On the visual system's architecture underlying binocular rivalry and motion perception.
Does it take two to tango?

Jeroen J.A. van Boxtel

On the visual system's architecture underlying binocular rivalry and
motion perception. Does it take two to tango?

Thesis, Faculty of Physics, Utrecht University, The Netherlands

Cover design: Jeroen Johan Adrien van Boxtel

Electronically published by the University Library, Utrecht University,
the Netherlands

ISBN: 978-90-393-4731-7

Printing: Ponsen & Looijen b.v., Wageningen.

On the visual system's architecture underlying binocular rivalry and motion perception.
Does it take two to tango?

Over de architectuur van het visuele systeem aan de basis
van binoculaire rivaliteit en bewegingszien.
Does it take two to tango?
(met een samenvatting in het Nederlands)

Tatanan arsitektur dari sistim visual yang mendasari rivalitas antar
dua mata dan persepsi terhadap suatu gerakan.
Perlukah keduanya untuk berdansa tanggo?
(dengan ringkasan dalam bahasa Indonesia)

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 16 januari 2008 des middags te 4.15 uur

door

Jeroen Johan Adrien van Boxtel
geboren op 25 april 1979
te Gouda

Promotor: Prof. dr. C.J. Erkelens
Co-promotor: Dr. R. van Ee

Het printen van dit proefschrift werd mede mogelijk gemaakt met financiële steun van de J.E. Jurriaanse Stichting

Contents

Chapter I

Introduction

I.1 Short overview of the architecture of the visual system

Vision comes naturally to us. However, the seeming effortlessness belies the very complex underpinnings of our everyday perceptions. With an estimated 100.000.000.000 or so neurons [218] and many many more connections between these neurons, the brain provides a phenomenal challenge for vision scientists trying to unravel the exact neural interactions that build up our perceptions. Nevertheless, in this thesis I will try to convince the reader that with some relatively crude measures (such as subjective reports of perception), we may get a glimpse of the architecture of the neural structures and their interactions that subserve our visual impressions.

The sheer complexity of vision is exemplified by the retina, the first, and arguably most simple stage in visual perception. Here, the schema of two receptor types building a scaffold for later cortical processing—rods for colorless high-sensitivity vision, and cones of colored lower-sensitivity vision—is clearly a large simplification. In fact, in the retina over 50 different cell types are currently known, with several more to be discovered [177], and processes such as motion detection [31, 32, 233] and contrast normalization [242] are inaugurated here.

A large part of the computational power of the brain, including the retina, is not due to successive operations of many individually-operating neurons, but instead mediated by the massive parallel processing and cross-interactions between many neurons. These cross-interactions are greatly facilitated by the brain's anatomical architecture. The architecture is such that neurons performing similar tasks are positioned near to each other [62, 78]. This organization allows the information gathered and sent by these neurons to be readily available for mutual cooperation, easing the computational load per neuron. Such grouping of neurons seems to be a general principle in the brain as it takes place at many levels, e.g. neurons with similar orientation sensitivities are located next to each other within the primary visual cortex [120, 121, 102], and neurons particularly sensitive to, for instance, faces are located in a specific brain region [128, 256].

Based on histological findings, the brain can be divided into many different brain regions (see Figure I.1). Within such regions, cells generally perform sim-

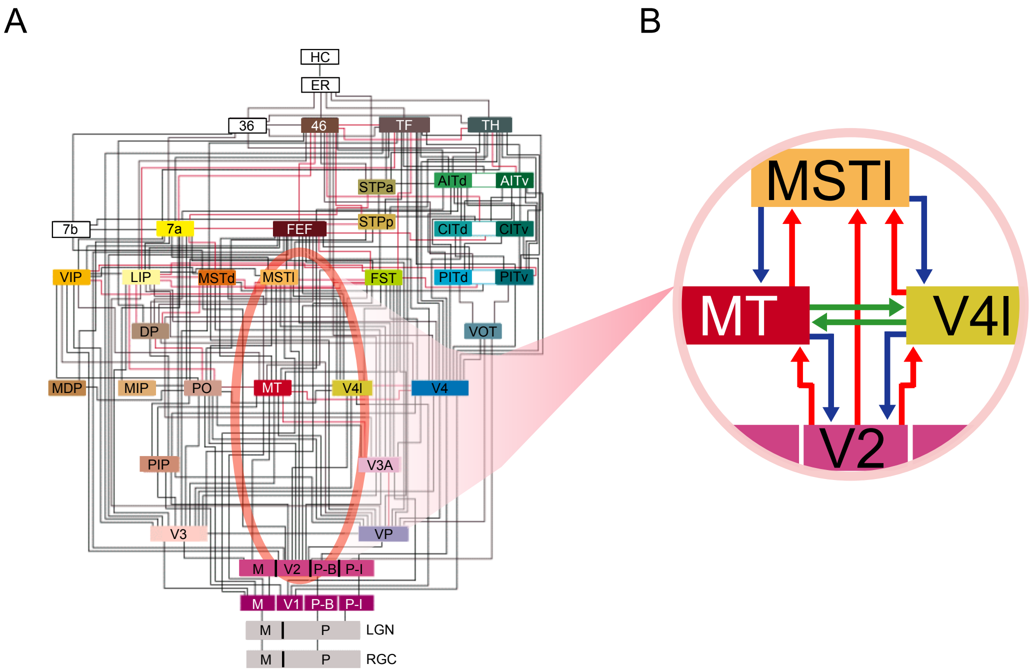


Figure I.1 | **A schematic of the architecture of the visual system** (A) *Different brain areas involved in the processing of visual information, and their inter-connections (adapted from Ref. [79]). The general hierarchical view is that visual information enters the visual system at the retina (RGC = retinal ganglion cells), before it is conveyed to the primary visual cortex (V1, with its subdivisions into M, P-Blob, and P-Interblob cells), after which signals may be sent to several higher visual areas (e.g., V2, the area behind V1, and the Medial Temporal (MT) area, which is sensitive to motion). Areas thought to perform processing of similar complexity are displayed in horizontal alignment.* (B) *Zoomed-in view of a part of the visual system. The connections between the different areas have been characterized as feedforward (red), feedback (blue), and lateral (green) connections.*

ilar tasks. Some of the brain areas involved in visual perception are portrayed in Figure I.1A. None of the regions operates on its own, and communication between and within these areas is of key importance to the formation of our perceptions. Figure I.1A shows some of the interactions as well, and as one can see, the architecture is quite complex. The interactions between brain regions (or within brain regions) may be divided into 3 classes (see Figure I.1B): feedforward connections, feedback connections and lateral connections. Feedforward connections transmit information from an early visual level (low in the computational architecture) to higher visual areas (red lines), feedback connections transmit information from higher visual areas to lower visual areas (blue lines), and lateral connections transmit information from one area to another area at the same level (green lines).

This presentation of the visual system, based on anatomical data, reveals a complex set of possible interactions between brain areas. The actual complexity is many times greater, however. This is because some connections are likely to be very important for many visual tasks, whereas others are for certain specific tasks only. Moreover, some interactions may be essential under some conditions and inessential under other conditions. In fact, they may even be deleterious for task performance under certain conditions. Such deleterious effects may lead to below-chance performance, or visual illusions, providing a wealth of information on the underlying visual processing.

From Figure I.1 one cannot derive the functionalities of the interconnections between brain regions. The functionality of the connections may be shown through neurophysiological studies and may occasionally be derived from psychophysical studies. In this thesis, I will present data obtained using a psychophysical approach supplemented with data from the neurophysiological literature, to obtain insights into the architecture of the visual system underlying motion perception and binocular/perceptual conflict resolution. The first two research chapters of this thesis are devoted to motion perception, while the next four chapters discuss visual conflict resolution. In the next sections, I will introduce both fields and some of their current research directions.

I.2 Motion processing

Motion extraction and initial processing

As mentioned before, some form of motion processing may already take place in the retina [32, 233]. It seems, however, that in primates, and probably humans, motion perception primarily has its roots in the primary visual cortex, and is refined at various other visual levels [48].

Generally, one believes that, in humans motion, is extracted from the flux of visual information by many different constellations of cells that are tuned to specific spatio-temporal correlations. In short, motion is a displacement in space during a certain time. One constellation of cells could be tuned to, say, a displacement of 0.1 mm (on the retina) in a time-period of, say, 30 ms. This constellation would be sensitive to a motion of 0.33 cm/s (on the retina), which is about 18 degrees of visual angle per second. Many such constellations, tuned to slightly different correlations, could cover the entire range of perceivable motions, and as such, support our perception of motion. How the brain performs these local correlations is still not completely clear, even though several computational models exist [221, 226, 1].

Extraction of motion signals alone is not sufficient for motion perception. The brain needs to know whether the speed is constant, or accelerating, or perhaps decelerating, and thus needs to retain past information to predict the future. Another question that needs to be tackled by the visual system is: is it the whole field of view that moves, or just a small patch; in other words, is the motion global or local? In order to answer these questions, the visual system needs to move beyond the extraction of local motion information, and to process the motion globally to create an integrative view of the moving world around it, which will involve several additional processing steps.

In this thesis, I will focus on one of these additional processing steps that lie after to stage of rudimentary motion-signal extraction. Specifically, I will try to unravel the possible neural architecture that underlies the formation of speed perception.

The motion aftereffect

A simple but powerful, yet not intuitively obvious, illustration of the auxiliary processing that is performed by the brain after motion signals are extracted, is the motion aftereffect (MAE; Figure I.2). The MAE is evident to the observer as an illusory directional motion in a stationary stimulus after prolonged adaptation to a motion in the opposite direction. Normally, when a static pattern is presented to the visual system, motion sensitive cells are activated (Figure I.2, pre-test, for simplicity only up- and down-ward selective cells are shown). However, this activation is not very outspoken—as indicated by the few vertical lines that represent spikes of activity of the neurons—,because neither up and down cells are tuned to stationary stimuli. When the pattern starts moving downward (Figure I.2, adaptation), the response of the 'down' cells increases dramatically, but over time will decrease slightly because of adaptation. Upon restimulation with a static pattern (Figure I.2, test) the 'down' cells are still adapted and will hardly be activated, while the 'up' cells fire at a baseline level. Because of this imbalance in activity,

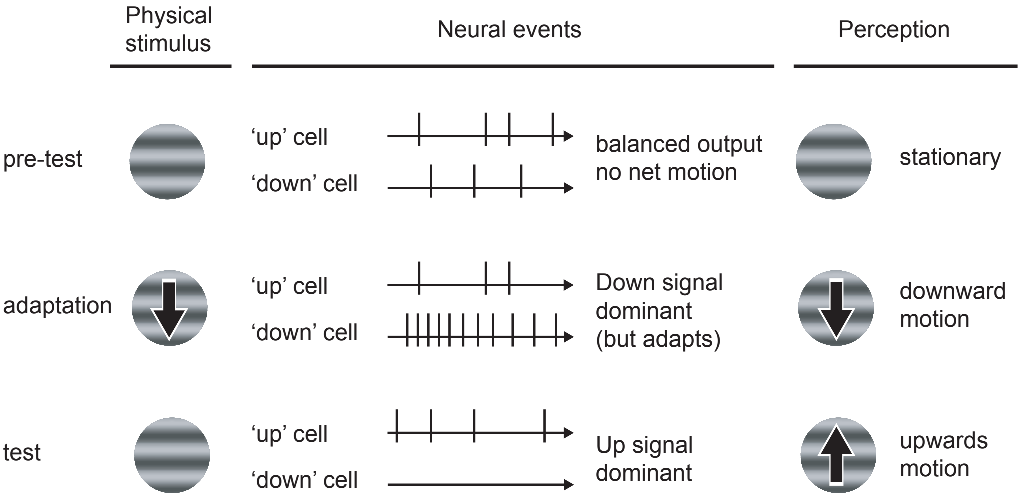


Figure I.2 | An explanation of the motion aftereffect (MAE) (Pre-test). *When confronted with a stationary stimulus, cells that are sensitive to up or down motions respond slightly.* (Adaptation) *When the pattern is moving downward, 'down' cells start firing vividly, and continuous to do so until the motion is stopped. Nevertheless, the activity decreases over time, due to adaptation.* (Test) *When the pattern is stopped, both cells will fire at a much reduced rate, but the 'down' cells more so than the 'up' cells' because they are adapted (and therefore less sensitive). Because the 'up' signal is dominant, the pattern is perceived to move upwards. [Adapted from Ref. [181]]*

the pattern will be perceived to move upwards [30]. The MAE therefore shows that apart from extracting motion signals, the visual system compares extracted signals in a subsequent process to generate a locally congruent motion percept [193, 219, 112].

An additional feature of the MAE is that, apart from the apparent (illusory) motion, one will often notice that the pattern's location nor its features change location. This interesting phenomenon shows that position and motion information are partially uncoupled in the visual system. Furthermore, this characteristic of the MAE makes it an interesting candidate to study the motion system *sensu stricto*, because of the possibility to exclude position tracking mechanisms, which otherwise may also give a sense of motion.

Local motion

Motion consists of two components: direction, and speed. It is well established that the primary visual cortex (V1) contains direction selective cells [120, 121, 102], which are responsive to motion in only a relatively small region of visual space (i.e. they have a local receptive field). It is thought that these cells underlie the perception of the direction of motion. However, the representation of speed is more contentious. The MAE is a useful tool to investigate the representation of speed, not least because of the possibility to silence position tracking mechanisms as potential confounding factors.

MAE studies have taught us that motion adaptation is specific to the direction of adaptation [178, 180, 231] and in some cases to the speed of the adapting motion or test stimulus [67, 22]. However, in other conditions the MAE is not tuned to speed [291, 211]. As it turns out, when a MAE is measured on a static test stimulus (as shown in Figure I.2), it is tuned to temporal frequencies [291, 211]. However, when it is measured on a dynamic (flickering or moving) test stimulus, it is speed tuned [22]. Additionally, MAEs of slow and fast speeds can occur simultaneously on a single test pattern, while two slow or two fast MAE will integrate [236, 263]. These differences have lead to the hypothesis that there are at least two speed tuned motion systems, one tuned to slow speeds (with MAEs elicited on static displays), and one tuned to high speeds (with MAEs elicited on dynamic displays).

Other work on motion perception, not always using MAEs, has also suggested that two or more motion systems exist. Distintions have been made between fast and slow motion systems [29, 57, 295, 75, 90, 91, 97, 107, 130, 101, 236, 261], long and short range motion systems [50], attentional and automatic motion systems [60], first and second order motion systems [202, 165], form and non-form based motion systems [212]. It seems likely that some of these proposed divisions are actually one and the same division, but probed with different stimuli (and consequenty, labeled differently). Furthermore, the conclusions have generally been drawn on the basis of a limited number of test conditions (most often two), and it is therefore possible that some of the putative distinctions may prove to be illusory when probed more extensively. It may be that our tendency to use a dichotomized nomenclature, which does not allow intermediate denominations, forces us to assume the existence of independent systems when we find a dichotomous data-set. However, instead of having measured two largely independent systems, activated by one or the other stimulus, one may actually have measured two different spots in a continuum, which may not be best described by for instance 'static/dynamic', but instead by, say, 'speed'.

In chapter II, I focus on the distinction between fast and slow motion systems, and I measure the dependence of the MAE on different adapting and test stimulus characteristics related to speed. In particular, the research was aimed at investigating whether the current data necessarily imply the existence of two independent motion systems, or that a single system with a broad and continuous sensitivity to speed may suffice. For motion direction perception during the MAE, it has been proposed that opposite direction-analysis is not necessary. Instead, a wide and near continuous range of direction selective units is thought to be involved [245, 178, 180, 285, 286]. In similar vein, I will argue that a single continuous mechanism can be used in the speed domain to explain many findings related to the putative fast and slow motion system dichotomy.

