarnemen van de twee monoculaire beelden. Als het voorkomen van deze meng-percepten over een langere periode van binoculaire rivaliteit wordt geregistreerd zien we dat dergelijke meng-percepten in de loop van de tijd steeds vaker voorkomen. Een algemeen geaccepteerde mechanistische interpretatie van het binoculaire rivaliteitsmechanisme gaat uit van twee populaties neuronen die elk één van de twee monoculaire beelden representeren. Zowel de perceptuele exclusiviteit als de perceptuele fluctuaties worden over het algemeen toegeschreven aan een mix van kruis-inhibitie tussen deze twee neuronale populaties en adaptatie binnen de populaties. Binnen deze gedachte zou de geobserveerde verhoging in het voorkomen van meng-percepten het gevolg kunnen zijn van een geleidelijke verzwakking van de kruis-inhibitie.

Een extra set experimenten laat zien dat de verzwakte inhibitie en het verhoogde percentage meng-percepten dat daaruit voorkomt niet passief herstellen, maar een hele specifieke binoculair overeenkomende, visuele input vereisen die dezelfde neuronen activeren als de binoculaire rivaliteitsstimulus. Het feit dat zulke specifieke activiteitspatronen vereist zijn om veranderingen van inhibitie-sterkte tot stand te brengen, suggereert de aanwezigheid van een zogenaamd 'anti-Hebbisch' synaptisch plasticiteitsmechanisme. De Hebbische synaps is een bekend neuronaal principe dat vaak wordt samengevat als 'neuronen die samen vuren, raken verbonden'. Dit betekent dat als de activiteit van een neuron bijdraagt aan het succesvol activeren van een tweede neuron, de connectiviteit tussen deze neuronen versterkt wordt en de correlatie tussen de activiteitspatronen van de twee neuronen sterker wordt. De anti-Hebbische synaps die we voorstellen in dit hoofdstuk is gebaseerd op een vergelijkbaar principe, maar omdat het hier niet gaat om excitatoire maar om inhibitoire synapsen leidt een versterkte connectiviteit juist tot een decorrelatie van activiteitspatronen. Met het aantonen van deze nieuwe vorm van door ervaring gedreven neuronale plasticiteit in het volwassen brein kan een belangrijke stap gezet worden in het leren begrijpen hoe binoculaire fusie en binoculaire rivaliteit voort kunnen komen uit een zelfde neuronale organisatie van binoculair zien.



---

## Bibliography





- Adams, D.L., & Zeki, S. (2001). Functional organization of macaque V3 for stereoscopic depth. *J Neurophysiol*, 86 (5), 2195-2203.
- Agüera y Arcas, B., Fairhall, A.L., & Bialek, W. (2003). Computation in a single neuron: Hodgkin and Huxley revisited. *Neural Comput*, 15 (8), 1715-1749.
- Ahmed, B., Hanazawa, A., Undeman, C., Eriksson, D., Valentiniene, S., & Roland, P.E. (2008). Cortical dynamics subserving visual apparent motion. *Cereb Cortex*, 18 (12), 2796-2810.
- Alais, D., & Blake, R. (1999). Neural strength of visual attention gauged by motion adaptation. *Nat Neurosci*, 2 (11), 1015-1018.
- Alais, D., & Blake, R. (2005). *Binocular Rivalry*. Cambridge, MA; The MIT Press.
- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol*, 14 (3), 257-262.
- Alais, D., & Melcher, D. (2007). Strength and coherence of binocular rivalry depends on shared stimulus complexity. *Vision Res*, 47 (2), 269-279.
- Alais, D., van Boxtel, J.J., Parker, A., & van Ee, R. (2010). Attending to auditory signals slows visual alternations in binocular rivalry. *Vision Res*, 50 (10), 929-935.
- Andersen, R.A., & Bradley, D.C. (1998). Perception of three-dimensional structure form motion. *Trends Cogn Sci*, 2 (6), 222-228.
- Anderson, B.L., Singh, M., & Fleming, R.W. (2002). The interpolation of object and surface structure. *Cogn Psychol*, 44 (2), 148-190.
- Andrews, T.J., & Purves, D. (1997). Similarities in normal and binocularly rivalrous viewing. *Proc Natl Acad Sci USA*, 94 (18), 9905-9908.
- Angrilli, A., Cherubini, P., Pavese, A., & Mantredini, S. (1997). The influence of affective factors on time perception. *Percept Psychophys*, 59 (6), 972-982.
- Assad, J.A., & Maunsell, J.H. (1995). Neuronal correlates of inferred motion in primate posterior parietal cortex. *Nature*, 373 (6514), 518-521.
- Attneave, F. (1968). Triangles as ambiguous figures. *Am J Psychol*, 81 (3), 447-453.
- Bair, W., Koch, C., Newsome, W., & Britten, K. (1994). Power spectrum analysis of bursting cells in area MT in the behaving monkey. *J Neurosci*, 14 (5 Pt 1), 2870-2892.
- Baker, D.H. (2010). Visual consciousness: the binocular rivalry explosion. *Curr Biol*, 20 (15), R644-646.
- Bakin, J., Nakayama, K., & Gilbert, C.D. (2000). Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *J Neurosci*, 20 (21), 8188-8198.
- Barlow, H.B. (1990). A theory about the functional role and synaptic mechanism of visual



- after-effects. In: C.B. Blakemore (Ed.) *Vision: coding and efficiency* (pp. 363-375). New York: Cambridge University Press.
- Barlow, H.B., & Földiák, P. (1989). Adaptation and Decorrelation in the Cortex. In: C. Miall, R.M. Durbin, & G.J. Mitchison (Eds.), *The Computing Neuron* (pp. 54-72). Wokingham, England: Addison-Wesley.
- Barry, S.R. (2009). *Fixing my gaze: A scientist's journey into seeing in three dimensions*. (p. 272). New York, New York, USA; Basic Books.
- Belitski, A., Gretton, A., Magri, C., Murayama, Y., Montemurro, M.A., Logothetis, N.K., & Panzeri, S. (2008). Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J Neurosci*, 28 (22), 5696-5709.
- Bell, C.C., Han, V.Z., Sugawara, Y., & Grant, K. (1997). Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature*, 387 (6630), 278-281.
- Bertelson, P., & Aschersleben, G. (2003). Temporal ventriloquism: crossmodal interaction on the time dimension. 1. Evidence from auditory-visual temporal order judgment. *Int Journal Psychophysiol*, 50 (1-2), 147-155.
- Blake, R. (2001). A Primer on Binocular Rivalry, Including Current Controversies. *Brain Mind*, 2, 5-38.
- Blake, R., & Logothetis, N.K. (2002). Visual competition. *Nat Rev Neurosci*, 3 (1), 13-21.
- Blake, R., O'Shea, R.P., & Mueller, T.J. (1992). Spatial zones of binocular rivalry in central and peripheral vision. *Vis Neurosci*, 8 (5), 469-478.
- Blake, R., Sobel, K., & Gilroy, L.A. (2003). Visual motion retards alternations between conflicting perceptual interpretations. *Neuron*, 39 (5), 869-878.
- Blake, R., Yang, Y.D., & Wilson, H.R. (1991). On the coexistence of stereopsis and binocular rivalry. *Vision Res*, 31 (7-8), 1191-1203.
- Blaser, E., Sperling, G., & Lu, Z.L. (1999). Measuring the amplification of attention. *Proc Natl Acad Sci USA*, 96 (20), 11681-11686.
- Boltz, M.G. (2005). Duration judgments of naturalistic events in the auditory and visual modalities. *Percept Psychophys*, 67 (8), 1362-1375.
- Born, R.T., & Bradley, D. (2005). Structure and function of visual area MT. *Annu Rev Neurosci*, 28, 157-189.
- Bossink, C.J., Stalmeier, P.F., & De Weert, C.M. (1993). A test of Levelt's second proposition for binocular rivalry. *Vision Res*, 33 (10), 1413-1419.
- Boynton, G.M. (2005). Attention and visual perception. *Curr Opin Neurobiol*, 15 (4), 465-469.



- Bradley, D., Chang, G., & Andersen, R.A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature*, 392 (6677), 714-717.
- Bradley, D., Qian, N., & Andersen, R.A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature*, 373 (6515), 609-611.
- Bradley, D.C., & Andersen, R.A. (1998). Center-surround antagonism based on disparity in primate area MT. *J Neurosci*, 18 (18), 7552-7565.
- Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10 (4), 433-436.
- Brascamp, J., Pearson, J., Blake, R., & van den Berg, A. (2009). Intermittent ambiguous stimuli: Implicit memory causes periodic perceptual alternations. *J Vis*, 9 (3), 3:1-23.
- Brascamp, J.W., Kanai, R., Walsh, V., & van Ee, R. (2010). Human middle temporal cortex, perceptual bias, and perceptual memory for ambiguous three-dimensional motion. *J Neurosci*, 30 (2), 760-766.
- Brascamp, J.W., Knapen, T.H.J., Kanai, R., Noest, A.J., van Ee, R., & van den Berg, A.V. (2008). Multi-timescale perceptual history resolves visual ambiguity. *PLoS ONE*, 3 (1), e1497.
- Brascamp, J.W., Knapen, T.H.J., Kanai, R., van Ee, R., & van den Berg, A.V. (2007). Flash suppression and flash facilitation in binocular rivalry. *J Vis*, 7 (12), 12:11-12.
- Brascamp, J.W., van Ee, R., Noest, A.J., Jacobs, R.H., & van den Berg, A.V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *J Vis*, 6 (11), 1244-1256.
- Brascamp, J.W., van Ee, R., Pestman, W.R., & van den Berg, A.V. (2005). Distributions of alternation rates in various forms of bistable perception. *J Vis*, 5 (4), 287-298.
- Brouwer, G., & van Ee, R. (2007). Visual cortex allows prediction of perceptual states during ambiguous structure-from-motion. *J Neurosci*, 27 (5), 1015-1023.
- Brouwer, G.J., Tong, F., Hagoort, P., & van Ee, R. (2009). Perceptual incongruence influences bistability and cortical activation. *PLoS ONE*, 4 (3), e5056.
- Brouwer, G.J., & van Ee, R. (2006). Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics. *Vision Res*, 46 (20), 3393-3402.
- Bruns, P., & Getzmann, S. (2008). Audiovisual influences on the perception of visual apparent motion: exploring the effect of a single sound. *Acta Psychol*, 129 (2), 273-283.
- Buhusi, C.V., & Meck, W.H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci*, 6 (10), 755-765.
- Buracas, G.T., & Albright, T.D. (1996). Contribution of area MT to perception of three-dimension shape: a computational study. *Vision Res*, 36 (6), 869-887.
- Burr, D., Banks, M., & Morrone, M. (2009). Auditory dominance over vision in the perception



- of interval duration. *Exp Brain Res*, 198 (1), 49-57.
- Burr, D., & Morrone, C. (2006). Time perception: space-time in the brain. *Curr Biol*, 16 (5), R171-173.
- Buschman, T., & Miller, E. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315 (5820), 1860-1862.
- Butts, D.A., Weng, C., Jin, J., Yeh, C.-I., Lesica, N.A., Alonso, J.-M., & Stanley, G.B. (2007). Temporal precision in the neural code and the timescales of natural vision. *Nature*, 449 (7158), 92-95.
- Caporale, N., & Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. *Annu Rev Neurosci*, 31, 25-46.
- Carandini, M., Barlow, H.B., O'Keefe, L.P., Poirson, A.B., & Movshon, J.A. (1997). Adaptation to contingencies in macaque primary visual cortex. *Philos Trans R Soc Lond B Biol Sci*, 352 (1358), 1149-1154.
- Carter, O.L., & Cavanagh, P. (2007). Onset rivalry: brief presentation isolates an early independent phase of perceptual competition. *PLoS ONE*, 2, e343.
- Carter, O.L., & Pettigrew, J.D. (2003). A common oscillator for perceptual rivalries? *Perception*, 32 (3), 295-305.
- Chen, K., & Yeh, S. (2009). Asymmetric cross-modal effects in time perception. *Acta Psychol*, 130 (3), 225-234.
- Chen, X., & He, S. (2004). Local factors determine the stabilization of monocular ambiguous and binocular rivalry stimuli. *Curr Biol*, 14 (11), 1013-1017.
- Chong, S., & Blake, R. (2006). Exogenous attention and endogenous attention influence initial dominance in binocular rivalry. *Vision Res*, 46 (11), 1794-1803.
- Chong, S., Tadin, D., & Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *J Vis*, 5 (11), 1004-1012.
- Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., Bradley, D.C., Smith, M.A., Kohn, A., Movshon, J.A., Armstrong, K.M., Moore, T., Chang, S.W., Snyder, L.H., Lisberger, S.G., Priebe, N.J., Finn, I.M., Ferster, D., Ryu, S.I., Santhanam, G., Sahani, M., & Shenoy, K.V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci*, 13 (3), 369-378.
- Crick, F., & Koch, C. (2003). A framework for consciousness. *Nat Neurosci*, 6 (2), 119-126.
- DeAngelis, G., Cumming, B., & Newsome, W.T. (1998). Cortical area MT and the perception of stereoscopic depth. *Nature*, 394 (6694), 677-680.



- DeAngelis, G., & Uka, T. (2003). Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol*, 89 (2), 1094-1111.
- Deco, G., & Obradovic, D. (1995). Decorrelated Hebbian Learning for Clustering and Function Approximation. *Neural Comput*, 7, 338-348.
- Ding, J., & Sperling, G. (2006). A gain-control theory of binocular combination. *Proc Natl Acad Sci USA*, 103 (4), 1141-1146.
- Dodd, J., Krug, K., Cumming, B., & Parker, A. (2001). Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J Neurosci*, 21 (13), 4809-4821.
- Donovan, C.-L., Lindsay, D.S., & Kingstone, A. (2004). Flexible and abstract resolutions to crossmodal conflicts. *Brain Cognition*, 56 (1), 1-4.
- Dosher, B.A., Sperling, G., & Wurst, S.A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Res*, 26 (6), 973-990.
- Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition*, 80 (1-2), 61-95.
- Driver, J., & Spence, C. (1998). Cross-modal links in spatial attention. *Philos Trans R Soc Lond B Biol Sci*, 353 (1373), 1319-1331.
- Driver, J., & Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Curr Biol*, 10 (20), R731-735.
- Duncan, R., Albright, T.D., & Stoner, G.R. (2000). Occlusion and the interpretation of visual motion: perceptual and neuronal effects of context. *J Neurosci*, 20 (15), 5885-5897.
- Eagleman, D.M. (2008). Human time perception and its illusions. *Curr Opin Neurobiol*, 18 (2), 131-136.
- Eby, D.W., Loomis, J.M., & Solomon, E.M. (1989). Perceptual linkage of multiple objects rotating in depth. *Perception*, 18 (4), 427-444.
- Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Res*, 43 (9), 1035-1045.
- Ernst, M.O., & Bülthoff, H.H. (2004). Merging the senses into a robust percept. *Trends Cogn Sci*, 8 (4), 162-169.
- Escher, M.C. (1992). *The Graphic Work of M.C. Escher*. Cologne, Germany; Benedikt Taschen Verlag.
- Fagot, J., Barbet, I., Parron, C., & Deruelle, C. (2006). Amodal completion by baboons (*Papio papio*): contribution of background depth cues. *Primates*, 47 (2), 145-150.



- Falconbridge, M., & Badcock, D.R. (2006). Implicit exploitation of regularities: novel correlations in images quickly alter visual perception. *Vision Res*, 46 (8-9), 1331-1335.
- Fang, F., & He, S. (2004). Stabilized structure from motion without disparity induces disparity adaptation. *Curr Biol*, 14 (3), 247-251.
- Felisberti, F., & Zanker, J. (2005). Attention modulates perception of transparent motion. *Vision Res*, 45 (19), 2587-2599.
- Földiák, P. (1990). Forming sparse representations by local anti-Hebbian learning. *Biol Cybern*, 64 (2), 165-170.
- Fontanini, A., & Katz, D.B. (2008). Behavioral states, network states, and sensory response variability. *J Neurophysiol*, 100 (3), 1160-1168.
- Forkman, B. (1998). Hens use occlusion to judge depth in a two-dimensional picture. *Perception*, 27 (7), 861-867.
- Freeman, E., & Driver, J. (2008). Direction of visual apparent motion driven solely by timing of a static sound. *Curr Biol*, 18 (16), 1262-1266.
- Freeman, E.D., & Driver, J. (2006). Subjective appearance of ambiguous structure-from-motion can be driven by objective switches of a separate less ambiguous context. *Vision Res*, 46 (23), 4007-4023.
- Fries, P., Nikolić, D., & Singer, W. (2007). The gamma cycle. *Trends Neurosci*, 30 (7), 309-316.
- Gawne, T.J. (2010). The local and non-local components of the local field potential in awake primate visual cortex. *J Comp Neurosci*, [Epub ahead of print].
- Geisler, W.S. (2008). Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol*, 59, 167-192.
- Georgeson, M., Yates, T., & Schofield, A. (2008). Depth propagation and surface construction in 3-D vision. *Vision Res*, 49 (1), 84-95.
- Getzmann, S. (2007). The effect of brief auditory stimuli on visual apparent motion. *Perception*, 36 (7), 1089-1103.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol Rev*, 84 (3), 279-325.
- Gibbon, J., Church, R.M., & Meck, W.H. (1984). Scalar timing in memory. *Ann N Y Acad Sci*, 423, 52-77.
- Gibson, J.J. (1950). *The Perception of the Visual World*. Boston, MA; Houghton Mifflin.
- Gigante, G., Mattia, M., Braun, J., & Del Giudice, P. (2009). Bistable perception modeled as competing stochastic integrations at two levels. *PLoS Comput Biol*, 5 (7), e1000430.



- Gilbert, C.D., & Wiesel, T.N. (1979). Morphology and intracortical projections of functionally characterised neurones in the cat visual cortex. *Nature*, 280 (5718), 120-125.
- Gilbert, C.D., & Wiesel, T.N. (1983). Clustered intrinsic connections in cat visual cortex. *J Neurosci*, 3 (5), 1116-1133.
- Gilbert, C.D., & Wiesel, T.N. (1989). Columnar specificity of intrinsic horizontal and cortico-cortical connections in cat visual cortex. *J Neurosci*, 9 (7), 2432-2442.
- Gilroy, L.A., & Blake, R. (2004). Physics embedded in visual perception of three-dimensional shape from motion. *Nat Neurosci*, 7 (9), 921-922.
- Glickstein, M., & Whitteridge, D. (1987). Tatsuji Inouye and the mapping of the visual fields on the human cerebral cortex. *Trends Neurosci*, 10 (9), 350-353.
- Gold, J.I., & Shadlen, M.N. (2007). The neural basis of decision making. *Annu Rev Neurosci*, 30, 535-574.
- Gonzalez, F., Perez, R., Justo, M.S., & Bermudez, M.A. (2001). Receptive field organization of disparity-sensitive cells in Macaque medial superior temporal cortex. *Eur J Neurosci*, 14 (1), 167-173.
- Goodale, M., & Milner, A. (1992). Separate visual pathways for perception and action. *Trends Neurosci*, 15 (1), 20-25.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci*, 10 (1), 14-23.
- Grondin, S. (2003). Sensory modalities and temporal processing. In: H. Helfrich (Ed.) *Time and Mind II: Information Processing Perspectives* (pp. 61-77). Germany: Hogrefe & Huber.
- Grondin, S., & McAuley, D. (2009). Duration discrimination in crossmodal sequences. *Perception*, 38 (10), 1542-1559.
- Grossberg, S., Yazdanbakhsh, A., Cao, Y., & Swaminathan, G. (2008). How does binocular rivalry emerge from cortical mechanisms of 3-D vision? *Vision Res*, 48 (21), 2232-2250.
- Grossmann, J., & Dobbins, A. (2003). Differential ambiguity reduces grouping of metastable objects. *Vision Res*, 43 (4), 359-369.
- Grossmann, J., & Dobbins, A. (2006). Competition in bistable vision is attribute-specific. *Vision Res*, 46 (3), 285-292.
- Guttman, S.E., Gilroy, L.A., & Blake, R. (2005). Hearing what the eyes see: auditory encoding of visual temporal sequences. *Psychol Sci*, 16 (3), 228-235.
- Hancock, S., & Andrews, T.J. (2007). The role of voluntary and involuntary attention in selecting perceptual dominance during binocular rivalry. *Perception*, 36 (2), 288-298.



- Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat Rev Neurosci*, 7 (7), 523-534.
- Hegdé, J., Fang, F., Murray, S.O., & Kersten, D. (2008). Preferential responses to occluded objects in the human visual cortex. *J Vis*, 8 (4), 16:11-16.
- Hinkle, D.A., & Connor, C.E. (2005). Quantitative characterization of disparity tuning in ventral pathway area V4. *J Neurophysiol*, 94 (4), 2726-2737.
- Hoffman, D.D. (2000). *Visual Intelligence*. New York: W.W. Norton & Company.
- Hol, K., Koene, A.R., & van Ee, R. (2003). Attention-biased multi-stable surface perception in three-dimensional structure-from-motion. *J Vis*, 3 (7), 486-498.
- Hollins, M. (1980). The effect of contrast on the completeness of binocular rivalry suppression. *Percept Psychophys*, 27 (6), 550-556.
- Hollins, M., & Hudnall, K. (1980). Adaptation of the binocular rivalry mechanism. *Invest Ophth Vis Sci*, 19 (9), 1117-1120.
- Houweling, A.R., & Brecht, M. (2008). Behavioural report of single neuron stimulation in somatosensory cortex. *Nature*, 451 (7174), 65-68.
- Howard, I.P., & Rogers, B.J. (2002). *Seeing in Depth (2: Depth Perception)*. Toronto; I Porteous.
- Hubel, D.H., & Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160, 106-154.
- Hubel, D.H., & Wiesel, T.N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *J Neurophysiol*, 26, 994-1002.
- Hubel, D.H., & Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J Physiol*, 195 (1), 215-243.
- Hubel, D.H., & Wiesel, T.N. (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *J Comp Neurol*, 146 (4), 421-450.
- Hubel, D.H., Wiesel, T.N., & Stryker, M.P. (1977). Orientation columns in macaque monkey visual cortex demonstrated by the 2-deoxyglucose autoradiographic technique. *Nature*, 269 (5626), 328-330.
- Hubel, D.H., Wiesel, T.N., & Stryker, M.P. (1978). Anatomical demonstration of orientation columns in macaque monkey. *J Comp Neurol*, 177 (3), 361-380.
- Huber, D., Petreanu, L., Ghitani, N., Ranade, S., Hromádka, T., Mainen, Z., & Svoboda, K. (2008). Sparse optical microstimulation in barrel cortex drives learned behaviour in freely moving mice. *Nature*, 451 (7174), 61-64.
- Hupe, J.-M., & Rubin, N. (2003). The dynamics of bi-stable alternation in ambiguous motion



- displays: a fresh look at plaids. *Vision Res*, 43 (5), 531-548.
- Jaekl, P.M., & Harris, L.R. (2007). Auditory-visual temporal integration measured by shifts in perceived temporal location. *Neurosci Lett*, 417 (3), 219-224.
- Jones, M.R., & McAuley, J.D. (2005). Time judgments in global temporal contexts. *Percept Psychophys*, 67 (3), 398-417.
- Kalarickal, G., & Marshall, J. (2000). Neural model of temporal and stochastic properties of binocular rivalry. *Neurocomputing*, 32-33, 843-853.
- Kamphuisen, A., van Wezel, R.J.A., & van Ee, R. (2007). Inter-ocular transfer of stimulus cueing in dominance selection at the onset of binocular rivalry. *Vision Res*, 47 (9), 1142-1144.
- Kanai, R., Bahrami, B., & Rees, G. (2010). Human Parietal Cortex Structure Predicts Individual Differences in Perceptual Rivalry. *Curr Biol*, 20, 1626-1630.
- Kanai, R., Knapen, T.H.J., van Ee, R., & Verstraten, F.A.J. (2007a). Disruption of implicit perceptual memory by intervening neutral stimuli. *Vision Res*, 47 (20), 2675-2683.
- Kanai, R., Paffen, C.L.E., Hogendoorn, H., & Verstraten, F.A.J. (2006). Time dilation in dynamic visual display. *J Vis*, 6 (12), 1421-1430.
- Kanai, R., Sheth, B.R., Verstraten, F.A.J., & Shimojo, S. (2007b). Dynamic perceptual changes in audiovisual simultaneity. *PLoS ONE*, 2 (12), e1253.
- Kanai, R., & Verstraten, F.A.J. (2005). Perceptual manifestations of fast neural plasticity: motion priming, rapid motion aftereffect and perceptual sensitization. *Vision Res*, 45 (25-26), 3109-3116.
- Kanai, R., & Verstraten, F.A.J. (2006). Attentional modulation of perceptual stabilization. *Proc Biol Sci*, 273 (1591), 1217-1222.
- Kanai, R., & Watanabe, M. (2006). Visual onset expands subjective time. *Percept Psychophys*, 68 (7), 1113-1123.
- Kanizsa, G. (1979). *Organization in Vision*. New York; Praeger.
- Kanizsa, G., Renzi, P., Conte, S., Compostela, C., & Guerani, L. (1993). Amodal completion in mouse vision. *Perception*, 22 (6), 713-721.
- Katz, L.C., & Crowley, J.C. (2002). Development of cortical circuits: lessons from ocular dominance columns. *Nat Rev Neurosci*, 3 (1), 34-42.
- Katzner, S., Nauhaus, I., Benucci, A., Bonin, V., Ringach, D.L., & Carandini, M. (2009). Local origin of field potentials in visual cortex. *Neuron*, 61 (1), 35-41.
- Keck, T., Mrsic-Flogel, T.D., Vaz Afonso, M., Eysel, U.T., Bonhoeffer, T., & Hübener, M. (2008). Massive restructuring of neuronal circuits during functional reorganization of adult visual



- cortex. *Nat Neurosci*, 11 (10), 1162-1167.
- Keetels, M., Stekelenburg, J., & Vroomen, J. (2007). Auditory grouping occurs prior to inter-sensory pairing: evidence from temporal ventriloquism. *Exp Brain Res*, 180 (3), 449-456.
- Kelly, R.C., Smith, M.A., Kass, R.E., & Lee, T.S. (2010). Local field potentials indicate network state and account for neuronal response variability. *J Comp Neurosci*, [Epub ahead of print].
- Khayat, P.S., Niebergall, R., & Martinez-Trujillo, J.C. (2010). Frequency-dependent attentional modulation of local field potential signals in macaque area MT. *J Neurosci*, 30 (20), 7037-7048.
- Kirmizi-Alsan, E., Bayraktaroglu, Z., Gurvit, H., Keskin, Y.H., Emre, M., & Demiralp, T. (2006). Comparative analysis of event-related potentials during Go/NoGo and CPT: decomposition of electrophysiological markers of response inhibition and sustained attention. *Brain Res*, 1104 (1), 114-128.
- Klink, P.C., Brascamp, J.W., Blake, R., & van Wezel, R.J.A. (2010). Experience-Driven Plasticity in Binocular Vision. *Curr Biol*, 20 (16), 1464-1469.
- Klink, P.C., Noest, A.J., Holten, V., van den Berg, A.V., & van Wezel, R.J.A. (2009). Occlusion-related lateral connections stabilize kinetic depth stimuli through perceptual coupling. *J Vis*, 9 (10), 20:21-20.
- Klink, P.C., van Ee, R., Nijs, M.M., Brouwer, G.J., Noest, A.J., & van Wezel, R.J.A. (2008a). Early interactions between neuronal adaptation and voluntary control determine perceptual choices in bistable vision. *J Vis*, 8 (5), 16:11-18.
- Klink, P.C., van Ee, R., & van Wezel, R.J.A. (2008b). General Validity of Levelt's Propositions Reveals Common Computational Mechanisms for Visual Rivalry. *PLoS ONE*, 3 (10), e3473.
- Knapen, T., Brascamp, J., Adams, W.J., & Graf, E.W. (2009). The spatial scale of perceptual memory in ambiguous figure perception. *J Vis*, 9 (13), 16:11-12.
- Knapen, T.H.J., Kanai, R., Brascamp, J.W., van Boxtel, J.J.A., & van Ee, R. (2007a). Distance in feature space determines exclusivity in visual rivalry. *Vision Res*, 47 (26), 3269-3275.
- Knapen, T.H.J., van Ee, R., & Blake, R. (2007b). Stimulus motion propels traveling waves in binocular rivalry. *PLoS ONE*, 2, e739.
- Koch, C. (2004). *The Quest for Consciousness*. Roberts & Company Publishers.
- Koch, C., & Segev, I. (2000). The role of single neurons in information processing. *Nat Neurosci*, 3 Suppl, 1171-1177.
- Kohn, A. (2007). Visual adaptation: physiology, mechanisms, and functional benefits. *J Neurophysiol*, 97 (5), 3155-3164.