Global motion

The previously discussed processes concern "local motion processing": How is a motion that is extracted from the visual information stream represented at a local site? A second question is, how is this information combined into a single motion percept? One of the many intermediate steps in that process is "global motion" processing. Global motion processing refers to the integration of many, possibly conflicting, motion signals into a consistent motion percept at one patch

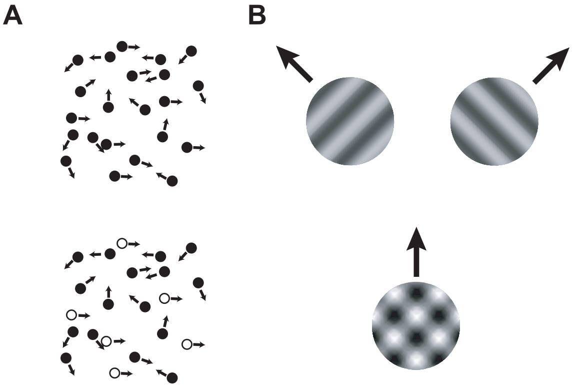


Figure I.3 | **Global motion extraction** (A). Top panel. *A stimulus with many dots seemingly all moving in random directions.* Bottom panel. *Nevertheless, the visual system quite easily extracts the dots that move in the same direction (marked white), and one perceives a rigid sheet of dots moving left, submerged in noise dots with random motion directions.* (B) Top panels. *A grating moving diagonally to the top-left, and another grating moving to the top-right.* Bottom panel. *The sum of the two gratings in the top panel results in a grid-like pattern that moves upwards.*

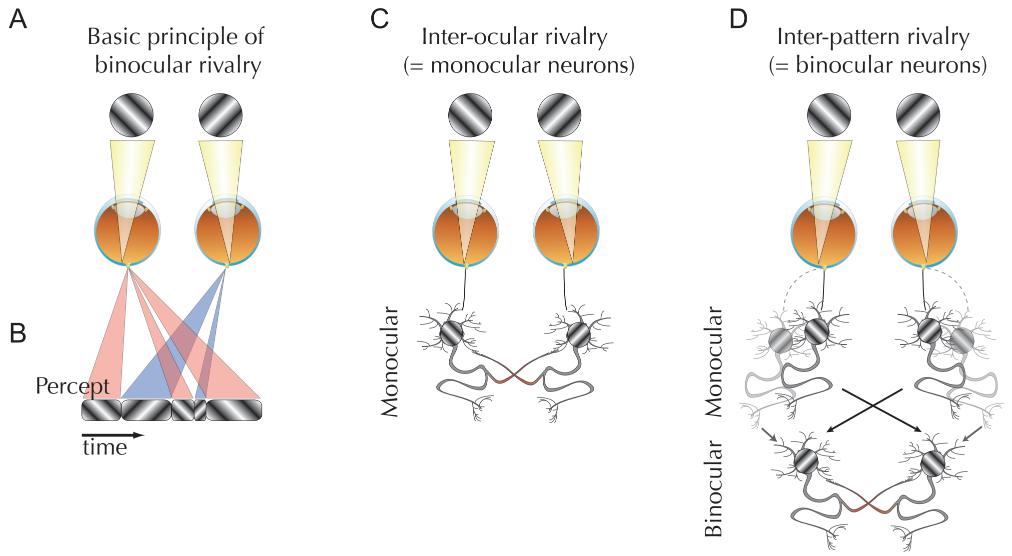


Figure I.4 | **Binocular rivalry** (A).*Binocular rivalry is induced by placing incompatible images in front of the two eyes.* (B) *the resultant percept from the configuration decribed in A is a random fluctuation of the percept between the two images.* (C) *The monocular (low-level) account of binocular rivalry.* (D) *The pattern-based (high-level) account of binocular rivalry*

of the visual field. It is generally accepted that local motion processing precedes global motion processing, both functionally and anatomically [193]. Local motion processing is thought to take place in V1 while global motion processing is often attributed to motion area V5/MT [193]. This again exemplifies the nature of the brain architecture in which cells with similar tasks cluster together.

Figure I.3A explains schematically what global motion processing entails. The top panel depicts a field of moving dots, with the dots moving in a seemingly random manner. If we had to guess a predominant direction of motion in this display by inspecting it point by point, we would be at a loss. The visual motion system extracts this predominant direction quite effortlessly [219], and groups the elements with a similar motion to form a single “object”. This is the result of “global motion” processing (Figure I.3A, bottom panel). Figure I.3B shows an other example of global motion processing. When presented with a diagonally-orientated grating (a blurred line pattern), moving to the top left (Figure I.3B), we will perceive motion to the top left. When presented with a grating moving to the right, we perceive it moving to the top right. Now, when presented with these patterns simultaneously (Figure I.3B, bottom panel), we do *not* perceive two gratings moving transparently through each other. Instead, we perceive a single grid pattern moving upwards (bottom panel) [2, 193]. Again, this is termed “global motion” processing (although the involved processes may actually differ from those involved in the previous example).

Interestingly, in global motion perception, as in local-motion perception, the idea that a slow and a fast speed system exist, has been developed. The existence of these systems has been proposed based on detection-in-noise experiments, and not on MAE experiments (although global-motion aftereffects do exist [238]). Briefly, in the detection-in-noise experiments it was found that low noise speeds (black dots in Figure I.3A) decrease the detection performance of low signal speeds (white dots in Figure I.3B), and, similarly, high noise speeds do so for high signal speeds. However, low noise speeds do not decrease detection performance of high speed signals, and high noise speeds do not decrease detection performance of low speed signals [75, 130]. These data prompted the conclusion that independent high and low speed systems exist in the visual motion system.

Since the above-mentioned studies used only two signal speeds, the possibility of a continuous speed distribution, instead of a bipartite system, could not be investigated. In chapter III this possibility is investigated, and I conclude that, most likely, speed is represented in a continuous manner in the visual system.

I.3 Visual rivalry

In the second part of this thesis we look into the visual architecture underlying the phenomenon of visual rivalry. Visual rivalry results when the eyes view stimuli that have multiple interpretations that are about equally valid solutions to the visual system. Rivalry is manifest as a continuing stream of perceptual alternations of 2 (or more) interpretations. In this thesis, I will focus on binocular rivalry. Binocular rivalry occurs when two incompatible images are viewed by the two eyes (Figure I.4A). The brain cannot combine the two images and chooses to let one image at the time to enter our awareness (schematically drawn in Figure I.4B).

The complete explanation of this arresting phenomenon has evaded us up to this moment. Throughout time, the debate has centered around the question whether binocular rivalry is the result of conflict resolution at a low or at a high level within the visual hierarchy (Refs. [114, 113, 34, 158, 269], and Figure I.4C and D). Put differently, the question has been: is binocular rivalry the result of monocular neurons detecting an interocular incongruency in the visual input which they solve among themselves, or is binocular rivalry primarily dependent on a conflict resolution process within and between higher-level neurons, which try to resolve the ambiguities based on pattern information or even image meaning.

Up until the 1960's, both these views on binocular rivalry were about equally dominant. Since the late 1960's the view of binocular rivalry as a low-level process has been predominant, largely due to the work of Levelt [153], even though importance of pattern information—suggesting the involvement of more central (i.e. cortical) processes—was also emphasized at that time [269]. The low-level view of binocular rivalry culminated in an especially influential review [34], in which it was hypothesized that binocular rivalry resulted from inhibitory interactions between monocular neurons (Figure I.4C). This view is based on several well established findings showing that the patch of eye containing the perceptually suppressed (i.e., invisible) pattern has a reduced sensitivity to all kinds of probe stimuli, irrespective of whether they match the rivalling patterns [82, 83, 198, 34]. On the basis of these findings it has been proposed that perceptual suppression is not dependent on the pattern of probe, but on its eye of origin (e.g. Ref. [34]). Furthermore, is has been reported that suppression-duration depends on low-level visual characteristics, such as contrast and luminance [153, 280, 129].

Since the late 1990's, influences of pattern information (which is eye-independent) on binocular rivalry have been discovered [138, 159, 205] and early work on this topic was rediscovered [10, 72]. These findings have reinstantiated an interest in pattern-based explanations of binocular rivalry. Some pattern-based explanations of rivalry are based on the assumption of eye-based conflict resolution, and the subsequent strengthening of perceptual suppression by pattern-based processing [8, 38]. In other explanations rivalry is based on global reinterpretations of the visual information initiated in brain areas positioned at the end of the visual processing stream, which switch between almost-fully formed perceptions [150, 167, 168, 243, 223]. The pattern-based (or binocular) view of binocular rivalry is clearly quite diverse, but in its most rudimentary form it argues that binocular rivalry is not primarily due to interactions between monocular neurons (see Figure I.4D).

Why study binocular rivarly? The study of binocular rivalry is interesting in its own right, because of the stunning perceptual consequences of rivalry. Alternatively, binocular rivalry may be used as a tool to investigate the formation of visual awareness [64]. Additionally, the rather conflicting theories of rivalry also provide a scaffold for research into the architecture of the visual system. Investigations of binocular rivalry, and especially those that investigate which conditions favor low-level versus high-level rivalry, may advance our knowledge of a part of the visual system that underlies our daily perceptions.

Research into binocular rivalry has a long history, and many aspects of it have been extensively studied, reviewed in [5, 38, 158, 251, 35]. However, one aspect of binocular rivalry has so far received relatively little attention: its dependence on temporal properties of stimulation. Some phenomena related to binocular rivalry, such as dichoptic masking[1] [188, 56] and flash suppression[2] [289], show a strong dependence on temporal parameters of stimulation, but these studies are only concerned with relatively brief stimulus sequences. Binocular rivalry, however, is a long-lasting binocular conflict, and may be impacted differently by temporal parameters.

In this thesis, I will present data from experiments using manipulations of temporal parameters of binocular rivalry stimuli. The approach of O'Shea and Crassini [209] was employed to investigate the occurence of binocular rivalry with stimuli that were shown only intermittently. From their study it was known that binocular rivalry ensues for intermittently shown stimuli even when the two eyes' images are temporally non-overlapping. This stimulation protocol yielded relatively normal rivalry when the stimuli were repeated within an $\sim$350 ms time window. Rather than being an epiphenomenon of the visual system, the contin-

---

[1] Dichoptic masking: brief ($<$ 500 ms) and conflicting stimuli are presented to the two eyes. One investigates whether the visibility of one of them depends on presentation-parameters of the other stimulus.

[2] Flash suppression: A binocular rivalry stimulus is preceded by a stimulus that causes one of the binocular rivalry stimuli to be adapted upon rivalry initiation. One investigates the dependence of the bias that is caused by the adaptation on the similarity between the adaptor and one of the rivalry stimuli.

ued occurrence of binocular rivalry in the absence of direct spatial and temporal conflict may provide essential insights into how, and at what level within the visual system, binocular rivalry is maintained. Up until now, the 1984 study of O'Shea and Crassini [209] has not been followed up by any study aimed specifically at studying this phenomenon. Chapter IV reports on a study in which I specify more accurately what determines the ability of continued rivalry in the absence of direct spatial and temporal conflict, and I speculate that a form-based parsing mechanism is involved in the process.

According to the eye-based account of binocular rivalry, binocular rivalry is placed in a special position, as it depends on conflict resolution between the eyes, whereas other forms of rivalry, such as monocular rivalry [54] and stimulus rivalry [159, 138], do not. In Chapter V it is reported that the form-based parsing mechanism described in Chapter IV also applies to other presentation protocols such as monocular rivalry and stimulus rivalry. Apart from supporting the view that the observed 350-ms limit is form-based and not eye-based, this finding also supports the idea that rivalry in general, including binocular rivalry, is dependent on a pattern-based conflict resolution process.

One of the experiments in Chapter IV showed that when a pair of stimuli is repeatedly presented in rapid sequence, normal rivalry ensues up to repetition periods of $\sim$350 ms. At slightly longer repetition periods, the percept was dominated by the first-presented stimulus (dichoptic masking). The near identity of the conditions leading to binocular rivalry and dichoptic masking allowed us to study the relationship between dichoptic masking and binocular rivalry, without the confound of major stimulus changes. It was found (Chapter VI) that a continuous transition occurs between binocular rivalry and dichoptic masking. At all times, dichoptic masking shows the same perceptual dynamics as binocular rivalry, suggesting a strong link between these phenomena.

As mentioned before, the visual system consists of many stages, and these stages communicate with each other. In binocular rivalry research, it has been suggested that feedback plays an important role in the structuring of visual perception during rivalry [3, 38, 7]. Chapter VII reports on the involvement of feedback information in binocular rivalry when rivalry involves conflict between two complex stimuli (i.e., house and face) and or between more simple grating stimuli. Adaptation to grating stimuli was found to affect the course of rivalry only for rivalrous stimuli that are matched to the adaptation stimulus in retinal position. For the more complex stimuli, adaptation also influenced the course of rivalry on positions matched in space (and therefore independent of the retina). This effect could be the result from a biased rivalry between populations encoding (elements of) house and face stimuli, or it could be the result from an attenuated feedback from those areas to low-level visual areas. The role of feedback in rivalry could be substantiated by the finding that adaptation to complex stimuli also influenced rivalry in spatial coordinates when rivalry took place between a grating and a face stimulus. This combination of rivalrous stimuli is thought to preclude rivalry at higher stages in the visual system, and therefore specifically bares the influences of feedback information.

I.4 Concluding remarks

In Chapter VIII, I will provide an overview of the different chapters and what they together may tell us about the architecture of the visual system underlying our motion perception and our perception of binocular rivalry, and possibly of our visual perceptual system in general.

Chapter II

A single system explains human speed perception

Jeroen J. A. van Boxtel
Raymond van Ee
Casper J. Erkelens

Motion is fully described by a direction and a speed. The processing of direction information by the visual system has been extensively studied. Much less is known, however, about the processing of speed. Although it is generally accepted that the direction of motion is processed by a single motion system, no such consensus exists for speed. Psychophysical data from humans suggest two separate systems processing luminance-based fast and slow speeds, whereas neurophysiological recordings in monkeys generally show a continuous speed representation, hinting at a single system. Although the neurophysiological findings hint at a single system, they remain inconclusive as only a limited amount of cells can be measured per study and possibly the putative different motion systems are anatomically separate. In three psychophysical motion adaptation experiments, we show that predictions based on the two-motion system hypothesis are not met. In stead, concurrent modeling showed that both here-presented and previous data are consistent with a single system subserving human speed perception. These findings have important implications for computational models of motion processing and the low-level organization of the process.

II.1 Introduction

The visual motion system has proven to be extremely versatile. This versatility has repeatedly been explained in terms of two (or more) separate and independent motion systems [29, 50, 57, 60, 75, 90, 91, 97, 107, 130, 202, 212, 101, 236, 261]. However, it is unlikely that all these motion systems exist, as this would lead to a very high number of motion systems (not to mention, all the possible interactions between those systems). We specifically investigated the widely supported division along the speed dimension: i.e. the claim that human perception of motion is subserved by a fast- and a slow-motion system[‡][57, 75, 90, 91, 97, 107, 130, 101, 236, 261]. The evidence for the proposed division in fast- and slow-motion systems comes from psychophysical and clinical research, which we briefly review.

Motion detection and discrimination thresholds for luminance-based (achromatic) and color-based (chromatic) motion are found to be identical for high-speed stimuli. However, it was found that at lower stimulus speeds, thresholds were higher for chromatic than for achromatic stimuli [57, 91]. These results were interpreted as evidence for three different motion systems: one system for slow achromatic motion, one for slow chromatic motion, and a combined fast motion system [57, 90, 91, 107].

Stimulus adaptation has proven to be powerful in revealing neural mechanisms underlying visual perception. Motion adaptation—caused by prolonged viewing of unidirectional motion—is demonstrated by illusory motion in a subsequently viewed motionless (or motion-balanced) pattern (called test pattern): the motion aftereffect (MAE). Psychophysical MAE reports [101, 236, 261] have shown that different MAEs are induced on different test stimuli. More specifically, for speed it has been reported that fast motion induces MAEs with rapidly refreshed test patterns (called dynamic), whereas slow motion does so with slowly refreshed patterns (called static). A strict separation between these two types of MAE has been reported in the temporal frequency domain [236]: at test stimulus refresh frequencies below 20 Hz a MAE of slow adapting motion was observed, whereas at refresh frequencies above 20 Hz a MAE of fast adapting motion was observed. Both these kinds of MAE were seen simultaneously when adaptation occurred to both slow and fast motion, and the test stimulus consisted of slowly and rapidly refreshed stimulus elements (dots), suggesting a complete independence of the slow and fast motion systems [236].

Some clinical studies are taken as evidence for a division in fast and slow motion systems. Patients with lesions in V1 are unable to consciously perceive speeds below 6 deg/s [29], whereas patients with lesions in MT are unable to perceive speeds above 6 deg/s [295], indicating that the slow and fast motion

---

[‡] We use the term motion system to refer to processes in the brain that give rise to motion percepts. Our use of motion systems is close to that of motion channels, in that motion channels perform identical calculations but with different parameters, leading to different sensitivities etc. However, it is of interest to note that authors often imply that the two motion systems respond to different kinds of stimuli. It is assumed that the slow-motion system responds specifically to static test stimuli, and the fast-motion system specifically to dynamic ones (e.g. Ref. [236]). Or that one activates different systems with colored (isoluminant) and luminance-defined motion stimuli (e.g. Ref. [107]), on the bases of the finding that both have different contrast-dependencies. A motion system may contain several motion channels.

systems could be located in different brain areas.

Contrary to the findings in psychophysical research, neurophysiological work in monkeys has so far not suggested different motion systems for fast and slow speeds. The changes in speed preference are generally found to be continuous when measuring neurons in each other's vicinity [156, 182], although occasional large jumps in speed preference occur. Although the neurophysiological findings hint at a single motion system, they remain inconclusive as only a limited amount of cells can be measured per study and, moreover, the putative different motion systems could be anatomically separate. Because of the nonconcordance of the neurophysiological and psychophysical studies the data are at present inconclusive.