- Kornmeier, J., Ehm, W., Bigalke, H., & Bach, M. (2007). Discontinuous presentation of ambiguous figures: How interstimulus-interval durations affect reversal dynamics and ERPs. *Psychophys*, 44 (4), 552-560.
- Kornmeier, J., Hein, C.M., & Bach, M. (2009). Multistable perception: when bottom-up and top-down coincide. *Brain Cognition*, 69 (1), 138-147.
- Kourtzi, Z., Krekelberg, B., & van Wezel, R.J.A. (2008). Linking form and motion in the primate brain. *Trends Cogn Sci*, 12 (6), 230-236.
- Kovács, I., Papathomas, T.V., Yang, M., & Fehér, A. (1996). When the brain changes its mind: interocular grouping during binocular rivalry. *Proc Natl Acad Sci USA*, 93 (26), 15508-15511.
- Krekelberg, G., van Wezel, R.J., & Albright, T.D. (2006). Interactions between speed and contrast tuning in the middle temporal area: implications for the neural code for speed. *J Neurosci*, 26 (35), 8988-8998.
- Lack, L.C. (1978). *Selective attention and the control of binocular rivalry*. The Hague, The Netherlands; Mouton.
- Laing, C.R., & Chow, C.C. (2002). A spiking neuron model for binocular rivalry. *J Comp Neurosci*, 12 (1), 39-53.
- Lamme, V.A., Supèr, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Curr Opin Neurobiol*, 8 (4), 529-535.
- Lamsa, K.P., Heeroma, J.H., Somogyi, P., Rusakov, D.A., & Kullmann, D.M. (2007). Anti-Hebbian long-term potentiation in the hippocampal feedback inhibitory circuit. *Science*, 315 (5816), 1262-1266.
- Landy, M.S., Maloney, L.T., Johnston, E.B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Res*, 35 (3), 389-412.
- Lankheet, M.J.M. (2006). Unraveling adaptation and mutual inhibition in perceptual rivalry. *J Vis*, 6 (4), 304-310.
- Lankheet, M.J.M., Borghuis, B.G., & van de Grind, W.A. (2005). Spike Timing Precision: Stimulus Dependence and Predictability. *Inv OphthVis Sci*, 46, E-Abstract 5684.
- Lee, J., & Maunsell, J.H.R. (2009). A normalization model of attentional modulation of single unit responses. *PLoS ONE*, 4 (2), e4651.
- Lee, S.-H., & Blake, R. (2004). A fresh look at interocular grouping during binocular rivalry. *Vision Res*, 44 (10), 983-991.
- Lee, S.-H., Blake, R., & Heeger, D.J. (2005). Traveling waves of activity in primary visual cortex during binocular rivalry. *Nat Neurosci*, 8 (1), 22-23.



- Lehky, S.R. (1988). An astable multivibrator model of binocular rivalry. *Perception*, 17 (2), 215-228.
- Leopold, D.A., & Logothetis, N.K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, 379 (6565), 549-553.
- Leopold, D.A., & Logothetis, N.K. (1999). Multistable phenomena: changing views in perception. *Trends Cogn Sci*, 3 (7), 254-264.
- Leopold, D.A., Wilke, M., Maier, A., & Logothetis, N.K. (2002). Stable perception of visually ambiguous patterns. *Nat Neurosci*, 5 (6), 605-609.
- Levelt, W.J. (1965). *On Binocular Rivalry*. Soesterberg; Mouton & Co.
- Levelt, W.J. (1966). The alternation process in binocular rivalry. *British Journal of Psychology*, 57, 225-238.
- Li, H.C., & Kingdom, F.A. (1999). Feature specific segmentation in perceived structure-from-motion. *Vision Res*, 39 (5), 881-886.
- Ling, S., & Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Res*, 46 (8-9), 1210-1220.
- Liu, J., & Newsome, W.T. (2006). Local field potential in cortical area MT: stimulus tuning and behavioral correlations. *J Neurosci*, 26 (30), 7779-7790.
- Livingstone, M.S. (1996). Differences between stereopsis, interocular correlation and binocularity. *Vision Res*, 36 (8), 1127-1140.
- Livingstone, M.S., & Hubel, D.H. (1984). Specificity of intrinsic connections in primate primary visual cortex. *J Neurosci*, 4 (11), 2830-2835.
- Logothetis, N.K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci*, 23 (10), 3963-3971.
- Logothetis, N.K., Leopold, D.A., & Sheinberg, D.L. (1996). What is rivalling during binocular rivalry? *Nature*, 380 (6575), 621-624.
- Logothetis, N.K., & Schall, J.D. (1989). Neuronal correlates of subjective visual perception. *Science*, 245 (4919), 761-763.
- Long, G., & Toppino, T. (2004). Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychol Bull*, 130 (5), 748-768.
- Lyons, G., Sanabria, D., Vatakis, A., & Spence, C. (2006). The modulation of crossmodal integration by unimodal perceptual grouping: a visuotactile apparent motion study. *Exp Brain Re*, 174 (3), 510-516.
- Maier, A., Logothetis, N.K., & Leopold, D.A. (2007). Context-dependent perceptual modula-



- tion of single neurons in primate visual cortex. *Proc Natl Acad Sci USA*, 104 (13), 5620-5625.
- Maier, A., Wilke, M., Logothetis, N.K., & Leopold, D.A. (2003). Perception of temporally interleaved ambiguous patterns. *Curr Biol*, 13 (13), 1076-1085.
- Maij, F., Brenner, E., & Smeets, J.B.J. (2009). Temporal information can influence spatial localization. *J Neurophysiol*, 102 (1), 490-495.
- Maimon, G., & Assad, J.A. (2009). Beyond Poisson: increased spike-time regularity across primate parietal cortex. *Neuron*, 62 (3), 426-440.
- Malach, R., Schirman, T.D., Harel, M., Tootell, R.B., & Malonek, D. (1997). Organization of intrinsic connections in owl monkey area MT. *Cereb Cortex*, 7 (4), 386-393.
- Martin, K.A., & Whitteridge, D. (1984). Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. *J Physiol*, 353, 463-504.
- Masse, N.Y., & Cook, E.P. (2008). The effect of middle temporal spike phase on sensory encoding and correlates with behavior during a motion-detection task. *J Neurosci*, 28 (6), 1343-1355.
- Mauk, M.D., & Buonomano, D.V. (2004). The neural basis of temporal processing. *Annu Rev Neurosci*, 27, 307-340.
- Maunsell, J.H., & Van Essen, D.C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J Neurophysiol*, 49 (5), 1148-1167.
- Maunsell, J.H.R., & Cook, E.P. (2002). The role of attention in visual processing. *Philos Trans R Soc Lond B Biol Sci*, 357 (1424), 1063-1072.
- Mayo, J.P., & Sommer, M.A. (2008). Neuronal adaptation caused by sequential visual stimulation in the frontal eye field. *J Neurophysiol*, 100 (4), 1923-1935.
- Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J Vis*, 4 (7), 539-551.
- Miller, C.T., Dibble, E., & Hauser, M.D. (2001). Amodal completion of acoustic signals by a nonhuman primate. *Nat Neurosci*, 4 (8), 783-784.
- Miller, G. (2006). Optogenetics. Shining new light on neural circuits. *Science*, 314 (5806), 1674-1676.
- Miller, S.M., Hansell, N.K., Ngo, T.T., Liu, G.B., Pettigrew, J.D., Martin, N.G., & Wright, M.J. (2010). Genetic contribution to individual variation in binocular rivalry rate. *Proc Natl Acad Sci USA*, 107 (6), 2664-2668.
- Mitchell, J.F., Stoner, G.R., & Reynolds, J.H. (2004). Object-based attention determines domi-



- nance in binocular rivalry. *Nature*, 429 (6990), 410-413.
- Mitchell, J.F., Sundberg, K.A., & Reynolds, J.H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron*, 55 (1), 131-141.
- Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cogn Brain Res*, 17 (1), 154-163.
- Moreno-Bote, R., Rinzel, J., & Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *J Neurophysiol*, 98 (3), 1125-1139.
- Morrone, M.C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nat Neurosci*, 8 (7), 950-954.
- Mueller, T.J. (1990). A physiological model of binocular rivalry. *Vis Neurosci*, 4 (1), 63-73.
- Mueller, T.J., & Blake, R. (1989). A fresh look at the temporal dynamics of binocular rivalry. *Biol Cybern*, 61 (3), 223-232.
- Murray, M.M., Foxe, D.M., Javitt, D.C., & Foxe, J.J. (2004). Setting boundaries: brain dynamics of modal and amodal illusory shape completion in humans. *J Neurosci*, 24 (31), 6898-6903.
- Nadler, J.W., Angelaki, D.E., & Deangelis, G.C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature*, 452 (7187), 642-645.
- Nagasaka, Y., & Wasserman, E.A. (2008). Amodal completion of moving objects by pigeons. *Perception*, 37 (4), 557-570.
- Nakayama, K., Shimojo, S., & Silverman, G.H. (1989). Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, 18 (1), 55-68.
- Nawrot, M., & Blake, R. (1989). Neural integration of information specifying structure from stereopsis and motion. *Science*, 244 (4905), 716-718.
- Nawrot, M., & Blake, R. (1991a). The interplay between stereopsis and structure from motion. *Percept Psychophys*, 49 (3), 230-244.
- Nawrot, M., & Blake, R. (1991b). A neural network model of kinetic depth. *Vis Neurosci*, 6 (3), 219-227.
- Nelson, S.B. (2004). Hebb and anti-Hebb meet in the brainstem. *Nat Neurosci*, 7 (7), 687-688.
- Noest, A.J., van Ee, R., Nijs, M.M., & van Wezel, R.J.A. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *J Vis*, 7 (8), 10:1-14.
- O'Shea, R.P., Sims, A.J., & Govan, D.G. (1997). The effect of spatial frequency and field size on the spread of exclusive visibility in binocular rivalry. *Vision Res*, 37 (2), 175-183.