Although the psychophysical data has been interpreted as showing the existence for two motion systems, the different studies propose different underlying mechanisms. The slow motion signals are thought to be carried by sustained firing cells [101, 236] or parvocellular cells [57, 90, 91, 107]; the fast motion signal are thought to be carried by transiently firing cells [101, 236], or magnocellular cells [57, 90, 91, 107]. At a level lying after the slow and fast motion systems, the motion signals are thought to converge again to form a unified perceptual experience [57] or, instead, they are thought to have "a private line to consciousness" [101].

We note that the classifications like static vs. dynamic, chromatic vs. achromatic and also first-order vs. second-order, depend on stimulus characteristics that may not be the characteristics the visual system uses. The dichotomized nomenclature, which does not allow intermediate conditions, forces one to assume that a finding of a dichotomy in the data using two different stimuli implies that two largely independent systems exist that are mainly activated by one or the other stimulus. However, one may actually have measured at two different spots in a continuum that is not best described by e.g. "static/dynamic", but instead by, say, "speed".

By employing psychophysical methods, corroborated by a mathematical model based on neurophysiological principles, we approached the question from both psychophysical and neurophysiological sides. In the first part of this report we show that the single-system model explains the experimental data on MAEs without necessitating a division in fast- and slow-motion systems. Instead, a single motion system—by which we mean a continuum of interconnected speed- and direction-tuned cells—suffices.

Model simulations gave the inspiration to execute three experiments, which are discussed in the second part of the report, together with explanatory model stimulations. We used an adaptation procedure to specifically gauge a system that adapts to motion, thereby excluding systems that may give rise to percepts of motion solely by means of position tracking, which are not motion systems sensu stricto (although such mechanisms can nevertheless be very important in motion perception [165]). With these experiments we demonstrate that the hypothesis of two independent motion systems fails both the necessity and the sufficiency criterion. The single motion system hypothesis is sufficient to explain the data (see discussion), which therefore suggests that a single motion system underlies human motion perception.

II.2 Model section: a single system explanation

The model is a simplified functional representation of motion-area MT, in which nearby locations have similar speed and direction tuning, which both change in a continuous manner [156, 182], and in which a motion opponency mechanism exists [112].

In the model we assume that a test stimulus elicits a reaction in the motion system. The reaction is altered by an adaptation-dependent modulation. We investigate if such a model, without assuming any discontinuity in the form of motion systems, or channels, is able to produce the known MAE-data.

Materials and Methods

Single motion system model

In the model we assume that a test stimulus elicits activation of the motion system. The activity is modulated by an adaptation-dependent factor. Responses to the test stimulus are modelled as spontaneous activity (arbitrarily set to unity) plus a test stimulus-induced activity, caused by the matching of different stimulus-elements between subsequent frames. The test stimulus-induced activity is described by a distribution of velocities that are present in the test stimulus, and is represented as a Gaussian blob ($Gb(\mu_{test}, \sigma_{test})$). μ_{test} is the speed that is predominant in the test stimulus, and is set to zero to simulate zero net motion, σ_{test} is the variable test velocity-distribution width (TVD width), which is larger for dynamic test displays than for static ones[§]. The final formula describing the test-induced activation is: $1 + cGb(\mu_{test}, \sigma_{test})$. The constant c is set to 8 to increase the influence of the test stimulus substantially above spontaneous activity (the exact value of k is unimportant for the qualitative results of the model).

The test-induced response is multiplied by a modulation factor defined by the preceding adaptation:

$$1 - \sum_{n=1}^{m}[(G(\mu_n, \sigma_n)_s - G(\mu_n, \sigma_n)_{-s})M'(\beta_n, \kappa_n^e, \kappa_n^i)],$$

where m is the number of adapting motions, and G is a Gaussian distribution, M' is a circular mexican-hat shaped function that we will describe below. The modulation factor is made up of two main parts. The first part, $(G(\mu_n, \sigma_n)_s - G(\mu_n, \sigma_n)_{-s})$, describes the adaptation profile along the speed dimension, which depends on motion opponency. $\mu_n$represents the speed at which the motion system is most adapted by stimulus n, and σ_n represents the spread in the speed dimension. Motion opponency in the human visual system may be

---

[§] We have followed the next reasoning. Suppose we have two frames. Each contains randomly positioned dots, and the frames differ from each other. By presenting the two frames in sequence, the motion system matches all (or many) of the dots of frame one with all (or many) of the dots in frame two. These spatial displacements, together with a certain time lag between the onsets of both frames, create a certain velocity distribution. Choosing the lag small (high refresh frequencies) gives rise to high speeds, choosing the lag to be large, but keeping the spatial displacements the same, creates low speeds. Therefore high refresh frequencies give rise to wider velocity distributions than low refresh frequencies in the stimuli we used.

accomplished by subtracting the activity of motion-sensors that are sensitive to opposite motions. We implemented it therefore as a subtraction of the value in the Gaussian distribution $G(\mu_n, \sigma_n)$ at position $-s$ from that at position s, for all values of s. In the function this is represented as a subtraction of $G(\mu_n, \sigma_n)_{-s}$ (the distribution's tail at negative s) from $G(\mu_n, \sigma_n)_s$ (the tail at positive s) for all $s \geq 0$. If σ were infinitely large, the visual system would be equally adapted to all speeds.

The second part of the modulation factor is $M'(\beta_n, \kappa_n^e, \kappa_n^i)$, which describes the adaptation profile along the direction dimension (d). $M'(\beta_n, \kappa_n^e, \kappa_n^i)$ is a Mexican hat-shaped function: $M(\beta_n, \kappa_n^e) - M(\beta_n, \kappa_n^i)$, where M is the von Mises distribution (here also dependent on κ), β is the mean direction, and κ is the concentration parameter (κ = 0, means a uniform distribution). The von Mises distribution is a circular analog of a normal distribution and is defined on the range of $\beta \in (0, 2\pi\rangle$: $M(\beta, \kappa) = e^{\kappa \cos(d-\beta)}/(2\pi I_0(\kappa))$, where I_0 is the modified Bessel function of the first kind of the order 0, and d is the direction-variable, along which the function varies. The first and second parts of the modulation-factor are multiplied to obtain an adaptation profile along both the speed and direction dimension, which is used to modulate the test-induced response.

We investigated the importance of the parameters κ^e and κ^i by repeating the simulation in figure II.1 with many different combinations of κ^e and κ^i. We found that the results could be obtained when parameter κ^e was smaller than about 0.6π, and κ^i was smaller than about 0.25π, corresponding to a width-at-half-height excitation and inhibition larger than about 90 degrees and 120 degrees, respectively. In all subsequent simulations, we set κ^e to 0.3π (width at half height~100 degrees) and κ^i to 0.1π (width at half height ~160 degrees), which are values close to those used in previous simulations with a model similar to ours [103], and not unlike known physiological data [11, 240].

Due to the later subtraction from 1, the physiological meaning of excitation and inhibition are actually inhibition and disinhibition, respectively [103]. σ and μ have units in deg/s, β in rad. For Figure II.1, II.3, II.4, local maxima were found, starting at random locations within twice the size of σ_{test} from the origin. This effectively models a noisy winner take-all mechanism. Thirty degrees of uniform noise was added to the obtained direction for illustrational purposes, 10 repetitions were executed.

Parameters Figure II.1: $\sigma_{1,2} = 6, \mu_1 = 2$ and $\mu_2 = 20, \beta_1 = -3\pi/4, \beta_2 = -\pi/4$. σ_{test} was varied. Contour plots were scaled separately.

Relation to other models

Our model was designed to answer a sufficiency question: Can the known MAE data be explained in a single-motion system framework? The model also served to answer a necessity question: is a two-motion system framework necessary to explain the known MAE data? To this end, the model was kept devoid of large amounts of detail, and included just a motion opponency stage [112], and a continuous representation of direction and speed [156, 182]. The model is closely related to the multiple-direction opponency model used by Grunewald and Landheet [103], as the interactions along the direction dimension are similar, but in

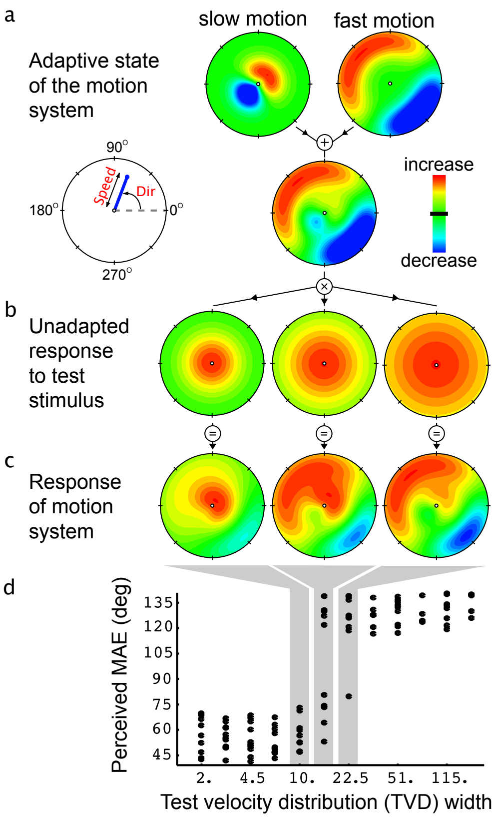


Figure II.1 | Model results showing that a single motion system explains the discontinuity in perceived direction of motion aftereffects (MAE). *The model stages:* (a) *modulation through adaptation, of the slow, fast and the combined components. Velocities are represented by a vector, the length of which is the speed, and the angle is the direction. Shades of red and blue represent increased and decreased activity, respectively.* (b) *response to test stimuli of different test velocity distribution (TVD) width in an unadapted state,* (c) *the final response of the motion system.*(d) *The single-system model reveals a discontinuity in the perceived MAE-directions, shown for different TVD, replicating previously reported data that led to the prevailing conclusion that slow and fast motion systems exist.*

our model the speed variable has been included, and the time dimension was not. The model framework is also similar to that of Wilson and Kim [286], which has a first level of opponent motion interactions and a second level of short-range excitation plus long-range inhibition along the direction dimension, much as in our model. The current model does not explicitly include the interactions among the Fourier components in the stimuli [139], although it does include them implicitly by defining MAEs in a velocity-space [267].

II.3 Model results

Discontinuous MAE data from a continuous model system

The motion system's response after adaptation was modeled by its response to the test stimulus in the non-adapted state multiplied by an adaptation-dependent modulation. The motion system's response to the test stimulus was explicitly modeled, taking into account the velocity content of the test stimulus, as static (low test-refresh frequency; low TRF) and dynamic (high TRF) test stimuli have highly dissimilar velocity contents. A Gaussian blob centered at zero velocity modeled the velocity content of the test stimulus (Figure II.1b). Its width, representing the width of the test stimulus velocity distribution (TVD), is the model's analogue of the experimental parameter TRF. The motion system's response after adaptation was read out by a noisy winner-take-all mechanism, giving perceived MAE directions.

Our single-system model reproduces the previously reported discontinuity in MAE directions (see Ref. [236], and our Figure II.1d). The discontinuity has been taken as evidence for two independent motion systems [101, 236, 261]. Using the model, we may explain the occurrence of the discontinuity as follows: adaptation causes an increased responsiveness at two separate spots (opposite the two adapting velocities [103]) in the direction-speed space. Only one of these spots (at the high speed) is a global peak (Figure II.1a, bottom panel). A ridge, slowly decreasing in height, runs from this peak to the spot opposite the slow adapting speed, after which it rapidly declines to basal levels (Figure II.1a, bottom panel). Without explicitly putting a decreased sensitivity to slow motions in this model, such behavior is observed. The model produces this behavior because the motion-opponency stage cancels out much of the low speed signals since much of the normally distributed adaptation of the velocity-sensitive cells falls on motion cells of opposite motion specificity when low speeds are used, but not when high speeds are used. The resulting larger adaptation at high velocities seems consistent with the literature that observes an ever-increasing responsiveness with increasing speed in motion area MT for the speed-range we used in this study [182]. This causes the following effect: with narrow TVDs, only the lower end of the ridge is read out, and a slow MAE is produced; with increasingly wide TVDs the high-speed peak will eventually absorb the lower peak, thus creating fast MAEs (Figure II.1b and II.1c).

It thus seems that previously reported influential experimental results [236] may be the footprints of a differential read-out of a continuous system, in which a dichotomous response distribution arises because of sufficiently large speed differences between the adapting motion components.

II.4 Psychophysical Tests of Model Predictions

If a mechanism similar to our model is operational in the visual system the following four predictions should be true: Prediction 1: the occurrence of a motion aftereffect after adaptation to slow or fast motion is not strictly dependent on the type of test stimulus. The static test stimulus activates mostly the lower part of the speed spectrum, whereas the dynamic test stimulus also activates large parts of the higher speed spectrum (Fig II.2, second row). The static test will therefore support a MAE of the slower of two adaptation motions, independent of the exact speed (Fig II.2, third row). The dynamic test will support a MAE of the faster of the two adapting motions. This will only hold if the faster motion induces a stronger activation than the slower motion (as shown in figure II.2), if this not the case one may observe two MAEs on a single test stimulus (as is found in experiment 2, Figure II.2). Prediction 2: With the right parameters, three different speed-dependent motion aftereffects can be induced with a single adaptation stimulus. Prediction 3: Using three different refresh frequencies for the test stimulus should produce a maximum-duration MAE at three different speeds. Prediction 4: The discontinuous transition from slow to fast MAEs should exist only for rather large speed differences in the adaptation stimulus. If speed differences are small no transition should be found, or it should be continuous rather than discontinuous. This prediction should hold even if one speed falls within the range of the putative slow motion system, and the other falls in the range of the fast motion system. None of the predictions are made by a dual motion system account for luminance-base motion that supposes independence of the two motion systems.

Prediction 1: The importance of relative speed differences

In experiment 1 we tested prediction 1: the type of MAE is not dependent on the type of test stimulus. We pitted slow (6.4 deg/s), medium (17 deg/s), and fast (28 deg/s) motions against each other in a pair-wise fashion, and tested with static and dynamic test stimuli. According to the prediction 1, and intermediate speed should show a MAE on a static test pattern when a faster motion is simultaneously adapted to, but it should show a MAE on a dynamic pattern when the concurrent adapting motion was a slower one.

Materials and Methods

Experiment Movement was presented at 225 deg, and 315 deg (counted counter clockwise from a horizontal rightwards direction) for slow and fast motions respectively. Used were: 6.4 vs, 17 deg/s, 17 vs. 28 deg/s, and 6.4 vs. 28 deg/s (diamonds in Fig II.3). A similar experiment has been performed before [236], and to be able to compare the data, we repeated the experiment with the speeds reported in that study (circles in Figure II.3): 1.3 deg/s vs 4 deg/s (plotted in Fig. II.3: slow vs medium), 4 deg/s vs 12 deg/s (plotted in Fig. II.3: slow vs fast), and 12 deg/s vs 36 deg/s (plotted in Fig. II.3: medium vs fast). Subjects GB, AK, and TK participated. Each direction was represented by 1,000 white dots (9 arcmin/dot) on a black background. The stimulus size was 18.4 degrees, it was surrounded by a 25-pixel wide empty space, followed by a 50-pixel wide border of randomly placed stationary dots, which were also visible during the test period. The test stimulus consisted of 3,000 randomly placed dots. In static test conditions the dots on the stimulus remain motionless on the screen for the duration of the test. In dynamic test conditions, patterns were refreshed at 33.3 Hz, meaning that every 30 ms all dots were removed and repositioned to random new locations. A test stimulus was presented for 4 seconds, after which the subject indicated the strongest perceived MAE-direction by means of a mouse-controlled dial. If subjects indicated that no MAE was seen, the datum-point was excluded from analysis. A fixation mark (size: 18 arcmin) was present, surrounded by 16-pixel wide empty space. 1 pixel = 2.3 arcmin; monitor frame rate: 100 Hz.

Viewing was binocular. A Macintosh PowerPC G4 drove the experiment. Images were presented on a LaCie electron22blueIV monitor. Subjects AK, TK, and GB were naïve as to the purpose of the experiment; subject JB is one of the authors. Subjects used a chin-rest to stabilize head position. Experimental procedures were reviewed and approved by the Institutional Review Board and subjects gave informed consent.

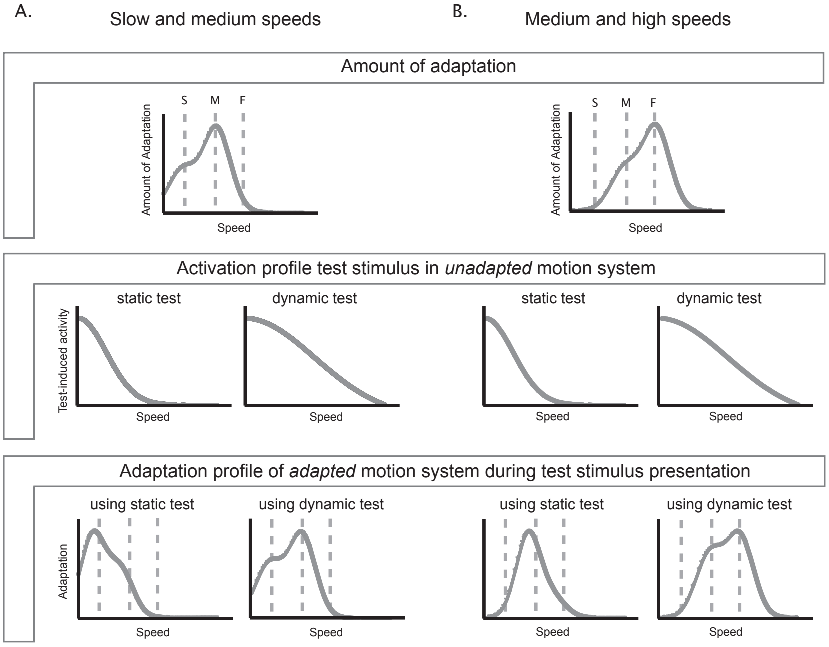


Figure II.2 | Schematic representation of the model predictions in cases of adaptation to slow and medium speeds (panel A), and medium and high speeds (panel B). *First row, the amount of adaptation caused by the stimuli. The model shows less adaptation to slower speeds, both in panel* (A) *and* (B)*. The dashed lines indicate the adaptation speeds for slow (S), medium (M), and fast (F) components, although only slow and medium speeds were presented in panel* (A)*, and medium and fast speeds in panel* (B.) *The second row shows activation caused by static (left panels in* (A) *and* (B)*) and dynamic (right panels in* (A) *and* (B)*) test stimuli. Dynamic patterns show a wider range of activation. The third row shows that final readout of the amount of adaptation by the test stimulus (a multiplication of the first row by the second row). A static test pattern causes a maximum adaptation to be read out at the slower speed of the two adaptation components; a dynamic pattern has a maximum at the higher speed. This pattern of results occurs in both panels* (A) *(slow-medium speed adaptation) and panel* (B) *(medium-high adaptation). The transition can be discontinuous (i.e. a jump) or continuous, depending on model/experiment parameters.*

Parameters used in the simulations Figure II.3. Red: $\sigma_{1,2} = 3$, Left panel: $\mu_1 = 1, \mu_2 = 20$, Middle panel: $\mu_1 = 1, \mu_2 = 9$, Right panel: $\mu_1 = 9, \mu_2 = 20$. Green: $\sigma_{1,2} = 3$, Left panel: $\mu_1 = 4, \mu_2 = 17$, Middle panel : $\mu_1 = 1, \mu_2 = 4$, Right panel: $\mu_1 = 17, \mu_2 = 20$. Blue: $\sigma_{1,2} = 6$, μ_1 and μ_2 as for the green plots. σ_{test} was 4 for static conditions, and 34 for dynamic conditions.

Results Experiment 1

Statistical tests showed that all pair-wise comparisons between static and dynamic conditions were significant (one-tailed t-test, all $p < 10^{-3}$; the results from the different static conditions did not differ from each other, nor did the different

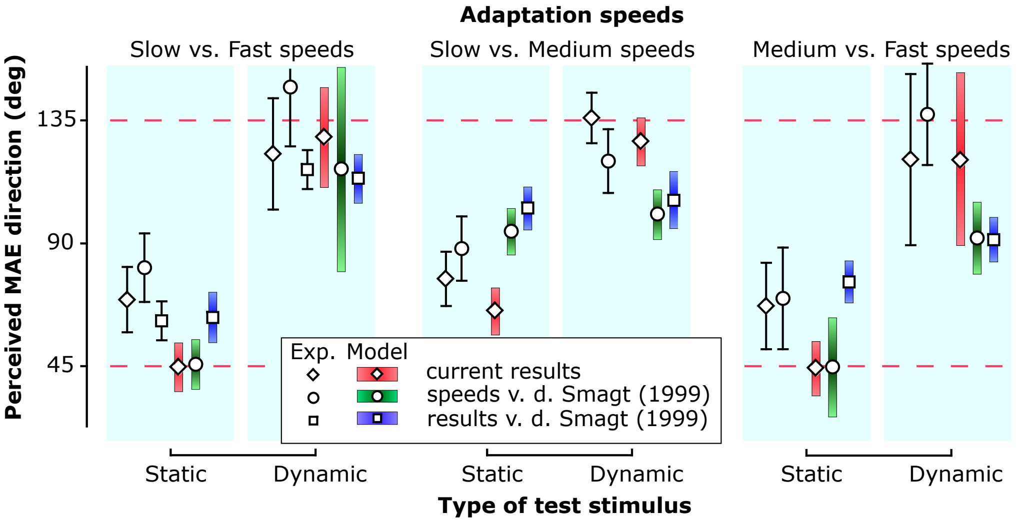


Figure II.3 | Perceived MAE direction (± standard deviation) for static and dynamic test stimuli. *Experimental (symbols only), and model results (having symbols matching the experimental data, and colored rectangular-shaped error bars) demonstrate a discontinuity in the MAE direction for all paired comparisons (slow-fast, slow-medium, and medium-fast speeds; diamonds). A previous influential paper [236] did not obtain such results (squares). However using their parameter settings, we still observed a distinct discontinuity (right panel; circles). The single-system model encompasses all data by changing the adapting speeds (red to green), and the spread of velocity adaptation (σ; green to blue).*

dynamic conditions differ significantly from each other: all $p > 0.1$, two-tailed t-test). Moreover, all means differed significantly from 90 degrees (all $p < 10^{-3}$, z-test). The results (Fig. II.3, diamonds) show that for all pairs the slower of the two motions was linked to a MAE on static test stimuli, whereas the faster was linked to a MAE on dynamic test stimuli, irrespective of their absolute speed (see also Ref. [267]). The medium speed (17 deg/s)—falling well within the proposed range of the fast motion system [261]—induced a MAE on static or on dynamic test stimuli, depending on the accompanying motion. These results speak against the existence of two motion systems. However, adapting speed differences are not the only factor determining the occurrence of the discontinuity in MAE direction, which is made clear by comparing our results with earlier research [236]. Similar speed differences between the adapting motions were used in this report and previous research, but the previous research found no discontinuity when pitting two rather slow (or two fast) but different motions against each other, whereas we did find a discontinuity. This discontinuity did not disappear when we repeated the experiment with speeds identical to those from Ref. [236] (Fig. II.3, right panel, circles; pair wise comparisons between static and dynamic conditions, $p < 10^{-3}$; all means were different from 90 degrees ($p < 10^{-2}$), except the slow vs. medium comparison $p > 0.7$).

Our model reproduces our two sets of data by just changing the adapting speeds (see methods) in similar ways as in the experiments (compare red (diamond) and green (circle) model results with corresponding experimental results, Fig. II.3). Without any further changes, it did not reproduce the results from Ref. [236]. However, the stimuli used in both experiments were rather different: random dot patterns are used in the present study, whereas previous research used random-pixel arrays. (These stimuli look like checkerboards with the checks randomly assigned a white or black color.) Displaced random-pixel arrays (used in Ref. [236]) contain more false matches, and thus noise, between subsequent views than do sparse random dot stimuli. Interestingly, when we introduce this notion into the model by increasing the range of speeds that adapts around the adapting speed (increasing σ in the model) the model reproduces the experimental data of Ref. [236] (blue model results in Fig. II.3).

There are therefore two factors determining whether a discontinuity in the MAE data occurs. First, the distance in speed between the two adapting motions. Second, the range of speeds that adapts in response to prolonged viewing of the adapting motion. In fact, the distance between the adapting speeds needs to be large relative to the width of the region that will be adapted by the stimuli to give rise to a discontinuity in the reported MAE directions with static and dynamic test stimuli, otherwise a continuous transition is observed, or no transition at all.

Prediction 2: MAEs in three directions

In a second experiment we tested prediction 2: adaptation to three motions may produce three MAEs. It is proposed that within the slow and fast systems motion information is averaged after adaptation, but between systems it is not [236]. This view is supported by the finding that adaptation to two motions of similar but not necessarily identical slow (or high) speeds but different direction will result in a MAE opposite to the average motion during adaptation, whereas adaptation to motion two components, of which is a slow and one is a high speed motion, leads to MAE opposite to one or the other component [236] (but see our experiment 1)¶. A prediction of a strict dual motion-system hypothesis is that a maximum of two MAEs can be obtained in response to any combination of adapting motions. However, under the condition that three simultaneously present adapting components are used, our single system model predicted three MAEs opposite the adapting motion components (Fig. II.4a).

Materials and Methods

Experiment Movement was presented simultaneously in following directions: 45 deg, 165 deg and 315 deg, moving at 6.4 deg/s, 17 deg/s, and 28 deg/s, respectively, or at 1 deg/s, 12.7 deg/s, and 28 deg/s. We avoided motion in the four cardinal directions as we found in pilot studies that subjects are biased towards reporting MAEs in these directions. We avoided these directions to give all adapting motion components a more equal chance to be reported. The components were not equally spaced along the 360 degrees as it turned out in pilot

---

¶Note that one need not expect any interaction between slow chromatic and achromatic motions if one assumes different systems to underlie these motion percept

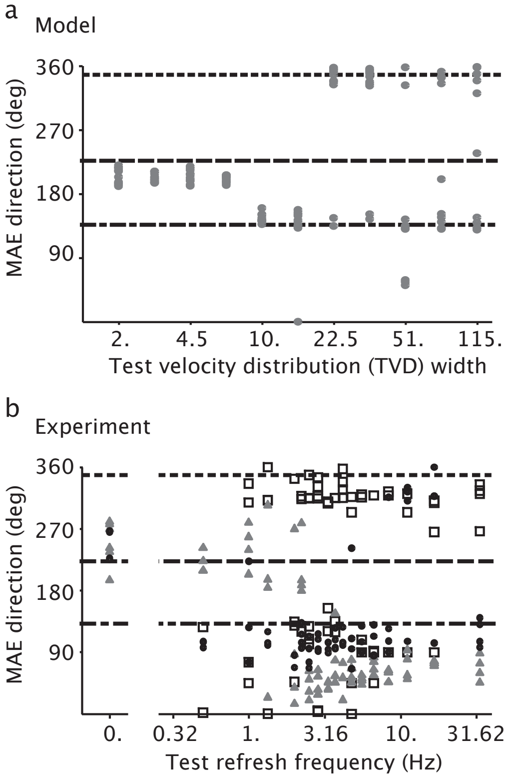


Figure II.4 | Perceived MAE directions depending on test stimulus refresh frequency (or its model analog: the test velocity distribution width) after adaptation to a stimulus containing three motions. *The lines show directions opposite the adapting motion directions. Different subjects are denoted by different symbols. Both in the model* (a) *and in the experiment* (b)*, three MAE directions were produced, contrary to the two MAE directions that would be predicted by a two-system account of the visual motion system. Inter-subject variability causes the somewhat noisy appearance of the experimental data. Along previous lines of reasoning the experimental data suggest the existence of three motion systems, but the model shows that a single motion system is sufficient to explain these experimental data.*

experiments that in such conditions subjects showed a continuous transition from the high-speed to the medium-speed MAE; the used motion directions resulted in discontinuous transitions in most subjects. Test pattern refresh frequencies varied. Other parameters were as in experiment 1.

Model parameters Fig. II.4. $\sigma_{1,2,3} = 3, \mu_1 = 1, \mu_2 = 9$ and $\mu_3 = 20, \beta_1 = \pi/4, \beta_2 = 7\pi/4$, and $\beta_3 = \pi/12$, σ_{test} was varied. $\sigma_{1,2,3}$ and $\mu_{1,2,3}$ are as those in the left panel in Fig. II.3 (experment 1).

Results Experiment 2

We found that three different MAEs—opposite the adapting motion components—were reported (Fig. II.4b), in accordance with the model predictions that were based on parameters used in experiment 1 (Fig. II.4a). The subject represented by the gray symbols does not show three groups of answers, but in stead shows a group of answers around 225 degrees (MAE of slow speed), and a group of answers that follows a continuous path from the intermediate speed MAE (at 345 degrees) to the fast speed MAE (at 135 degrees) as the TRF increases, without reaching either solution completely. Note here that the MAE-direction is a circular variable. These results conform the single-system account (see Prediction 4). Furthermore, subjects indicated that sometimes two different MAEs were seen simultaneously on a single test pattern. Both findings are in agreement with a single-system hypothesis.

Prediction 3: Perceived MAE durations for intermediate test stimulus refresh frequencies

In experiment three we tested prediction 3: intermediate TRFs show their longest MAEs for intermediate adapting speeds. (Predictions for a two-system hypothesis—proposed on the basis of the described results [261]—could not be made, as they depend on the precise mechanisms that one assumes to underlie the division. So far, the proposed mechanisms are not articulated sufficiently precisely.) We tested our prediction experimentally (see Methods) with TRFs of 0, 19, and 75 Hz.

Materials and Methods

Experiment We used a setup and procedures similar to Ref. [261]. Adapting movement was 1 to 29 pix/frame (3 to 58 deg/s). Stimulus size was 10.3×10.3 deg^2 (256×256 pix^2). Test-refresh frequencies (TRF) were 0, 19 or 75 Hz. The 19 Hz TRF was chosen close to the transition point between the putative slow and fast motion systems [236] to maximize the responses of both systems. During the entire experiment a 16 pixel-sized fixation cross was surrounded by a 28-pixel wide square (intermediate grey). The monitor refresh-rate was 75 Hz, 1 pixel was 2.4 arcmin wide. Subjects could indicate that no MAE was seen, this was analyzed as a zero second MAE. Subject GB performed 4 repetitions, AK and JB each 5.

Results Experiment 3

Prediction 3 was borne out by the findings: intermediate TRFs led to distinct peaks in MAE duration at intermediate adapting speeds for all subjects (Fig. II.5). At first glance it may be argued that some curves show a bimodal profile. However, this interpretation is dismissed as for almost all these cases such an interpretation depends critically on a single data point that lies above or below the value expected from a unimodal dataset, which often (but not always) has a rather large s.e.m. as compared to the other data-points. Moreover, if one assumes that two systems cause two peaks, one should also see a small peak for the static curve at about 25 deg/s (especially for observer AK). Such is not seen. Furthermore, if one adheres to the reasoning that many researchers follow when proposing the dual-motion system, Fig. II.4 would suggest that three systems exist (as it shows three different MAE-directions), and thus that Fig. II.5 should show 3 peaks. Also, this is not observed.

To determine precisely the mechanism causing the different positioning of the peaks of the thee curves, the experimental data were fitted with a function based on our model, of which the parameters have physiological interpretations: $c[G(\mu,\sigma)_s - G(\mu,\sigma)_{-s}]G(\mu_{test},\sigma_{test})$, where G is a Gaussian distribution. μ and σ characterize the adaptation state of the visual system, μ represents the speed at which it most strongly adapts, σ represents the spread, and σ_{test} represents the TVD-width, μ_{test} represents the overall speed in the test stimulus, which is set to zero. μ, σ, and σ_{test} are expressed in degrees per second. All parameters

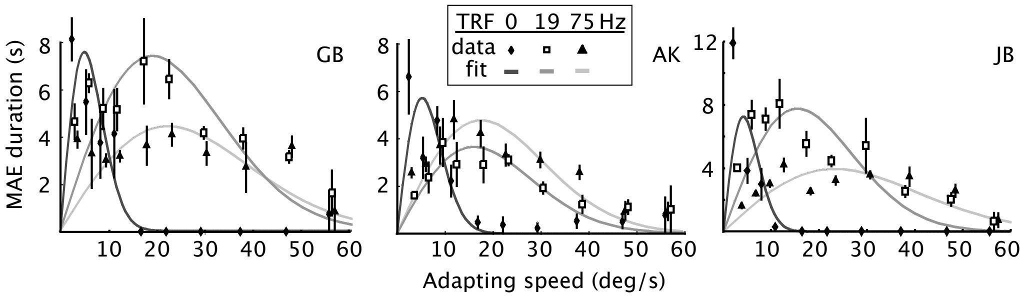


Figure II.5 | Perceived MAE durations (±1 s.e.m.) as a function of adapting speed for three subjects. *Test patterns were static (diamonds), refreshed at 19 Hz (squares), or refreshed at 75 Hz (triangles). Fits through these data sets are shown (from red to yellow: 0, 19, 75 Hz). Previously [261], fast and slow motion systems were proposed on the basis of the finding that static and dynamic test stimuli exhibit the longest MAE at vastly different adapting speeds. The widely accepted division in fast and slow motion systems is at odds with here-presented data, as intermediate TRFs (between static and very dynamic, 19 Hz) show the longest MAE durations for intermediate adapting speeds.*

were free to vary (except for the scaling factor c). If different motion systems were at the origin of the experimentally found results, one would expect μ and/or σ to increase with increasing TRF. If differences in the velocity content of the test stimulus were responsible for the found results, as predicted by the single motion system hypothesis, one would expect σ_{test} to increase with increasing TRF. The data from the different sessions were fitted separately (exemplar fits through the average data are shown in Fig. II.5 (curves)). Averaged values for μ, σ and σ_{test} for 0, 19, and 75 Hz were: $\mu = \{32, 15, 18\}$, $\sigma = \{29, 26, 27\}$, $\sigma_{test} = \{5, 25, 30\}$. Linear regression analyses on all collected μ, σ, and σ_{test} values showed that only σ_{test} was significantly positively related to TRF (p $<$ .0005). These data suggest that the different positioning of the curves is not related to differences between adaptation states (characterized by μ and σ and associable with different motion systems), but instead to differences in the velocity content of the various test stimuli (defined by σ_{test}). Results were very similar among subjects.