- Ooi, T., & He, Z. (1999). Binocular rivalry and visual awareness: the role of attention. *Perception*, 28 (5), 551-574.
- Orbach, J., Ehrlich, D., & Vainstein, E. (1963). Reversibility of the Necker Cube: Iii. Effects of Interpolation on Reversal Rate of the Cube Presented Repetitively. *Percept Mot Skills*, 17, 571-582.
- Orbach, J., Zucker, E., & Olson, R. (1966). Reversibility of the Necker Cube. VII. Reversal rate as a function of figure-on and figure-off duration. *Percept Mot Skills*, 17, 615-618.
- Paffen, C., Alais, D., & Verstraten, F.A.J. (2006). Attention speeds binocular rivalry. *Psychol Sci*, 17 (9), 752-756.
- Paffen, C.L.E., Naber, M., & Verstraten, F.A.J. (2008). The spatial origin of a perceptual transition in binocular rivalry. *PLoS ONE*, 3 (6), e2311.
- Pariyadath, V., & Eagleman, D.M. (2008). Brief subjective durations contract with repetition. *J Vis*, 8 (16), 11:1-6.
- Park, J., Schlag-Rey, M., & Schlag, J. (2003). Voluntary action expands perceived duration of its sensory consequence. *Exp Brain Res*, 149 (4), 527-529.
- Parker, A. (2007). Binocular depth perception and the cerebral cortex. *Nat Rev Neurosci*, 8, 379-391.
- Parker, A., Krug, K., & Cumming, B. (2002). Neuronal activity and its links with the perception of multi-stable figures. *Philos Trans R Soc Lond B Biol Sci*, 357 (1424), 1053-1062.
- Parker, A., & Newsome, W.T. (1998). Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci*, 21, 227-277.
- Pastukhov, A., & Braun, J. (2008). A short-term memory of multi-stable perception. *J Vis*, 8 (13), 7:1-14.
- Pearson, J., & Brascamp, J. (2008). Sensory memory for ambiguous vision. *Trends Cogn Sci*, 12 (9), 334-341.
- Pearson, J., & Clifford, C.W.G. (2004). Determinants of visual awareness following interruptions during rivalry. *J Vis*, 4 (3), 196-202.
- Pearson, J., & Clifford, C.W.G. (2005). When your brain decides what you see: grouping across monocular, binocular, and stimulus rivalry. *Psychol Sci*, 16 (7), 516-519.
- Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10 (4), 437-442.
- Penney, T.B., Gibbon, J., & Meck, W.H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *J Exp Psychol Hum Percept Perform*, 26 (6), 1770-1787.



- Penney, T.B., Meck, W.H., Roberts, S.A., Gibbon, J., & Erlenmeyer-Kimling, L. (2005). Interval-timing deficits in individuals at high risk for schizophrenia. *Brain Cogn*, 58 (1), 109-118.
- Penton-Voak, I.S., Edwards, H., Percival, A., & Wearden, J.H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *J Exp Psychol Anim B*, 22 (3), 307-320.
- Peterson, M., & Hochberg, J. (1983). Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception. *J Exp Psychol Hum Percept Perform*, 9, 183-193.
- Pettigrew, J.D. (2001). Searching for the Switch: Neural Bases for Perceptual Rivalry Alternations. *Brain Mind*, 2, 85-118.
- Pinna, B. (2008). A new perceptual problem: the amodal completion of color. *Vis Neurosci*, 25 (3), 415-422.
- Qian, N. (1997). Binocular disparity and the perception of depth. *Neuron*, 18 (3), 359-368.
- Quijan Quiroga, R., & Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat Rev Neurosci*, 10 (3), 173-185.
- Raemaekers, M., van der Schaaf, M., van Ee, R., & van Wezel, R. (2008). Widespread fMRI activity differences between perceptual states in visual rivalry are correlated with differences in observer biases. *Brain Res*, 1252, 161-171.
- Ramachandran, V.S., & Anstis, S.M. (1983). Perceptual organization in moving patterns. *Nature*, 304 (5926), 529-531.
- Rauschenberger, R., Liu, T., Slotnick, S.D., & Yantis, S. (2006). Temporally unfolding neural representation of pictorial occlusion. *Psychol Sci*, 17 (4), 358-364.
- Recanzone, G.H. (2003). Auditory influences on visual temporal rate perception. *J Neurophysiol*, 89 (2), 1078-1093.
- Repp, B.H., & Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *J Exp Psychol Hum Percept Perform*, 28 (5), 1085-1099.
- Reynolds, J.H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annu Rev Neurosci*, 27, 611-647.
- Reynolds, J.H., & Heeger, D.J. (2009). The normalization model of attention. *Neuron*, 61 (2), 168-185.
- Rockland, K.S., & Lund, J.S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *J Comp Neurol*, 216 (3), 303-318.



- Roe, A.W., Parker, A.J., Born, R.T., & Deangelis, G.C. (2007). Disparity channels in early vision. *J Neurosci*, 27 (44), 11820-11831.
- Roelfsema, P. (2006). Cortical Algorithms for Perceptual Grouping. *Annu Rev Neurosci*, 29, 203-227.
- Rose, D., & Summers, J. (1995). Duration illusions in a train of visual stimuli. *Perception*, 24 (10), 1177-1187.
- Rousseau, L., & Rousseau, R. (1996). Stop-reaction time and the internal clock. *Percept Psychophys*, 58 (3), 434-448.
- Roy, J.P., Komatsu, H., & Wurtz, R.H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *J Neurosci*, 12 (7), 2478-2492.
- Rubin, N. (2001). Figure and ground in the brain. *Nat Neurosci*, 4 (9), 857-858.
- Rubin, N., & Hupe, J.-M. (2005). Dynamics of Perceptual Bistability: Plaids and Binocular Rivalry Compared. In: D. Alais & R. Blake (Eds.), *Binocular Rivalry* (pp. 137-154). Cambridge, MA; The MIT Press.
- Rust, N.C., & Movshon, J.A. (2005). In praise of artifice. *Nat Neurosci*, 8 (12), 1647-1650.
- Saalmann, Y., Pigarev, I., & Vidyasagar, T. (2007). Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. *Science*, 316 (5831), 1612-1615.
- Sanabria, D., Soto-Faraco, S., Chan, J.C., & Spence, C. (2005). Intramodal perceptual grouping modulates multisensory integration: evidence from the crossmodal dynamic capture task. *Neurosci Lett*, 377 (1), 59-64.
- Sanabria, D., Soto-Faraco, S., Chan, J.S., & Spence, C. (2004). When does visual perceptual grouping affect multisensory integration? *Cogn Affect Beh Neurosci*, 4 (2), 218-229.
- Schwartz, B.L., & Sperling, G. (1983). Luminance controls the perceived 3-D structure of dynamic 2-D displays. *Bull Psychonom Soc*, 21 (6), 456-458.
- Sejnowski, T.J. (2003). The once and future of Hebb synapse. *Can Psychol*, 44 (1), 17-20.
- Sengpiel, F., Blakemore, C., & Harrad, R. (1995). Interocular suppression in the primary visual cortex: a possible neural basis of binocular rivalry. *Vision Res*, 35 (2), 179-195.
- Sereno, M.E., & Sereno, M.I. (1999). 2-D center-surround effects on 3-D structure-from-motion. *J Exp Psychol: Hum Perc Perf*, 25 (6), 1834-1854.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cogn Brain Res*, 14 (1), 147-152.
- Sheinberg, D.L., & Logothetis, N.K. (1997). The role of temporal cortical areas in perceptual organization. *Proc Natl Acad Sci USA*, 94 (7), 3408-3413.



- Shapiro, A., Curtu, R., Rinzel, J., & Rubin, N. (2007). Dynamical characteristics common to neuronal competition models. *J Neurophysiol*, 97 (1), 462-473.
- Singer, W., & Gray, C.M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci*, 18, 555-586.
- Slotnick, S., & Yantis, S. (2005). Common neural substrates for the control and effects of visual attention and perceptual bistability. *Cogn Brain Res*, 24 (1), 97-108.
- Sohn, W., & Seiffert, A.E. (2006). Motion aftereffects specific to surface depth order: beyond binocular disparity. *J Vis*, 6 (2), 119-131.
- Spillmann, L., & Werner, J.S. (1996). Long-range interactions in visual perception. *Trends Neurosci*, 19 (10), 428-434.
- Stollenwerk, L., & Bode, M. (2003). Lateral neural model of binocular rivalry. *Neural Comput*, 15 (12), 2863-2882.
- Sugita, Y. (1999). Grouping of image fragments in primary visual cortex. *Nature*, 401 (6750), 269-272.
- Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, 421 (6926), 911.
- Sussillo, D., & Abbott, L.F. (2009). Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63 (4), 544-557.
- Suzuki, S., & Grabowecky, M. (2007). Long-term speeding in perceptual switches mediated by attention-dependent plasticity in cortical visual processing. *Neuron*, 56 (4), 741-753.
- Suzuki, S., & Peterson, M. (2000). Multiplicative effects of intention on the perception of bi-stable apparent motion. *Psychol Sci*, 11 (3), 202-209.
- Svoboda, K., & Yasuda, R. (2006). Principles of two-photon excitation microscopy and its applications to neuroscience. *Neuron*, 50 (6), 823-839.
- Thompson, P.G., & Movshon, J.A. (1978). Storage of spatially specific threshold elevation. *Perception*, 7 (1), 65-73.
- Tiesinga, P., Fellous, J., & Sejnowski, T. (2008). Regulation of spike timing in visual cortical circuits. *Nat Rev Neurosci*, 9, 97-107.
- Tong, F., Meng, M., & Blake, R. (2006). Neural bases of binocular rivalry. *Trends Cogn Sci*, 10 (11), 502-511.
- Toppino, T. (2003). Reversible-figure perception: mechanisms of intentional control. *Percept Psychophys*, 65 (8), 1285-1295.
- Treue, S., Andersen, R.A., Ando, H., & Hildreth, E.C. (1995). Structure-from-motion: percep-