Prediction 4: Discontinuous transitions disappear with small speed differences

The model suggested that discontinuous transitions should not be observed when speed differences between the adaptation components are small relative to the width of adaptation in the speed dimension (σ in the model). We did not test prediction 4 in a separate experiment, but assuming that adaptation processes differ between subjects, and that some subjects adapt a wider range of speed-sensors than others in identical conditions, we predicted continuous transitions for some subjects and discontinuous transitions for others, while stimulus conditions are the same. Indeed, in experiment 2 (Fig. II.4b), the observer denoted by the gray triangles shows a continuous transition between fast en medium speeds, whereas the other subjects show a two-peaked distribution. Also, subject FV in Ref. [236] shows a continuous transition between slow and fast MAEs whereas the other two subjects do not.

II.5 Discussion

What are the mechanisms underlying our conscious perception of speed? Several studies have proposed a division in fast- and slow motion systems, but have put forward different accounts of where in the brain and how this division is realized and until what stage it is sustained. It has been proposed that the division may occur early in the visual system, at the level of magno- and parvocellular cells [57, 107, 101, 236], and may continue to exist until much of the motion processing is achieved and then come together at a perceptual level [57] or even that the two systems remain separated until the very end, and each have a private line to consciousness [101]. Our results do not support claims of two separate motion systems, but instead show that a single motion system can explain the human perception of speed. We show that dichotomous data obtained with different stimuli can not be interpreted as evidence of a bipartite (motion) system.

We have demonstrated that for the system processing luminance-based motion—arguably the most important source of motion information—(1) no strict link exists between adapting speed and type of test stimulus on which the strongest MAE is reported (Fig. II.3 , diamonds for medium speeds); (2) simultaneous adaptation to three motions can bring about three different MAEs (Fig. II.4); (3) static test patterns support the longest MAE after adaptation to slow motion, dynamic test patterns do so for fast motions, but intermediately dynamic patterns support the longest MAE at intermediate adapting speeds (Fig. II.5); (4) depending on the speed differences among the different adaptation components and on the subject, discontinuous transitions become continuous transitions; (5) a single-system model explains all data (Fig. II.1, II.3, II.4). The model shows (Fig. II.1) that discontinuities in experimental data do not necessarily indicate discontinuities in the visual processing. Discontinuous data are also produced by a single system along a sampled parameter dimension (e.g. the adapting velocity (Fig. II.5 , and Ref. [261]) and TVD (Fig. II.1, II.3, II.4 , and Ref. [236])) when the adaptation processes—taking place in two separate subpopulations—do not overlap extensively. None of the findings support the assumption of two independent speed-tuned motion systems.

Our results show how in different conditions transparent motions may lead to unidirectional and multi-directional aftereffects, depending on various stimulus parameters, and that all results are consistent with a single motion system. Our data further show that large-scale motion integration takes place in both direction and speed dimensions. Integration along the direction dimension may be precluded by large speed differences in the adapting stimuli, which is consistent with the research on plaid stimuli that has shown that the coherent motion perception (formed by motion integration) of the plaid decreases as the speed differences between its component motions increases [132].

There exist, however, a large body of data that is interpreted as evidence for a distinction in slow and fast motion systems, both psychophysical and clinical. Below we will go to some length in discussing these data to show that the data are not necessarily indicative of two motion systems, and may be incorporated in a single motion system account.

Slow chromatic and fast achromatic motion systems have been proposed in the psychophysical literature [57, 90, 91, 97, 107]. Reaction time to motion onset [57], and motion identification thresholds [91] were found to increase more steeply for chromatic stimuli than for achromatic stimuli when using increasingly low physical stimulus speeds (or temporal frequencies). However, these increases were nearly identical for both conditions when replotted against perceived speed [57]. These data were interpreted to mean that there exist separate motion systems for chromatic and achromatic motions, and that the outcome of these systems converge at a subsequent level [57]. However, the differences between chromatic and achromatic conditions are well explained by assuming different contrast-transduction properties of early temporal filters without assuming different motion systems [186, 187]. Also, the discovery of low and high temporal frequency channels [9, 14, 249] has incited claims that slow and fast motion systems exist. However, it has been suggested on theoretical [105, 249, 215] and experimental [69, 249] grounds that these channels operate not independently but jointly to create a representation of speed at a later stage. These findings can therefore be incorporated in the one-system hypothesis.

Research on motion detection in noise has shown that motion detection is impaired when noise dots move with similar speeds as signal dots, but not when both speeds are very different. Such was found for fast and slow signal speeds [75, 130]. These results—although interpreted as evidence for two motion systems—are fully compatible with our findings that, under certain conditions, subpopulations of a single motion system may adapt to (and thus process) fast and slow motion without much interaction. Such a mechanism may also be the key to explain the data reported in Ref. [101], who showed that presenting different motions to the two eyes results in binocular rivalry if both motions were slow, or fast, but resulted in transparency if one motion was fast and the other slow (*cf.* [132]).

In clinical studies, it has been suggested [80] that there exist two separate pathways to motion area MT: one pathway projecting directly from LGN to MT (for fast motions, >6deg/s), the other one relayed through V1. Although much of the data suggesting the existence of the LGN-MT pathway has been criticized [24], a direct LGN-MT pathway seems to have been found, recently [235]. The available data however suggests that the direct koniocellular LGN-MT projection [235] is not involved in motion processing per se [23, 189, 200, 225], but may function as a fast attention-grabbing (and possibly directing) device. Such a device should be specifically sensitive to fast motions [61, 80] as they pose a greater threat to the survival of an organism (embodied in approaching predators, or e.g. cars and trains), for which the more time-consuming motion calculations in V1 should be bypassed. Therefore, the finding of a direct LGN-MT projection is not in conflict with our findings.

One other possibility is that slow- and fast-motion systems are located in anatomically separate brain areas. Consistent with this idea is the finding that a patient with a bilateral lesion in MT could not see motions faster than about 6deg/s [295], whereas a patient with V1 lesions could not see motions slower than about 6deg/s [29]. However, patients with V1 lesions were unable to determine the direction of motion of the stimulus as it contained little position information [23]. Similarly, the patient with MT lesions did show only saccadic pursuit for fast motions, and she did not perceive a motion aftereffect after prolonged periods of motion stimulation [295]. The rudimentary motion perception for both types of lesion seems to depend on a position tracking (and not a motion) system.

V3A is also an area that is found to be motion sensitive. Much is still unknown about its motion specificity, but it seems to have speed sensitivities almost identical to MT [61], which speaks against a division of labor between V3A and MT along the speed dimension. Its inferior activity to motion stimuli relative to MT [254], seems to suggest that V3A is not very specific for motion, which dovetails nicely with the suggestion that it processes 3D shape or layout [255] which can be derived from multiple cues among which motion information. Our results add to this idea that a single functional entity underlies the found MAEs, and it is then most parsimonious to assume that this is also a single anatomical substrate.

Overall, a single motion system is well able to explain the available data on motion processing. Our study reconciles a large body of research that lent support to the single-system account of motion perception with the psychophysical literature that repeatedly reported evidence for a division in fast and slow motion systems. We have demonstrated that a two-system hypothesis is neither sufficient (Experiment 1, 2 and 3) nor necessary (Model) to explain the data on luminance-base motion perception. Therefore we argue—opposing previous psychophysical studies [57, 75, 90, 91, 97, 107, 130, 101, 236, 261], but in line with currently known neurophysiological findings in monkeys [156, 182, 214]—that a single luminance-based motion system can explain human speed perception. This finding accords well with the generally accepted view of a single system for the processing of motion direction [12, 178, 179].

Acknowledgements

The authors would like to thank Dr. W.A. van de Grind for critical discussions and P. Schiphorst and T. Knapen for technical assistance.

Chapter III

A single motion system suffices for global-motion perception

Jeroen J. A. van Boxtel
Casper J. Erkelens

Global-motion perception is the perception of coherent motion in a noisy motion stimulus. Thresholds for coherent motion perception were measured for different combinations of signal and noise speeds. Previous research [75, 130] showed that thresholds were elevated when signal and noise speeds were similar, but not when they were different. The regions of increased threshold values for low and high signal speeds showed little overlap. On the basis of this evidence two independent speed-tuned systems were proposed: one for slow and one for fast motion. However, in those studies only two signal speeds were used. We expanded the results by measuring threshold-curves for four different signal speeds. Considerable overlap of the threshold-curves was found between conditions. These results speak against a bipartite global-motion system. Model simulations indicate that present and previous experimental results can be produced by a single motion system provided that the mechanisms within it are speed-tuned.

III.1 Introduction

Motion information is abundant, and therefore a useful and often-used source of information. Research on motion perception has suggested that the processing of motion information from the retina to perception consists of at least two processing steps [2, 276, 234, 294, 133]. Motion signals are initially picked up by cells that have receptive fields spanning less than a degree [121, 222, 293], hence this step is called the local motion-processing step. Often the gathered motion information is rather noisy, and pooling of velocity information may be essential to obtain a reliable indication of the motion of the observed object. Such a pooling process is thought to occur at a global-motion processing stage in which information from several local-motion processing units are combined [193, 224, 244]. In this study we are interested in the representation of velocity (i.e. direction and speed) at the global-motion processing level. We are particularly interested in the question whether global-motion processing is achieved by one or several (independent) systems.

MT is the cerebral location where one thinks global-motion processing takes place[‖]. Indeed, cells that process global motion have been found in MT [193, 224, 239, 244]. In MT the representation of direction changes smoothly over the cortex [175]. This structural map is probably the resultant of an evolutionary pressure that pushes neurons that functionally interact closer to each other [62]. However, no clear-cut evidence exists for speed maps. Nevertheless, nearby cells have a tendency to have a similar speed preference [156]. These neurophysiological results are suggestive of a single motion system processing the whole range of directions and speeds, but they do not exclude the possibility of more motion systems.

Psychophysical data also suggests that direction information is processed in a single system. Specifically, motion aftereffect (MAE) studies have found that adaptation to two motions of different directions (but identical speeds) results in a single motion aftereffect opposite to the vector-average of the two adapting components, suggesting that all the direction information is integrated [236, 178] Furthermore, when two moving planes are presented simultaneously the difference in direction of motion is perceived larger than veridical. This phenomenon is called motion repulsion or direction repulsion [178, 68, 51, 115, 133, 176], and shows that long-range interactions exist in the direction-dimension. Motion repulsion takes place for angles between two motion vectors of up to 120 degrees [178, 51], again indicating that direction information is processed by a single system.

On the other hand, other psychophysical data does give support to claims that multiple systems exist for speed processing. First, it is found that motion detection [57] and discrimination [90] thresholds are higher for chromatic than for

---

[‖] Area MST is also important in global-motion processing, but seems to be implicated in more complex motion analysis than MT (e.g. [48, 260]).

achromatic motion when stimulus speed is low, but these thresholds are identical when stimulus speed is high. These data suggest that there may be three motions systems: one for slow chromatic motion, one for slow achromatic motion, and one for fast (chromatic and achromatic) motion [57, 90, 107]. Second, MAE experiments show that for certain speed combinations of two motion-adaptation components, the resulting motion aftereffect is not opposite to the vector average of these components. Adaptation to two adapting components of different speeds leads to two MAEs of differing speed, which may be brought about separately or simultaneously [236, 263]. These data have been interpreted as evidence for independent slow- and fast-motion systems. Within each motion system direction is integrated over the full 360 degrees, but between systems no such integration takes place. Third, research on motion detection in noise has shown that motion detection is impaired when noise and signal dots move with similar speeds, but not when they move at very different speeds. Such was found for fast and slow signal speeds [75, 130]. Again the data was interpreted as evidence for independent fast and slow motion systems, this time for global-motion processing.

Although the psychophysical data is repeatedly interpreted in terms of two (or sometimes more) motion systems, direct neurophysiological evidence for such a division is absent. Moreover, some psychophysical counter-evidence has been reported (Refs. [186, 187], and Chapter II). These studies show that the results that were interpreted as evidence for high and low speed systems may be reinterpreted in terms of a more parsimonious single-system explanation (see discussion section). For the last line of evidence (i.e. motion detection in noise) no such studies were found. Inspired by the previous studies reporting counter-evidence, we set out to investigate if the two systems that are proposed for global-motion processing could be united into a single system. Measuring thresholds for global-motion perception, we find that the detection of global motion is impaired by noise that moves at similar speeds as the to-be-detected coherent motion, as reported previously [75, 130]. However, when using more than two signal speeds, we find no indication of a division in fast and slow motion systems. Noise has a systematic degenerative effect on global-motion detection over wide ranges in the speed dimension, and the effect shows a tight link to the signal speed used and not to a particular speed range, which would be expected if it were linked to a speed-specific motion system. We developed a single-system model to investigate if the obtained pattern of results could be the result of a single continuous system, and found that this indeed appears to be the case.

Before continuing, we would like to clearly define the terminology we adopt in this paper. The word “system” is used, as it is generally done, to imply that the processing of the information in question differs from that of other information. The processing of two systems is independent, and may involve different mechanisms/algorithms, and may also take place in different brain areas and often involves different modalities. Two motion systems respond therefore to different kinds of stimuli. For example, in MAE experiments it is assumed that the slow-motion system responds specifically to non-dynamic test stimuli, and the fast-motion system specifically to dynamic ones (e.g. Ref. [236]). Or that one activates different motion systems with colored (isoluminant) and luminance-defined motion stimuli (e.g. Ref. [107]), on the bases of the finding that both have different contrast-dependencies. The word “channel” has three main usages, (1) it is the term for the psychophysically obtained profiles of the threshold-curves using masking-procedures; (2), it may be used to describe a set of units (e.g. neurons) that in a functional way form an entity that spans a specific and rather well-delineated region of the investigated domain; for example such a channel would be the main entity that codes for speeds of, say, 3 to 7 deg/s. The entity would be the direct neuronal implementation of the psychophysically measured channel and would as such have the same tuning characteristics as the psychophysically measured channel. It would consist of a fixed set of units that responds as a whole to part of the investigated domain. Other channels consist of a completely different set of units, which may however partly overlap in their tuning to e.g. speed. When these channels are shown to be truly independent, one could argue that they represent separate (independent) systems. (3) The word channel may also refer to a single unit (neuron) that codes for a limited part of the investigated domain. As is clear from this list of definitions, the word channel is used to describe different things that are obviously related but may happen at quite different stages in the visual system, which complicates discussions on these issues. We will use the terms “psychophysical channel”, “neurophysiological channel”, and “neuron” or “unit” for definitions 1, 2, and 3, respectively[**].

III.2 Experiment 1

Previous studies [75, 130] investigating the detection of coherent motion in noise, have found that noise impairs global-motion detection only when the speeds of noise dots were similar to that of the signal dots (carrying the coherent motion signal). The following paradigm was used. Stimuli consisted of seventy-five dots, containing a variable number of signal dots that defined the coherent motion, the other dots being noise dots. The number of signal dots necessary to correctly identify a movie-sequence as containing a coherent motion, was measured and called the detection threshold. Signal and noise dots had identical speeds, but the noise dots had an individually assigned random direction of motion. The thresholds in this condition functioned as comparison for conditions in which additional noise was added. The seventy-five additional noise dots also moved in random directions, but had a speed that varied over trials, which could be different from or identical to the speed of the signal dots. In the case that the speed of the additional noise dots was identical to the signal (and the other noise) dots,

---

[**] Note that within a system several channels may be found.

thresholds increased as expected. However, when the speeds differed, thresholds rapidly fell to comparison levels, as if no additional noise was added. The resulting threshold curve is a psychophysical channel.

Using slow (1.2 deg/s) and fast (9.6 deg/s) signal speeds there was little or no overlap between the psychophysical channels (i.e. the speed ranges where the increased thresholds were found) [75, 130]. On the basis of this finding the authors concluded that at least two independent speed-tuned global-motion systems exist[††]. However, these authors used just two signal speeds, which were rather far apart. To test if the regions of increased thresholds are indeed separable along the speed dimension, and identifiable as separate and independent systems, we decided to increase the number of signal-speed conditions. Furthermore, we decided to measure thresholds for higher values of noise speed, to investigate if thresholds indeed remained elevated for high-noise conditions, when high signal speeds were used [75, 130]. We show in this report that these separate channels may not be interpreted as independent motion systems when their profiles show little overlap.

Method

Observers

Three subjects with normal or corrected-to-normal vision participated as observers, two were experienced (JB, first author; TK), one (AK) was an inexperienced observer. AK was naïve as to the purpose of the experiment.

Stimulus

The stimuli were brief motion sequences of random dot patterns that did or did not contain a coherent motion embedded in noise. All sequences contained 150 white dots that were divided into two groups: a signal-containing group of dots, and a group of additional noise dots, each 75 dots large. Dots were of high luminance (about 70 cd/m^2), sized 3.1 arcmin, and were displayed on a black background (about 0.06 cd/m^2). The dots were randomly positioned on the first frame, and had an infinite life-time. Of the first group a variable number of dots was assigned to be signal. Those signal dots moved en group in a certain direction (chosen randomly every trial), and with a given signal speed. The other dots of the first group were noise dots that moved in random directions (changing every frame-transition) and had the same speed as the signal dots. At each frame transition, dots from the signal-containing group were assigned anew to be signal or noise. The second group of dots (the additional noise dots) moved in random directions (changing every frame transition) and with a certain speed (not necessarily the same as that of the signal-containing group). The display was 10.3° by

---

[††] The authors did not mean psychophysical channels as they were speculating that the two systems could have functions that a single system could not have (Ref. [75], p. 1574) and could possibly involve different cortical processing (Ref. [130], p. 3041).

10.3° of visual angle and was surrounded by a single row of high-contrast stationary and filled squares (12.6 arcmin), that were positioned 25.2 arcmin from each other, and had a 70% chance of being visible (i.e. about 30% of the positions were left empty). Motion dots that moved outside the display were wrapped around. Each sequence consisted of 30 frames, refreshed at 75Hz, yielding 400ms long movie sequences. A small fixation mark (size: 3.1 armin) was present at the center of the display.

Procedure

A two-interval forced-choice procedure was used to determine the number of signal dots needed for a coherent motion to be detected. One of the two intervals contained a coherent motion signal, whereas the other contained only noise movements. The two intervals were separated by a 500 ms blank period. Observers indicated in which interval a coherent motion was perceived. We used the Quest procedure [273] to find the number of signal dots needed for the observer to correctly identify in 75% of the trials the sequence with coherent motion, called the threshold for coherent motion perception. The procedure started with 18 coherently moving dots, and was aborted when, within a staircase, the Quest procedure gave two successive threshold-estimations that differed by less than 0.001 dots (this value was determined in tests that showed that the then-reached threshold (in number of dots) would remain the same irrespective of the subjects subsequent answers).

Thresholds were obtained for all combinations of 12 additional-noise speeds and 4 signal speeds (noise speeds: 0.6, 0.8, 1.1, 1.6, 2.1, 2.9, 4.1, 5.6, 7.8, 10.6,14.5, 20.0 deg/s; signal speeds: 1.1, 2.1, 5.1, 10.6 deg/s). For every signal speed, the whole set was divided into 3 sessions, containing 4 interleaved Quest staircases each (session 1: 0.6, 1.6, 4.1, 10.6 deg/s; session 2: 0.8, 2.1, 5.6, 14.5; session 3: 1.1, 2.9, 7.8, 20.0 deg/s). The order of these sessions was random between subjects, and also within subjects between the three repetitions of each condition. It has been shown that the used procedure really taps a speed-tuned system and not a stepsize-tuned system, and that the results are not dependent on refresh-frequency or perceived contrast [75] in the range of parameters used in the present experiment.

Apparatus

Viewing was binocular. A Macintosh PowerPC G4 drove the experiment. Images were presented on a LaCie electron22blueIV monitor. Subjects used a chin-rest to stabilize head position. Experimental procedures were reviewed and approved by the Institutional Review Board and each subject gave informed consent.

	Signal speed in deg/s (log value in brackets)			
	1.1 (0.10)	2.1 (0.74)	5.1 (1.63)	10.6 (2.36)
Parameter				
ω	11.91	9.09	6.98	6.67
α	3.11	3.73	5.46	6.07
μ (log)	0.37	1.08	2.02	2.46
σ (log)	0.53	0.49	0.57	0.59
R^2	0.77	0.91	0.96	0.95

Table III.1: **Values of the fit parameters, and the R^2-values of these fits**

Results and discussion

Figures III.1 and III.2 show the results for one of the observers and of the averaged data, respectively. Shown are the threshold values (number of signal dots) for coherent motion perception plotted against the speed of the additional noise dots. Thresholds are low over rather broad ranges, but peak when the speed of the additional noise dots is similar to the speed of the signal-containing group of dots. Considerable overlap between the curves exists. The size of the effects is of the same order as those reported previously [75].

We fitted log-Gaussian curves to the mean results to identify the positions of the peaks of the different psychophysical channels and assess the width of the speed tuning. The fit-function is $\omega + \alpha G(\mu, \sigma)$, where ω is the offset, α the amplitude, μ is bias (i.e. the position of the maximum threshold), σ the width of the curve, and G is a Gaussian function in log-space. All four parameters were free to vary. The curves fitted through the averaged data are shown in Figure III.2. Table III.1 shows the obtained parameter values in the different conditions. From these results it can be concluded that with increasing signal speed, detection of coherent motion is easier (i.e. offset lower), also there exists a tight link between μ (bias) and the signal speed. The values of α (amplitude) increased with increasing signal speed, which means that best and worst performance deviated more and more from each other with increased signal speed. The values of σ were constant among all four conditions. A constant σ indicates that speed-tuning curves have a constant width in log-space (here being about 0.5 Log-units, about 0.8 octaves) independent of the preferred speed, which is indeed found for macaque MT cells [182]. Edwards et al. [75] and Khuu & Badcock [130] compared curves with signal speeds of 1.2 deg/s and 9.6 deg/s, and found that the regions of increased threshold values hardly overlapped. They argued that because of the absence of overlap, at least two largely independent global motion systems could be assumed to exist: one for slow motions and one for fast motions. Our dataset, however, also included two intermediate signal speeds. The finding that the two extreme channels in our study do not overlap suggests that the populations of neurons coding for 1.2 deg/s and 10.6 deg/s are not overlapping (i.e. operate independently) for our specific stimuli. However, when we follow the line of reasoning

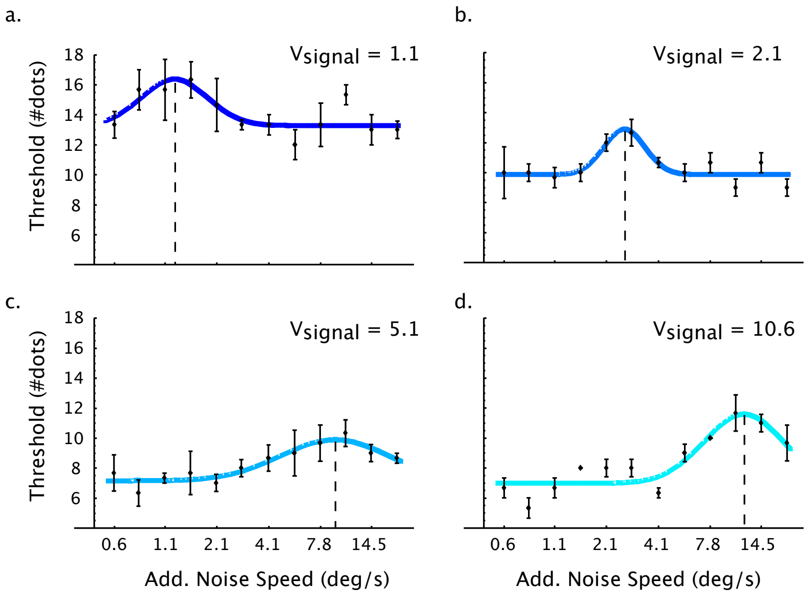


Figure III.1 | Threshold values—in number of signal dots—for the detection of coherent motion in noise for subject TK. *Data points are means ± 1 SEM, in conditions where the signal speed is* (a) *1.1°/s;* (b) *2.1°/s;* (c) *5.1°/s; and* (d) *10.6°/s. For illustrational purposes log-normal fits through the data are also shown. The vertical dashed line shows the speed of the additional noise dots at which the highest thresholds are obtained. It can be seen that a strict link exists between the used signal speed and the additional noise speed at which the maximum is reached.*

used by previous authors, we should conclude that the psychophysical channels at 1.1 deg/s and 2.1 deg/s are not independent of each other, and that those at 2.1 deg/s and 5.1 deg/s are not independent of each other, and that the 5.1 deg/s and 10.6 deg/s channels are not independent of each other. The evidence that the channel at 1.1 deg/s and that at 10.6 deg/s are independent of each other (in the current conditions), which might indicate separate speed-tuned systems, is overshadowed by the evidence that the series of channels are in fact part of a single system. Furthermore, there is no indication of a discontinuity in the data. Instead, it seems that the position of the peak increases in a continuous manner with increasing signal speed.

Finally, it was previously reported that with high signal speeds the thresholds of coherent motion detection remained high even when noise speed was much higher than signal speed [75, 130]. We measured thresholds over a more extended range, and found that for sufficiently high noise speeds thresholds drop again (see Figure III.2 a, bottom right panel), suggesting that the absence of a drop in thresholds [75, 130] was due to a too restricted range of noise speeds.

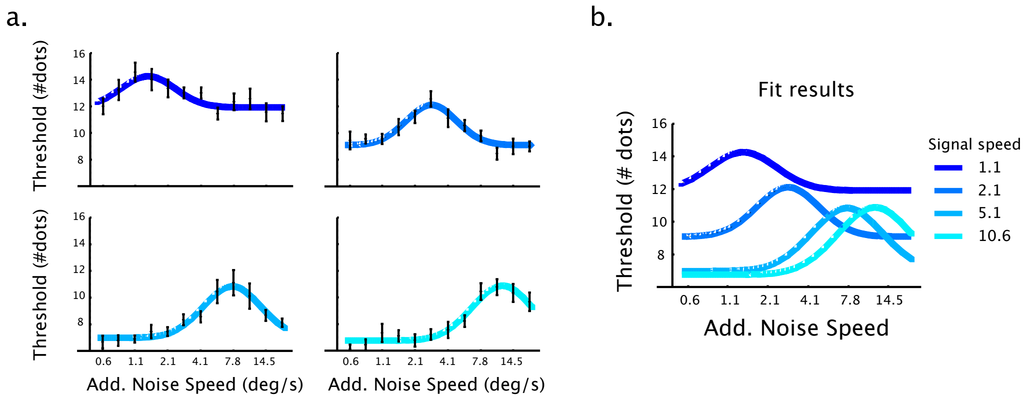


Figure III.2 | Threshold values in number of signal dots for the detection of coherent motion in noise for all subjects combined. *Data points are means ± 1 SEM.* (a) *Data is presented as in Figure III.1 , but now for all the subjects together.* (b) *The fits from panel* a *are replotted together. The fits show a steady migration of the top to higher additional noise speeds when using increasingly high signal speeds. Two other points to note are that the offset (i.e. the minimum threshold value) of the curves decreases with increasing signal speeds, and the width of the curves (at the point halfway from the maximum to the offset) is rather constant.*

Model

The present experimental data oppose a division in fast and slow motion systems. In order to explain present and previous data, we decided to model a motion system based on the single-system assumption. How could motion be represented in a motion system? When an object moves, it casts an array of signals on the retina that all have a certain speed and direction. Motion area MT processes both direction and speed information. Direction information is represented in an organized manner. Neurons responsive to similar directions are situated in cortical columns [12]. Moreover, motion direction preference changes smoothly over the cortex [175], and within MT connections exist preferentially between neurons of similar orientation preference ([174]; orientation preference is tightly coupled to direction preference in MT, e.g. Ref. [175]). The representation of speed is much less well understood, but speed is represented in area MT [214, 217, 182, 156], and it seems that neurons positioned near to each other have similar speed preferences [70, 156].

A given motion stimulus activates neurons responsive to a variety of directions and speeds. More specifically, neurons optimally responsive to the stimulus velocity will be activated most, but neurons optimally responsive to velocities slightly different from that of the stimulus, will be activated nonetheless, but to a lesser degree. In this way a population activity may emerge as depicted in Figure III.3b. A global-motion stimulus contains a coherent motion signal, but also lots of noise, caused by spurious dot matching between successive frames.

Both the noise and the coherent motion signal will activate the motion system. In the present experiment and those performed previously [75, 130], noise velocities were mostly limited to a certain speed-band, without a predominance of any particular direction signal. Two examples of a population activity caused by this kind of noise are shown in Figure III.3c.

The extraction of motion signals from luminance changes on the retina can be seen as local motion extraction. A global-motion system may operate independently of this local-motion system, but it has been show that this is not the case. The global motion system processes information obtained by the local motion system (e.g. Ref. [130, 237, 244, 192]).

Therefore the next step in our model, the global-motion extraction step, builds on the motion signals obtained by the local-motion system. We added the activity caused by signal and noise dots, and for every direction-speed combination we calculated a measure of how distinctly the model was activated at that direction-speed combination relative to other direction-speed combinations (see methods). High values mean a large distinctiveness for a direction-speed combination, which would translate into a low threshold for detection. Therefore, to compare the model and experimental results, we took the inverse of the maximum value of distinctiveness to obtain a measure comparable to a threshold for detection.

Methods

Motion representation

A motion stimulus of a particular direction-speed combination was assumed to activate a wide collection of neurons which themselves are maximally responsive to either the direction-speed combination of the stimulus, or a combination close to it. This dispersion of activity can be caused by both noise in the visual system and by (excitatory) interactions among neurons (e.g. Ref. [49, 174]).

The overall activation of the local motion sensors caused by the signal (i.e. the coherent motion) was modeled as:

$$S(d,s) = (G(s,\sigma)^{+} - G(s,\sigma)^{-})M(d,k)$$

being dependent on direction (d) and speed (s) of the signal dots, σ is the spread of the signal in the speed dimension, M is the von Mises distribution (a circular analogue of the normal distribution) modeling the spread in the direction dimension, k is the concentration parameter of this distribution (with k=0 representing a uniform distribution). G represents a Gaussian distribution in log-space. $(G(s,\sigma)^{+} - G(s,\sigma)^{-})$ represents motion opponency; if the spread along the speed dimension is large, a stimulus that moves in a 'plus' direction (say rightwards), will also active some neurons sensitive to 'minus' motions (say leftwards), in the motion opponency stage the activity in the left direction is subtracted from the activity in the right direction to obtain a clearer directional signal (see Figure III.6). Note that when σ is relatively small (as it is in our model), each stimulus will activate only part of the entire speed range in the system. It may therefore happen that two motions that are sufficiently different activate two non-overlapping populations of units in the system, causing them not to influence each others processing.

The activity caused by the noise dots was modeled as:

$$N(s) = (G(s, \sigma)^{+} - G(s, \sigma)^{-})M(d, k)$$

k was set to zero, yielding a uniform circular distribution, meaning that the activity caused by the noise dots was not modulated along the direction dimension. The total activity (T) caused by the global-motion stimulus was modeled as:

$$T(d, s) = b + S(d, s) + N(s),$$

where b is a baseline activity.

Measure of distinctiveness

The global-motion system is assumed to base its processing on the information provided by the local motion system, described in the previous section.

Based on the local motion information the global motion stage in the model computes a measure of distinctiveness (D) for every direction-speed combination as follows. It takes the value of activity at a certain speed and direction, and it divides this value by the value in the direction-speed space diametrically opposite to it (see two example dots in Figure III.3d):

$$D(d, s) = T(d, s)/T(d + \pi, s).$$

To obtain a measure comparable to a threshold for detection the inverse of the distinctiveness measure was taken, which we will call the model's threshold value.

The model has 3 free parameters: k, σ, and b. k was set to $3\pi/10$, but its precise values did not change the results qualitatively. σ was set to 0.5, close to the widths obtained in the Gaussian fits in the previous experiment. b was arbitrarily set to 5; the qualitative results are independent of the exact value of b. Parameter b influences the signal-to-noise ratio, as b is part of the noise. It will show that for low speeds, signal-to-noise ratios will be low (partly due to b), and thresholds higher. With low bs this effect if greatly attenuated, but still present.