- tual evidence for surface interpolation. *Vision Res*, 35 (1), 139-148.
- Treue, S., Husain, M., & Andersen, R.A. (1991). Human perception of structure from motion. *Vision Res*, 31 (1), 59-75.
- Treue, S., & Martinez Trujillo, J. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399 (6736), 575-579.
- Treue, S., & Maunsell, J.H.R. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *J Neurosci*, 19 (17), 7591-7602.
- Tso, D.Y., Gilbert, C.D., & Wiesel, T.N. (1986). Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *J Neurosci*, 6 (4), 1160-1170.
- Tzounopoulos, T., Kim, Y., Oertel, D., & Trussell, L.O. (2004). Cell-specific, spike timing-dependent plasticities in the dorsal cochlear nucleus. *Nat Neurosci*, 7 (7), 719-725.
- Uka, T., & Deangelis, G.C. (2003). Contribution of middle temporal area to coarse depth discrimination: comparison of neuronal and psychophysical sensitivity. *J Neurosci*, 23 (8), 3515-3530.
- Ullman, S. (1979). *The Interpretation of Visual Motion*. Cambridge, MA; The MIT Press.
- 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.
- van Dam, L., & van Ee, R. (2005). The role of (micro)saccades and blinks in perceptual bi-stability from slant rivalry. *Vision Res*, 45 (18), 2417-2435.
- van Dam, L., & van Ee, R. (2006a). Retinal image shifts, but not eye movements per se, cause alternations in awareness during binocular rivalry. *J Vis*, 6 (11), 1172-1179.
- van Dam, L., & van Ee, R. (2006b). The role of saccades in exerting voluntary control in perceptual and binocular rivalry. *Vision Res*, 46 (6-7), 787-799.
- van de Grind, W.A., van der Smagt, M.J., & Verstraten, F.A.J. (2004). Storage for free: a surprising property of a simple gain-control model of motion aftereffects. *Vision Res*, 44 (19), 2269-2284.
- van der Smagt, M.J., & Stoner, G.R. (2008). Occlusion and the solution to visual motion ambiguity: Looking beyond the aperture problem. *J Vis*, 8 (2), 4:1-12.
- van Ee, R. (2005). Dynamics of perceptual bi-stability for stereoscopic slant rivalry and a comparison with grating, house-face, and Necker cube rivalry. *Vision Res*, 45 (1), 29-40.
- van Ee, R., Noest, A.J., Brascamp, J.W., & van den Berg, A.V. (2006). Attentional control over either of the two competing percepts of ambiguous stimuli revealed by a two-parameter



- analysis: means do not make the difference. *Vision Res*, 46 (19), 3129-3141.
- 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. *J Neurosci*, 29 (37), 11641-11649.
- van Ee, R., van Dam, L.C., & Brouwer, G.J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Res*, 45 (1), 41-55.
- van Ee, R., van Dam, L.C.J., & Erkelens, C.J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *J Vis*, 2 (9), 597-607.
- van Wassenhove, V., Buonomano, D.V., Shimojo, S., & Shams, L. (2008). Distortions of subjective time perception within and across senses. *PLoS ONE*, 3 (1), e1437.
- van Wezel, R.J.A., & Britten, K.H. (2002). Motion adaptation in area MT. *J Neurophysiol*, 88 (6), 3469-3476.
- Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the ‘unity assumption’ using audiovisual speech stimuli. *Percept Psychophys*, 69 (5), 744-756.
- Vatakis, A., & Spence, C. (2008). Evaluating the influence of the ‘unity assumption’ on the temporal perception of realistic audiovisual stimuli. *Acta Psychol*, 127 (1), 12-23.
- Verstraten, F.A., Fredericksen, R.E., Grüsser, O.J., & van de Grind, W.A. (1994). Recovery from motion adaptation is delayed by successively presented orthogonal motion. *Vision Res*, 34 (9), 1149-1155.
- Verstraten, F.A.J., & Ashida, H. (2005). Attention-based motion perception and motion adaptation: what does attention contribute? *Vision Res*, 45 (10), 1313-1319.
- von der Heydt, R., Zhou, H., & Friedman, H.S. (2000). Representation of stereoscopic edges in monkey visual cortex. *Vision Res*, 40 (15), 1955-1967.
- Von Helmholtz, H. (1924). *Treatise on physiological optics*. (J.P.C. Southall, Transl). Rochester, New York; The Optical Society of America.
- Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. *J Exp Psychol: Hum Perc Perf*, 26 (5), 1583-1590.
- Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cogn Brain Res*, 22 (1), 32-35.
- Vul, E., Krizay, E., & MacLeod, D.I.A. (2008). The McCollough effect reflects permanent and transient adaptation in early visual cortex. *J Vis*, 8 (12), 4:1-12.
- Wada, Y., Kitagawa, N., & Noguchi, K. (2003). Audio-visual integration in temporal perception. *Int J Psychophysiol*, 50 (1-2), 117-124.



- Walker, J.T., & Scott, K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones and gaps. *J Exp Psychol: Hum Perc Perf*, 7 (6), 1327-1339.
- Wallace, M.T., Roberson, G.E., Hairston, W.D., Stein, B.E., Vaughan, J.W., & Schirillo, J.A. (2004). Unifying multisensory signals across time and space. *Exp Brain Res*, 158 (2), 252-258.
- Wandell, B.A., Dumoulin, S.O., & Brewer, A.A. (2007). Visual field maps in human cortex. *Neuron*, 56 (2), 366-383.
- Wannig, A., Rodriguez, V., & Freiwald, W.A. (2007). Attention to Surfaces Modulates Motion Processing in Extrastriate Area MT. *Neuron*, 54 (4), 639-651.
- Watanabe, K., Nijhawan, R., Khurana, B., & Shimojo, S. (2001). Perceptual organization of moving stimuli modulates the flash-lag effect. *J Exp Psychol: Hum Perc Perf*, 27 (4), 879-894.
- Watanabe, M., Tanaka, H., Uka, T., & Fujita, I. (2002). Disparity-selective neurons in area V4 of macaque monkeys. *J Neurophysiol*, 87 (4), 1960-1973.
- Watanabe, T., & Cole, R. (1995). Propagation of local motion correspondence. *Vision Res*, 35 (20), 2853-2861.
- Watson, A.B., & Pelli, D.G. (1983). QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys*, 33 (2), 113-120.
- Wearden, J.H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": application of a model of the internal clock in humans. *Q J Exp Psychol B*, 51 (2), 97-120.
- Weigelt, S., Singer, W., & Muckli, L. (2007). Separate cortical stages in amodal completion revealed by functional magnetic resonance adaptation. *BMC Neuroscience*, 8 (70), 1471-2202.
- Welch, R.B., & Warren, D.H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychol Bull*, 88 (3), 638-667.
- Wichmann, F., & Hill, N. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys*, 63 (8), 1293-1313.
- Wilson, H.R. (2003). Computational evidence for a rivalry hierarchy in vision. *Proc Natl Acad Sci USA*, 100 (24), 14499-14503.
- Wilson, H.R. (2007). Minimal physiological conditions for binocular rivalry and rivalry memory. *Vision Res*, 47 (21), 2741-2750.
- Wilson, H.R., Blake, R., & Lee, S.-H. (2001). Dynamics of travelling waves in visual perception. *Nature*, 412 (6850), 907-910.



- Witten, I.B., & Knudsen, E.I. (2005). Why seeing is believing: merging auditory and visual worlds. *Neuron*, 48 (3), 489-496.
- Wolfe, J.M. (1984). Reversing ocular dominance and suppression in a single flash. *Vision Res*, 24 (5), 471-478.
- Womelsdorf, T., Anton-Erxleben, K., Pieper, F., & Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat Neurosci*, 9 (9), 1156-1160.
- Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *J Vis*, 7 (10), 2:1-5.
- Yacoub, E., Harel, N., & Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proc Natl Acad Sci USA*, 105 (30), 10607-10612.
- Yacoub, E., Shmuel, A., Logothetis, N., & Uğurbil, K. (2007). Robust detection of ocular dominance columns in humans using Hahn Spin Echo BOLD functional MRI at 7 Tesla. *Neuroimage*, 37 (4), 1161-1177.
- Yang, Y., & Blake, R. (1995). On the accuracy of surface reconstruction from disparity interpolation. *Vision Res*, 35 (7), 949-960.
- Yang, Y., Rose, D., & Blake, R. (1992). On the variety of percepts associated with dichoptic viewing of dissimilar monocular stimuli. *Perception*, 21 (1), 47-62.
- Yao, H., Shi, L., Han, F., Gao, H., & Dan, Y. (2007). Rapid learning in cortical coding of visual scenes. *Nat Neurosci*, 10 (6), 772-778.
- Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J.C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature*, 414 (6861), 302-305.
- Yazaki-Sugiyama, Y., Kang, S., Câteau, H., Fukai, T., & Hensch, T.K. (2009). Bidirectional plasticity in fast-spiking GABA circuits by visual experience. *Nature*, 462 (7270), 218-221.
- Zhang, F., Gradinaru, V., Adamantidis, A.R., Durand, R., Airan, R.D., de Lecea, L., & Deisseroth, K. (2010). Optogenetic interrogation of neural circuits: technology for probing mammalian brain structures. *Nat Protoc*, 5 (3), 439-456.



# Acknowledgements

Dankwoord





The four years of research that have led to this thesis have sometimes been hectic, busy, educational, challenging, tiresome, or frustrating, but they were generally very enjoyable. There are many people that contributed to these sentiments and they deserve to be properly acknowledged. The problem with explicit acknowledgements, however, is that one is highly likely to forget to mention someone. In an attempt to avoid such a mistake, I would like to start out by thanking everyone I worked with over the past couple of years, and additionally everyone that took the trouble of reading (parts of) this thesis. Even if you started out by just reading this *Acknowledgements* chapter, I am confident that you will eventually find the time and courage to also read the very interesting research chapters that I have worked on so hard for so many hours over the past couple of years. While this first acknowledgement conveniently covers anyone that is likely to notice a forgotten name in the first place, there are a number of people that I would still like to mention a bit more explicitly.

First of all, the group of human and non-human primates that served as observers in one or several of my experiments are responsible for the fact that there was something to write about in the first place. I am grateful for the many hours they spent in the dark, producing data, while staring at a computer screen and/or listening to rather uninteresting beeps. The efforts and endurance of Sukke and Appie are especially appreciated.

I am also grateful for the way my supervisors and advisors have guided me through this PhD project. Richard, you gave me a lot of freedom to address the research questions that I thought would be most worth pursuing. At the same time, you were always prepared to critically question the proposed experimental set-up or the value of my hypotheses. This often forced me to reconsider whether this ‘fantastic’ new idea really made as much sense as I might have initially thought. I remember that in my first week of working at your lab you told me that meeting the requirements for a PhD were first and foremost my own responsibility, but that help and supervision would always be available upon request. For me, this arrangement could not have worked any better. Since I was practically always chasing the scientific puzzles I was most interested in myself, it was very easy to stay motivated and put in the necessary hours of work. When the times for the lab were getting rougher as a result of some rather dramatic external decisions, you kept us well-informed and managed to keep us relatively free of worries about our own working situations. You also taught me a lot about ‘the way things work in science’, which is not always as straightforward and transparent as one would like to think. Your style of loose supervision, with all the support and feedback one could ask for, has probably played an important role in the type of researcher I have become. At the very least, it made for a highly pleasant experience working in your lab.

Bert, unfortunately we never really managed to make our discussion moments as regular as we intended them to be. The times we did meet, however, I greatly benefited both from the necessity to summarize what I had recently been doing and what I intended to do next, as well as from your feedback on these plans. You often had just the right amount of distance from my projects to understand what I wanted to achieve but to also ask the crucial questions that had not even crossed my mind yet. Furthermore, your prompt feedback on many initial versions of my manuscripts has been greatly appreciated.

Martin, when you moved from the other side of the building to the portacabin where my desk was also situated, it quickly became customary to walk into your office for a short chat. You would always make time if I had technical questions about a particular analysis, a programming problem, or a research question for which I was looking for the proper (old) refer-



ences of which I was certain that they had to exist. We also talked about less scientific topics such as interesting mountainbike trails in the nearby forests, GPS software, the often frustrating developments in the world of academics, or simply the news of the moment. I really enjoyed teaching the practical session on the neurophysiology of the catfish electroreceptor with you, and I hope you will be able to successfully introduce this subject in the curriculum of Wageningen University where you are working now.

The longest conversations I have had at the lab's coffee maker invariably involved André. It was not exceptional to walk into the meeting room to get a cup of coffee and to get caught in some lively and highly interesting conversation, only to return to my desk two or three cups later. These conversations often took the form of long and passionate pseudo-monologues, that had in common that they usually had very little to do with my thesis work. They were, however, really fun distractions from the problems that I spent most of my time on. Besides, a little bit of general knowledge has never hurt anyone, has it? On a more scientific note, I learned a lot from discussing my ideas and data with you. Your computational model of perceptual rivalry has played an important role in the way I think about the neuronal underpinnings of conscious visual perception and your near encyclopedic knowledge of the neurophysiology and perception literature has saved me a lifetime of searching for relevant literature.

Raymond, your enthusiasm at the start of every single project has been inspiring. The many discussions we have had over the years have greatly influenced the way in which I approach and formulate my scientific questions, hypotheses, and explanatory mechanisms. Early in my PhD project, your feedback has significantly improved my writing skills, while your tendency to look beyond the borders of our university has encouraged me to visit many national and international conferences and keep in contact with colleagues all over the world.

My interest in neuroscience in general, and in sensory systems in particular, really began to grow during my third year of studying biology. Wim van de Grind's course *Perception and Behaviour* contained lectures on the echolocation system of bats and the closed-loop sensorimotor system by which toads identify and catch their prey. At that moment, I realized that the effortlessness of sensation starkly contrasts with the elegant complexity of the underlying neural mechanisms, and I decided that I wanted to learn more about these mechanisms. During the years of my thesis work, Wim kept coming to our lab-meetings to provide feedback on the discussed work, mock the fMRI 'revolution', and remind us all of the relevant experimental work that was often already performed in the years before I was even born. Wim, your input has always been greatly appreciated and I can only hope that the adult brain contains enough neural plasticity for you to cope with the consequences of the recent unfortunate events.

The environment in which one's PhD research is performed is an important influence that shapes one's style of dealing with scientific problems. Utrecht University has a strong tradition in perception research with a clear mechanistic way of approaching the problems. Thinking about solutions in terms of models and neuronal mechanisms seems natural to me now, but I realize that this is probably a (partial) result of the many lab-meetings our neurobiology group used to have together with the Physics of Man group that also worked on (visual) perception. Apart from the weekly lab-meetings, these were also the people with whom we organized lab-barbecues and afternoons of watching soccer matches in the conference room. Both the formal and the informal contributions of Jan, Tomas, Jeroen, Dagmar, Gijs, Maartje, Casper, Raymond, Bert, Wim, Douwe, and Ger-Jan are greatly appreciated. The 'NEST', our



idyllically situated lab, was the best place to work and I would like to thank my fellow NEST-inhabitants Richard, Martin, André, Erik, Matthijs, Anna, Jonathan, Jacob, Michiel, Roger and Jeanette for creating and maintaining such an awesome atmosphere.

Anna, we shared a room for the entire period of our PhD projects. Since we started our work at approximately the same time, we could often share the frustrations and joys of doing neurobiological research. It was great to be able to directly discuss data analysis issues and test possible explanations against a critical and well-informed one-person audience. At the time of writing this chapter, you have already successfully defended a very nice thesis and I hope to follow your example shortly.

It is impossible to mention the NEST and not think of Theo. The number of times Theo explicitly solicited for an honorary mentioning in the *Acknowledgements* section of my thesis almost drove me to the point of ‘forgetting’ to mention his name altogether. However, with his unique sense of humor (interpret that any way you want), the never-ending stream of stories from the olden days (especially after a few drinks), and his general likability he has been an important determinant the NEST’s character and he deserves to be mentioned. At the NEST, animal caretaking was in the competent hands of Theo, Hans and Henk, while Nico and Joost performed peri-surgical assistance. From the Kruyt building, skilled technical support was provided by Ed, while Miriam helped with administrative tasks. Also situated in the Kruyt building were our ‘fish-research’ colleagues Rob, Frank and Lonneke. Thank you all for your general support and the more technical help during the catfish practical sessions.

It is certainly true that you can learn a lot about your own research from supervising good students on longer research projects. The excellent work that was done by Miesje, Vivian, Jorrit and Luuk has played an important role in the formation of this thesis. By now, some of you have already started your own PhD work and I am confident that you will produce very nice publications and dissertations in the near and slightly more distant future.

The Helmholtz Institute, of which our Neurobiology group was part, has been an inspiring surrounding to study the neural mechanisms of visual perception. Without going through an extensive list of names, I thank all Helmholtz researchers for creating a strong basis for perception research. I have always greatly enjoyed the Helmholtz lectures, retreats and informal gatherings. As a member of the Helmholtz PhD Students Council, I had great fun working together with Roger, Tanja, Titia, Mijke, Jeroen, Maarten, Mirela, Krista, Christa, Willemijn, Devika, Chris, Dennis and Alexander.

The two weeks I spent at the Visual Neurosciences Bootcamp in the Rauischholzhausen castle were fantastic, and I learned a lot. I thank the organizers Karl, Jochen and Frank for putting such a nice educational program together, but more importantly, I thank my fellow attending students for creating such an awesome vibe and for continuing to do just that in the many international meetings that followed our initial period in ‘Castle Curly Lettuce’. Arjen, Alexander, Marianne, Donatas, Matteo, Elisabeth, Anna, Milena, Ariel, Anja, Carsten, Gesche, Tom, Yixing, Femke, Toni, Jan, Anushree, Katrin, Therese, Jose, Paola, Linus, Patrick, Chris, Elisabeth, Sebastien, Patricia, Erin, Jen, Kelly, and Jan, I hope to see you all over drinks and/or science for many more times to come.

One of the chapters in this thesis was initiated at Vanderbilt University in Nashville. I thank Randolph for having me in the lab as a visitor and Jan, Katy, Sam, Xian, Jascha, Janneke, Eunice, Min-Suk, David, Ben, Rosanne, and the other locals for making me feel welcome and including me both in their curricular and extracurricular activities (in retrospect I consider



---

myself very lucky that the taxi-company found my passport and got it back to me so quickly in the middle of the night).