Results and discussion

In Figure III.4 we show the results of the model. The model results echo those of the experiment in its main characteristic, namely that the top of the curves

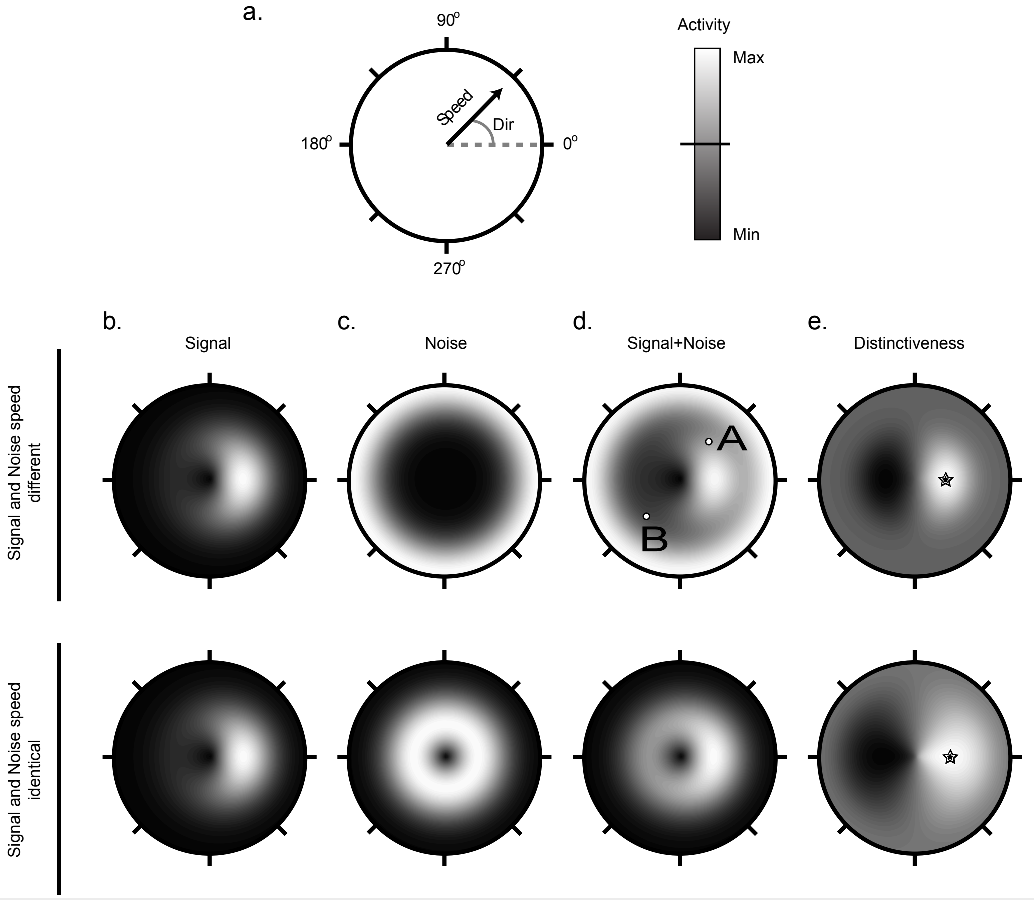


Figure III.3 | The representation of velocity in the model. (a) *Left, the polar plot shows a velocity vector, extending from the origin of the plot to a certain point in the velocity-space. The vector has an angle relative to a comparison direction (here right-horizontal), which is the direction component of the velocity. The length of the vector is the speed. In panels b-e, the top part shows a model condition in which signal and noise speed are different, the bottom part shows a condition in which they are identical. Grayscales are rescaled per plot.* (b) *The motion signal is directed (here to the right), and the representation of it shows a spread in both direction and speed dimensions. There exists only one maximum in the representation, and it is positioned at the presented stimulus speed.* (c) *The noise is non-directed, and contains only a certain band of speeds. In the top part of the panel, this band of speeds is positioned at rather high values, in the bottom part it is centered at the signal speed.* (d) *The sum of signal and noise signals.* (e) *The distinctiveness measure (see methods). It is calculated at the ratio of a certain point (e.g. point* A *in panel d) and the point diametrically opposite to it (point* B *in panel d). If done for all points, one obtains the figures shown here. The star indicates the point of maximum distinctiveness.*

moves to higher additional noise speeds with higher signal noise speeds (compare Figure III.2b and Figure III.4). Clearly, a single continuous motion-system model is capable of producing the present and previous psychophysical findings of psychophysical channels that move to higher speeds ranges when higher signal speeds are used. The model produces the feature of psychophysical channels without having separate neurophysiological entities (i.e. neurophysiological channels) representing each individual channel. Although the model contains units that are tuned to specific speeds, these units are not combined together as entities that span an a priori defined range of speeds (i.e. a neurophyiological channel) in order to produce the psychophysical channels.

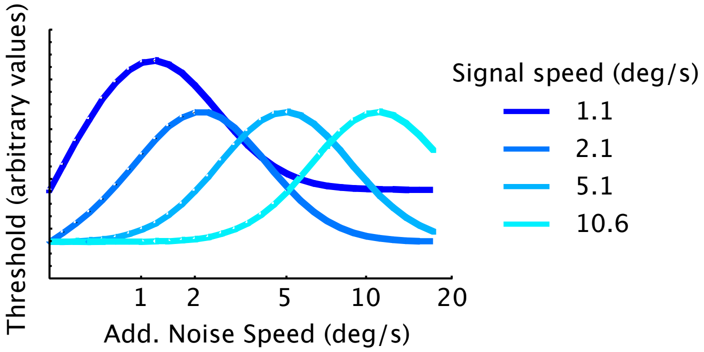


Figure III.4 | **Model results, simulating the experimental conditions.** *The different curves show the different signal-speed conditions. The model shows clear resemblance to the experimental data (see Figure III.2b). Notably, it shows a lawful shift in the position of the top of each curve towards higher noise speeds for higher signal speeds. Moreover, the results also show an increased offset for low signal-speed conditions, and a constant width of the different curves.*

Why does the model show these results? Figure III.5 shows a graph of the activity along the speed dimension, in the direction of the strongest signal; it is a horizontal slice through Figure III.3d ("signal + noise"). Figure III.5a shows the composition of the total activity in the case that noise and signal speeds are the same. Figure III.5b shows the same plots for the condition that noise speed is larger than the signal speed. The portions of activity that are caused by signal and noise have been denoted by "S" and "N", respectively. One can see that at various positions along the velocity dimension the signal makes up a large part of the total activity when signal and noise speeds are different. Its relative contribution to the total activity is smaller when noise and signal speeds are the same. Consequently, the distinctiveness of the motion signal is lower in Figure III.5a than in Figure III.5b, which is depicted in Figure III.5c. As a result of the low distinctiveness value, the threshold will be higher when noise is similar to the signal speed than when noise speed is different from signal speed, explaining why increased thresholds are found when noise and signal speeds are similar.

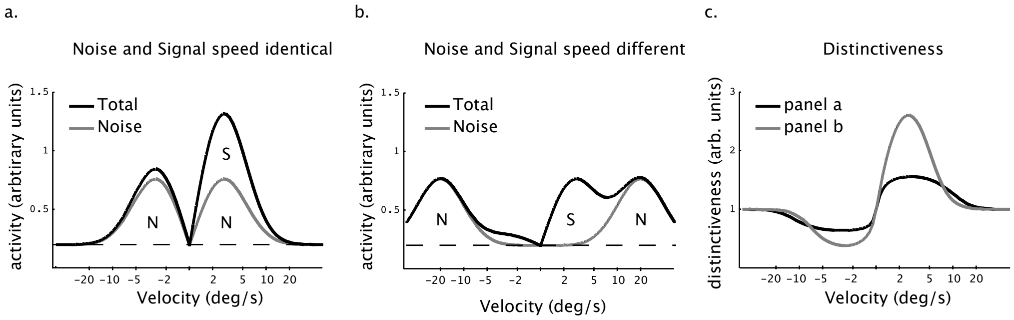


Figure III.5 | Explanation of the increased thresholds when signal and noise speeds are similar compared to conditions where they differ. (a) *A condition where noise and signal speed are identical. Shown are, the total activity (T) and the part of it that is caused by noise dots (N) and signal (S) dots. The figure is a slice through figure III.3d, slicing through the origin and the maximum; negative and positive velocities in this graph represent therefore motions in opposite directions (the absolute value of the x-axis being the speed of the motion). The horizontal dashed line represents the baseline (parameter b) in the model.* (b) *As in panel* a*, but now for a condition in which signal and noise speeds differ. By comparing panel* a *and* b*, one can see that the signal (S) constitutes a large part of the total activity over a wide range of velocities in* b *(from about 1 to 5 deg/s), whereas in a it is always accompanied by an increase in noise (in both "positive" and "negative" directions). Comparing the activities in opposite directions, as is done when calculating the distinctiveness measure* (c)*, it is clear that the relative contribution of signal (S) is lower in the case when signal and noise speeds are identical, compared to when signal and noise speeds differ. Therefore the final threshold (the inverse of the distinctiveness) will be high for conditions with identical signal and noise speeds.*

Interestingly, the model also reproduces the finding that with low signal speeds, the lowest obtainable threshold is rather high compared to the lowest thresholds of the other signal speeds. Figure III.6 explains how the model gives rise to the raised thresholds. Figure III.6a shows the distribution of activity along the speed dimension at the motion opponency stage (the noise is left out for clarity). The gray curve shows the distribution of activity ($G^{\pm}$) caused by a motion stimulus of low (left panel) and high (right panel) speed. Part of the activity is present in the negative part of the graphs (G^{-}). For low speeds this part is larger than for high speeds. In the motion opponency stage this activity is subtracted from the activity in the positive part of the graph (G^{+}). The final activity after the motion opponency stage ($G^{+} - G^{-}$) is represented by the black curve, which is the difference between the gray dashed ($G^{-}_{mirrored}$) and continuous line. The top of the high-speed curve is higher than that of the low-speed curve, because it is less inhibited by its negative tail (G). This difference in maximum activity is found back in Figure III.6b where the total activity ($T(d, s)$) is depicted for both high-speed (black line) and low-speed stimuli (gray line). Figure III.6b is a slice through the velocity space similar to the one shown in Figure III.3d, and the small peak at the negative side of the graph is some activity that has spread over 180 degrees from the opposite side. Figure III.6c depicts the measure of distinctiveness for both low-speed- and high-speed stimuli, derived from the total activity. In our model calculations, the measure was only calculated for the point of highest activity (e.g. Figure III.6b), but one can see that this point coincides with the point of highest distinctiveness. One can see that the low-speed curve has a lower maximum than the high-speed curve. The smaller maximum results from the fact that the signal is relatively small compared to the noise (mainly the baseline activity) in the low-speed condition, since the signal was reduced by the motion opponency stage. Importantly, a low value of distinctiveness translates into a high threshold of detection, explaining why low-speed stimuli are more difficult to detect. The modeling results suggest that the motion opponency stage plays a major role in the increase of thresholds for low signal speeds.

In previous reports [75, 130] it was stated that measurements were executed in a range (which is identical to our used speed range) in which coherence thresholds did not vary before performing the experiment, while we find a general decrease in thresholds with increased stimulus speed. In support of our finding, at least one and possibly two of the three subjects in the Edwards et al. [75] study showed a decrease of thresholds with increased stimulus speed, just as our subjects. The thresholds of Khuu & Badcock [130] indeed seem not to vary a lot, but they were only measured at two points along the speed dimension. One additional point is that Khuu & Badcock [130] used rotational and radial motion and not translational motion, which may also have caused the difference in results. From a theoretical point of view, however, one should expect an increase of thresholds at low signal speeds, because in the limit of zero speed it is impossible to detect motion, and therefore thresholds should be immeasurably high.

III.3 General Discussion

In this study we examined the existence of independent fast and slow speed-tuned global-motion systems. These independent motion systems have been proposed for both translational [75] and rotational and radial [130] global-motion processing. The existence of such independent systems was proposed on the basis of coherent-motion detection thresholds. Coherent motion detection was found to be impaired when noise was added with similar speeds as the signal speed (i.e. the speed of the coherent motion). On the other hand such impairments were absent when noise and signal speeds were very different. Previous studies had used low-speed (1.2 deg/s) and high-speed (9.6 deg/s) coherent motion stimuli, and found that the increases of the thresholds for coherent motion detection (i.e. psychophysical channels) in these cases occurred in different speed ranges that did hardly overlap. These results pushed the idea of two independent motion systems: one for fast motions and one for slow motions.

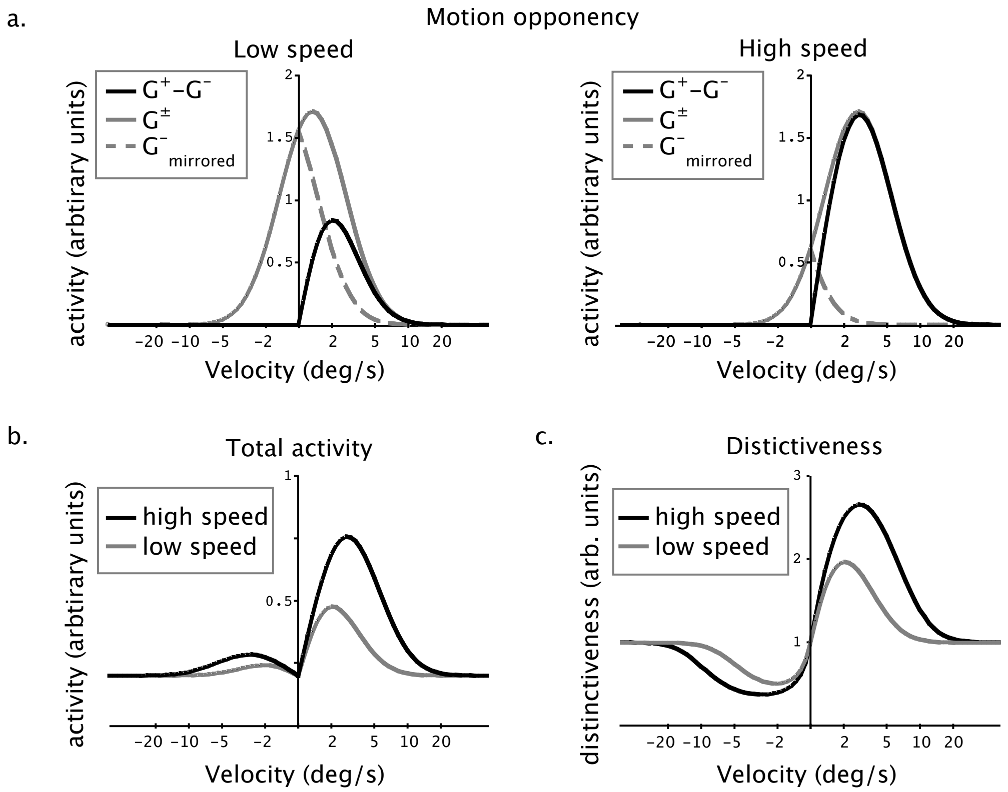


Figure III.6 | Explanation of the increased thresholds at low signal speeds. *At low signal speeds the smallest obtained threshold is much higher than at high signal speeds.* (a) *Left panel, a low signal-speed condition; right panel, a high signal-speed condition. The gray continuous line represents the activity caused by a moving stimulus before the motion opponency stage. A motion stimulus does not only activate neurons responsive to the speed of the stimulus, but also the neurons that are maximally responsive to slightly different speeds. A certain part of this activation is caused in neurons that are actually most responsive to motions in opposite directions (the part of the tail in the negative half of the graph). In the motion opponency stage this activity will reduce the activity of cells responsive to the direction of motion of the stimulus. Therefore, in the model the tail in the negative half of the graph is subtracted from the one in the positive half, yielding the activity after the motion opponency stage (black line: $G^+ - G^-$). The subtracted amount in the high-speed condition is however much less than in the low-speed condition, because the part of the tail in the negative half of the graph is much smaller.* (b) *The resulting total activity in high-speed conditions (black line), and low-speed conditions (gray line).* (c) *The distinctiveness values for high-speed (black line) and low-speed (gray line) conditions. The maximum distinctiveness value is lower in the low-speed condition.*

However, these studies used just two signal speeds. In the present study four different signal speeds were used covering the range from 1.2 deg/s to 10.6 deg/s. We found extensive overlap between the different psychophysical channels. The two extreme psychophysical channels in our study do indeed not overlap (suggesting independence). However, all other combinations of psychophysical channels do overlap, suggesting they resulted from the activation of the same range of speed-tuned neurons and are part of a single system. No indication was found for a dichotomy along the speed axis: the four different signal speeds we used led each to an increase of threshold at very specific values for noise speeds, tightly linked to the signal speed used. These results are not indicative of the existence of two independent global-motion systems [‡‡].

It may be argued that the two global-motion systems are not fully independent, but show some sort of interaction (as suggested previously [66]). However, if one proposes interactions between the two systems, the two-systems claim loses much of its value.

With our model we were able to show that a single global motion system is able to produce the present and previous data. We assumed that speed is represented in a continuous manner, from very low to very high speeds. Interactions among speed-tuned units exist, but are limited to a small speed range. This last point causes noise to have a detrimental effect on coherent motion perception only when its speed is within this range. When comparing curves that were obtained with signal speeds that lie sufficiently far apart, the limited range of the influence of noise speeds will automatically lead to an absence of overlap between the regions of increased threshold. The model indicates that this absence does not need to be a signature of two motion systems, but could very well be explained by a more parsimonious single-system account having a continuous range of speed-tuned units. The psychophysical channels were produced by our continuous model without having separate neurophysiological entities (i.e. neurophysiological channels) representing each psychophysical channel.

Position of the global motion system

The extraction of global motion from local motion signals is generally thought to take place in motion area MT [193, 224, 244]. Although our results do not constitute any direct evidence for the placement in MT, when we compare our curves of the different signal-speed conditions (which are based on the combined activity of many motion sensitive neurons) to the speed tuning curves of single neurons, the tuning widths of our curves are similar to, though somewhat sharper, than

---

[‡‡] We note that Khuu & Badcock [130] used rotational and radial motion, whereas we used translation motion. Our experimental results may not be directly extrapolated to radial and rotational motion, especially since area MT is not very sensitive to rotational motion and area MST is [260]. However our results strongly suggest that the interpretation of Khuu & Badcock [130] may need revision. Moreover, our model does not necessarily need to be interpreted in terms of translational motion, and may well be used to explain the results obtained with rotational and radial motion.

those found in single cell recordings in monkey MT [182]. Of course one should be careful when comparing two such very different processes as single cell measurements and psychophysically measured perceptual performance.

However, our model is able to reproduce the experimental psychophysical data to a great extent, by incorporating the known physiology of MT for direction representation and assuming that speed is represented in a continuous manner in MT as well. The correspondence between the model results and the experimental results further support the idea that speed is represented together with direction in a continuous manner in MT, in accordance with some motion aftereffect data [231].

Motion opponency

Our model data resemble quite closely the data from our human subjects. The model did not only show increased thresholds around the signal speeds in different conditions (i.e. psychophysical channels), but also a general increase of thresholds at low signal speeds. A finding that follows from the model is that low speeds should be more difficult to detect than high speeds signals, which we find in our global-motion stimuli, but is also reported for other motion stimuli [57]. The occurrence of this increased difficulty was the result of interplay between a motion opponency stage and the existence of a baseline activity in motion-sensitive units in the model. Without the motion opponency this behavior would not be observed, suggesting that the motion opponency stage is an important step in (both local and global) motion perception. Indeed many models of local motion processing include a motion opponency stage [111, 234, 133], and evidence is reported from functional imaging in humans [112].

Relation to other studies supporting independent fast and slow motion systems

Several other studies have reported evidence for independent motion systems. These studies have been reviewed more extensively in Chapter II than will be done here. Here, we focus on evidence from psychophysical studies that are most related to the findings in this chapter.

The evidence for fast and slow motion systems comes mainly from two kinds of study: differences in motion perception in chromatic and achromatic stimuli, and motion aftereffects with static and dynamic test stimuli.

It has been reported that reaction time to motion onset is faster for slowly moving achromatic stimuli than for chromatic stimuli, whereas no such difference exists for fast moving stimuli [57]. Thresholds for motion identification were also found to be high for slow speed chromatic stimuli, but low for slow speed achromatic stimuli [90]. And it has been reported that perceived speed was contrast dependent for slow chromatic (isoluminant) stimuli, but contrast-independent for fast chromatic and all achromatic stimuli [107]. On the basis of these results, one proposed three different motion systems, one for slow chromatic motion, one for slow achromatic motion, and one for fast (chromatic and achromatic) motion [90, 107].

However for reaction time to stimulus onset the increases found for both types of stimuli were nearly identical when the results were plotted against perceived (and not physical) speed [57], indicating that these results stem from a single system. Indeed, the differences between chromatic and achromatic conditions are well explained by assuming different contrast-transduction properties of early temporal filters [186, 187]. The difference is therefore not necessarily due to two different motion systems but is well explained by differences in the involved low-level detection mechanisms.

The second major group of studies reporting evidence for fast and slow motion systems has found differences in motion aftereffect (MAE) durations on static and dynamic test stimuli. A MAE is an illusory motion perception in a direction opposite to a moving stimulus that had been viewed for prolonged periods of time. Such MAEs are seen on stimuli presented after the prolonged adaptation (called test stimuli). The test stimuli may be static or contain balanced motion signals (such that the net motion is zero). These last test stimuli are called dynamic. It is found that adaptation to slow motion gives rise to motion aftereffects with static test stimuli, whereas adaptation to fast motion does so with dynamic test stimuli [236, 263, 261].

In Chapter II we have shown that the seeming dichotomy may be explained in terms of a single motion system, much like we have done in this chapter. Briefly, it was shown that static stimuli only read out the motion adaptation to slow motions, because static stimuli only activate units sensitive to slow motion. Dynamic test stimuli contain many high-speed signals and will therefore also read out the adaptation of those units sensitive to high speeds.

III.4 Conclusion

The results of the current study indicate that there is little evidence for independent global-motion systems for fast and slow motion. Data interpreted previously as evidence for this division is well explainable in a single motion system model, within which each speed-tuned neuron covers part of the entire speed-domain. Other lines of evidence from achromatic/chromatic motion experiments, and motion aftereffect experiments have already been shown to fit a single motion system account. In the absence of direct evidence for independent slow and fast motion systems, one should embrace the more parsimonious explanation assuming a single motion system covering the full range of perceived speeds and directions.

Chapter IV

Temporal form parsing limits binocular conflict resolution in dynamic visual stimuli

Jeroen J.A. van Boxtel
David Alais
Casper J. Erkelens
Raymond van Ee

Presenting the eyes with spatially mismatched images causes a phenomenon known as binocular rivalry: a stochastic fluctuation of awareness whereby each eye's image alternately determines perception. Binocular rivalry is used to study interocular conflict resolution and the formation of conscious awareness from retinal images. We show that the conflicting images do not need to be presented continuously or simultaneously to elicit binocular rivalry (see also Refs. [209, 288]). Very brief stimulus presentations separated by large intervals up to 350 ms still elicit rivalry, even when the conflicting images are temporally non-overlapping. This continuation of rivalry in the absence of direct spatial conflict reveals that a temporally sluggish process underlies rivalry. We go on to characterize this process by showing that it is independent of low-level information such as interocular timing differences, contrast-reversals, stimulus energy, and eye-of-origin information. This suggests the spatial factors initiating rivalry relate more to higher-level form information. Systematically comparing the role of form and motion reveals that this temporal limit is determined by form conflict rather than motion conflict. Together, our findings demonstrate that binocular conflict resolution depends on temporally coarse form-based processing, probably originating in the ventral visual pathway.

IV.1 Introduction

We addressed a central problem in perception: how to parse a continuous stream of information into meaningful events, resolve ambiguities, and shape awareness. The time domain is important because in natural viewing conditions we make several saccadic eye movements each second and therefore incompatible binocular inputs may need to be resolved at the same rate. Binocular information may survive temporal blanks of up to 150-300 ms [209, 20, 73, 147], but at larger delays a 'temporal parsing limit' is reached above which the visual events become separately individuated percepts. We wished to learn whether the temporal parsing limit in rivalry is determined by inter-ocular delays as such, or by the temporal period of repeating form information, and whether it is influenced by other stimulus manipulations (such as contrast polarity, eye-of-origin information, and stimulus energy) that keep form information intact but disturb low-level information. We also examine whether the temporal parsing limit to binocular rivalry requires form conflict by testing with stimuli that conflict in motion but not form.

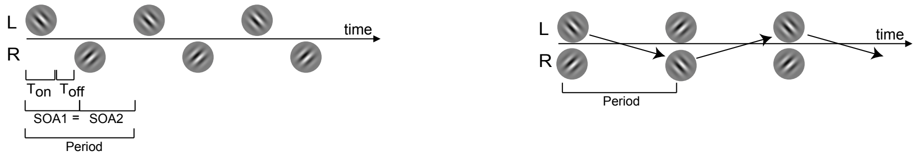
IV.2 Results

Determining the temporal limit of binocular rivalry

We determined the temporal limit for inducing binocular rivalry (BR) with temporally non-overlapping stimuli (Figure IV.1a; Ref. [209]). Brief grating presentations were delivered to the eyes in a repeating train of impulses, while varying the repetition period of the patterns between trials. Importantly, although the gratings were spatially conflicting, they were temporally interleaved between the eyes so that they were never simultaneously present. With very long repetition periods, every flashed grating was perceived individually (i.e. the visual stream of information was parsed into individual events), so that the observer perceived rapid and regular switches in orientation. With very short repetition periods (i.e. high stimulation frequency, approximating continuous presentation), normal BR was perceived, that is, a slow and irregular alternation between the two eyes' images. Between these qualitatively different perceptual manifestations, there was a transition-point indicating the temporal limit of BR with an average value of 377 ± 55 ms (4 subjects, here as well as in the remainder of the article, these values refer to the mean ± between-subject s.e.), consistent with previous estimations [209].

The independence from interocular delays

Given the temporal resolution of the visual system (psychophysically measured as high as 30-80 Hz [100]), one may assume that every flashed pattern in experiment 1 was parsed from the visual stream. Since the stimulus onset asynchronies



A. Experiment 1: temporal limit of binocular rivalry

C. Experiment 3: dependence on eye of origin

B. Experiment 2: dependence on inter-ocular delays

D. Experiment 4: dependence on contrast polarity

Figure IV.1 | Temporal layout of the separate experiments. (A.) *In experiment 1 the temporal limit of binocular rivalry was determined. Stimuli were gabor patches of orthogonal orientation in the two eyes.* T_{on} *is the presentation duration of the patterns.* T_{off} *is the time between two presentations.* T_{on} *and* T_{off} *together make up the stimulus onset asynchrony (SOA), which is identical for the pattern sequence from left (L) to right (R) eye, and vice versa. In this way we could determine at what repetition period binocular rivalry ceased to occur.* (B.) *Experiment 2: dependence on inter-ocular delays. Parameters as in experiment 1, except that the stimulus in R followed immediately the stimulus in L (or vice versa). Inserting various numbers of blank frames varied the period of repetition. This stimulus was used to determine whether the temporal limit observed in experiment 1 depends on the period of repetition or in stead on the SOA-values.* (C.) *Experiment 3: dependence on eye of stimulation. Stimuli in L and R are presented synchronously, but the orientations were swapped between the eyes on every new presentation (see arrows).* (D.) *Experiment 4: dependence on contrast polarity. Stimuli as in experiment 1, except that on every new presentation the contrast polarity was reversed (for illustrational purposes indicated with "+" and "−" signs).*

(SOAs) between two eye-patterns were identical in experiment 1 (i.e., the patterns were symmetrically interleaved: see Figure IV.1A), any inhibitory interactions between successive competing patterns would have been approximately balanced (see Figure IV.2A), causing each of the two patterns to be perceived about 50% of the time. However if every flashed pattern was indeed parsed from the visual stream, introducing asymmetrically interleaved gratings (Figure IV.2B) should result in strong perceptual biases. This prediction follows from dichoptic masking studies that have shown that inter-ocular inhibition operates nearly instantaneously, leading to strong inhibition at short SOAs and increasingly weak inhibition at longer SOAs [56]. Therefore, we predicted (Figure IV.2C), in accordance with Experiment 1, that rapidly alternating symmetric stimuli (see icon at the left-bottom of the graph) resulted in a 0.5 predominance of the lagging stimulus, but that as soon as any asynchrony is introduced (by increasing one of the two SOAs, icon on the right) a bias to the leading stimulus would be obtained (assuming forward masking).

We tested this prediction in experiment 2 using repetitive dichoptic masking

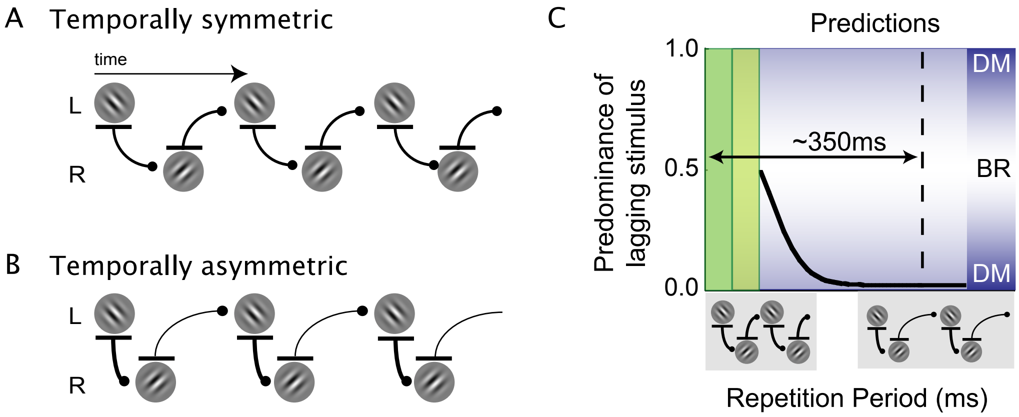


Figure IV.2 | Descriptions of possible interactions between the two eyes' images and the predicted perceptual consequences. (A) *and* (B)*: interactions between temporally resolved patterns in the two eyes, in cases of temporally symmetric* (A) *and asymmetric* (B) *stimulation. The presumed inhibitory interactions are drawn (lines with disk-heads), assuming inhibitory interactions are strongest at shorter and decrease at longer delays [56]. The strength of inhibition is indicated by the width of the lines.* (C) *The predicted perceptual consequences of asymmetric stimulation, assuming forward masking. The ratio of the percept duration of the lagging stimulus over total percept duration is shown on the y-axis ("Predominance of lagging stimulus"). The vertical dashed line represents the approximate position of the temporal limit found in Experiment 1. For temporally symmetric presentations (left, see icon bottom), the inhibitory interactions between the patterns are equal and therefore the percept is unbiased (a bias of 0.5, as in Exp. 1). For temporally asymmetric presentations (see icon on the bottom right) perception is biased towards the temporally leading stimulus because its inhibitory forces are stronger than the inhibition it receives from the lagging stimulus (see* B*). The dark and light green rectangles represent the presentation of the leading and lagging patterns, respectively. BR = Binocular rivalry, DM = Dichoptic masking.*

stimuli (Figure IV.1B; one SOA was ~50ms, the other variable), expecting large biases in the reported percept towards one of the patterns. Interestingly, and opposite to expectation, we observed that BR was unbiased. This result was independent of the inter-ocular SOAs, and held at all repetition periods up to 348 ± 33 ms (Figure IV.3A,B)—very close to the limit found in Experiment 1. The perceptual dynamics gave dominance durations typical of binocular rivalry with continuously presented stimuli, as well as the characteristic gamma distribution of dominance times (see Figure IV.3C). The lack of bias in experiment 2 suggests that the temporal limit for rivalry is not dictated by direct inter-ocular inhibition, and instead points to the proposal made above of a temporal parsing limit. In addition, it was only at temporal periods >~350 ms that we found a definite bias, i.e. dichoptic masking, towards one of the two patterns. Since it is only above this temporal limit that individual events are parsed from the visual stream (Exp. 1), it demonstrates that dichoptic masking does not occur before the parsing stage.

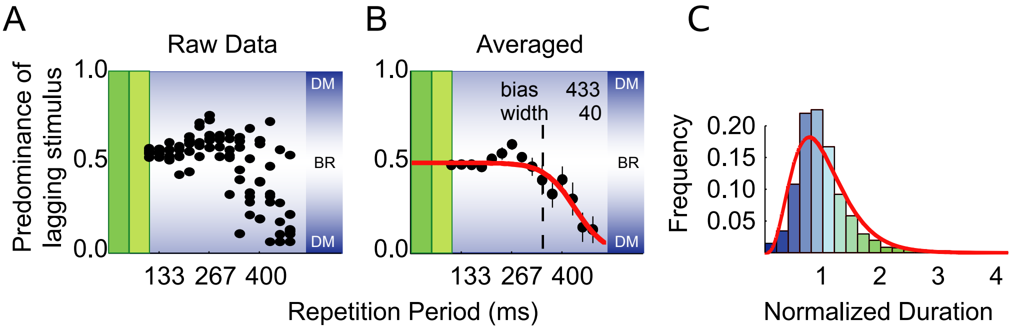


Figure IV.3 | **Support for a true parsing limit.** *The parsing limit identified in Experiment 1 is largely uninfluenced by asymmetric stimulation between the eyes, and therefore direct inter-ocular inhibition. This subject's data shows an unbiased perception for all periods up to about 433 ms, only for greater periods, inter-ocular inhibition takes the upper hand.* (A) *Data for single trials.* (B) *Means (± s.e.m.). The vertical dashed line represents the temporal limit found in Experiment 1. The small excursion towards unity near a period of 267 ms did not occur for other observers. Averaged over 4 subject the temporal limit is 348 ± 33 ms (mean ± between-subject s.e.).* (C) *Dominance duration distribution for the conditions with repetition periods of 213-267 ms. Even though the patterns in the two eyes are not delivered simultaneously, a gamma distribution characteristic of binocular rivalry is obtained. Normalized dominance durations were obtained by dividing percept durations by the average percept duration for each subject individually. Gamma fit parameters: shape: 4.2; scale: 0.24. The non-normalized mean percept durations are also similar to durations for conventional binocular rivalry: 2.6 