There is a life outside of science as well, but sometimes this is easy to forget. I admit that I have the unfortunate and disturbing habit of wanting to finish every single objective that I set myself right away. Whether it is a novel data analysis approach, the programming of a new experiment, or the revision of a manuscript; once I realize how I want to do things, it more or less has to be done and finished as soon as possible. Such an impatient nature relies upon external influences to stay within sanity ranges and I am extremely thankful for all my friends and family that have kept providing such influences over the years. Thank you so much for the good times, the invitation, the discussions, the fun, the music and the drinks; for your general support, for listening to my scientific rants with either sincere or well-simulated interest, and for reminding me that there is more to life than work. Please keep doing that. I need it sometimes.

Sita, I hope you know how important you are for my mental health. It would take too much space to list all your specific contributions to both me and my thesis, but I would particularly like to thank you for listening, disagreeing and (occasionally) agreeing; for understanding and being critical; for enduring, reviewing, and putting things in perspective; for contributing to a nice, relaxing home; for all the fun, silliness, unconditional support and love. Thank you.



---

## Publications





## Journal papers

- Klink, P.C., van Wezel, R.J.A., & van Ee, R. (In preparation). United we sense, divided we fail: Neural mechanisms of contextual inference in multistable perception. *Invited review article as part of a special issue proposal titled "Multistability in perception: Binding sensory modalities", submitted to Philosophical Transactions of the Royal Society B: Biological Sciences.*
- Oleksiak, A., Klink, P.C., Postma, A., van der Ham, I.J.M., Lankheet, M.J.M. & van Wezel, R.J.A. (In preparation). Parietal area 7a acts as a matched filter in visual working memory.
- Klink, P.C., Oleksiak, A., Lankheet, M.J.M., & Van Wezel, R.J.A. (Submitted). Intermittent stimulus presentation stabilizes neuronal responses in macaque area MT.
- Lankheet, M.J.M., Klink, P.C., Borghuis, B.G., & Noest, A.J. (Submitted). Quasi-periodic spiking allows temporal information encoding by interspike intervals in catfish electroreceptor.
- Oleksiak, A., Klink, P.C., Postma, A., van der Ham, I.J.M., Lankheet, M.J.M. & van Wezel, R.J.A. (Under revision). Spatial summation in macaque parietal area 7a follows a winner-take-all rule. *Journal of Neurophysiology*.
- Klink, P.C., Montijn, J.S., & Van Wezel, R.J.A. (2010). Crossmodal duration perception involves perceptual grouping, temporal ventriloquism and variable internal clock rates. *Attention, Perception & Psychophysics*, [Epub ahead of print].
- Oleksiak, A., Postma, A., van der Ham, I.J.M., Klink, P.C., & van Wezel, R.J.A. (2010). A review of lateralization of spatial functioning in nonhuman primates. *Brain Research Reviews*, [Epub ahead of print].
- Klink, P.C., Brascamp, J.W., Blake, R., & Van Wezel, R.J.A. (2010). Experience-driven plasticity in binocular vision. *Current Biology*, 20(16), 1464-1469.
- Klink, P.C., Noest, A.J., Holten, V., Van den Berg, A.V., & Van Wezel, R.J.A. (2009). Occlusion-related lateral connections stabilize kinetic depth through perceptual coupling. *Journal of Vision*, 9(10):20, 1-20.
- Klink, P.C., van Ee, R., & van Wezel, R.J.A. (2008). General validity of Levelt's propositions reveals common computational mechanisms for visual rivalry. *PLoS ONE*, 3(10): e3473.
- Klink, P.C., Van Ee, R., Nijls, 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, 1-18.



Klink, P.C. (2008). Some spikes are more informative than others. *Journal of Neuroscience*, 28(19), 4844-4845.

## Conference abstracts

Klink, P.C., Lankheet, M.J.M., & Van Wezel, R.J.A. (2011). Intermittent motion stimuli stabilize neuronal responses in area MT: Implications for the perceptual stabilization of visual ambiguities. *Vision Science Society Annual Meeting*, Naples FL, USA.

van Wezel, R.J.A., Lankheet, M.J.M., & Klink, P.C. (2010) The time-course of perceptual stabilization induces response stabilization in macaque area MT neurons. *Society for Neuroscience*, 74.12.

Klink, P.C., Noest, A.J., & Lankheet, M.J.M. (2010). Spike-pair triggered averaging reveals a function for spike generation in neural coding. *The Federation of European Neuroscience Societies Forum*, Amsterdam, the Netherlands.

Klink, P.C., Brascamp, J.W., Blake, R., & Van Wezel, R.J.A. (2010). Plasticity of interocular inhibition with prolonged binocular rivalry. *Vision Science Society Annual Meeting*, Naples FL, USA.

Klink, P.C., Lankheet, M.J.M., & Van Wezel, R.J.A. (2009). The neural signature of stabilized ambiguous structure-from-motion in macaque area MT. *Society for Neuroscience*, 558.19.

Klink, P.C. (2009). Perception of ambiguous figures in spatiotemporal context. *8th Dutch EndoNeuroPsycho Meeting*, Doorwerth, the Netherlands.

Montijn, J., Klink, P.C., & Van Wezel R.J.A. (2009). Audiovisual integration in crossmodal duration perception. *8th Dutch EndoNeuroPsycho Meeting*, Doorwerth, the Netherlands.

van der Velden, L., Klink, P.C., & Van Wezel R.J.A. (2009). Long-term adaptation to motion stimuli in area MT of the Macaque Monkey. *8th Dutch EndoNeuroPsycho Meeting*, Doorwerth, the Netherlands.

Klink, P.C., Noest, A.J., & Van Wezel, R.J.A. (2009). Amodal spatial facilitation resolves local ambiguities of kinetic depth. *Journal of Vision*, 9(8):275, 275a.

Klink, P.C., & van Wezel, R.J.A. (2009). Area MT and the stabilized perception of bistable structure-from-motion. *European Primate Neurobiology Meeting*, Goettingen, Germany.



Klink, P.C., Van Ee, R., & Van Wezel, R.J.A. (2008). Levelt's propositions generalized for bistable structure-from-motion: Common computational mechanisms of visual rivalry. *Perception 37 ECVP Abstract Supplement*, page 28.

Klink, P.C., Van Ee, R., Noest, A.J., & Van Wezel, R.J.A. (2007). Early interactions among perceptual history, stimulus timing and voluntary control determine perceptual decisions in bistable vision. Society for Neuroscience, 507.14.

Noest, A.J., Van Ee, R., Van Wezel, R.J.A., Klink, P.C., & Nijs, M.M. (2007). Bistable percept-choice dynamics explained by early interactions between stimulus timing, voluntary bias, and perceptual history. *Perception 36 ECVP Abstract Supplement*.

## Book chapters

Klink, P.C., van Wezel, R.J.A., & van Ee, R. (In preparation). The future of binocular rivalry research: Reaching through a window on consciousness. In: S.M. Miller (Ed.) *The Constitution of Consciousness* (Advances in Consciousness Research) [Provisional title].

## Other publications

Klink, P.C. (2007). Attention vs. Contrast for the Single Neuron: Does the analogy hold? *Journal of Neuroscience*, eLetter.

## Curriculum Vitae

Peter Christiaan Klink was born in Papendrecht, the Netherlands on October 8, 1980. Here, he attended high school at CSG De Lage Waard, where he graduated in 1999. He went on to study Biology (BSc) at Utrecht University and specialized in Neurobiology and Zoology. He was president of the Utrecht Biologists Society in the years 2002/2003. In this period he also served as a member of the Faculty Council of Utrecht University's Biology department and as an editor for the department's periodical *bioSCOPE*.

In 2003, he enrolled in Utrecht University's prestigious MSc program in Neuroscience and Cognition, where he combined the tracks Experimental & Clinical Neuroscience and Cognitive Neuroscience. As a part of this program, he carried out a six-month research internship at the Functional Neurobiology department, where he studied the neural mechanisms of implied motion perception under the supervision of Jeanette Lorteije and Richard van Wezel.

As a second research internship, he spent three months at the Neuroscience department of the Erasmus University Medical Center in Rotterdam to get acquainted with the neurophysiological and computational principles of cerebellar vestibular processing under the supervision of Chris de Zeeuw. He subsequently went to the Neurological Sciences Institute of Oregon Health Sciences University in Beaverton, Oregon, USA for six months. Here, he performed single unit recordings from cerebellar Purkinje cells and interneurons under the supervision of Neal Barmack and Vadim Yakhnitsa. He also learned to perform microsurgical procedures in mice.

Upon his return to the Netherlands, he wrote his MSc-thesis *The Neurobiology of Visual Consciousness*, which was shortlisted for the Vliegenthart thesis award. He graduated from his master's program with a *Cum Laude* distinction in 2006 and started working as a PhD-student in the department of Functional Neurobiology at Utrecht University shortly thereafter. Here he joined a research project funded by a grant from the Utrecht University High Potential Program to study the neural mechanisms that underlie conscious visual perception supervised by Richard van Wezel and Bert van den Berg. In his studies, he used human psychophysics, computational modeling approaches, and electrophysiological recordings in awake behaving monkeys. The results of these studies form the basis of this thesis.

During his time as a PhD-student, he attended Summer schools in France and Germany, visited conferences and colleagues all over Europe and the USA, and collaborated with Raymond van Ee, Martin Lankheet, André Noest, Jan Brascamp and Randolph Blake. He also participated in the Graduate School of Life Sciences and was a member of the Helmholtz Institute's PhD-students Council.

In April 2010, he joined the department of Psychopharmacology at Utrecht University where he will work as a postdoctoral fellow with Richard van Wezel in setting up Utrecht University's first two-photon neuroimaging research group. Here, he aims to study the neural mechanisms of learning, memory and perception at the neural network level, both *in vitro* and *in vivo*, while possibly interacting with these processes using pharmacological agents.

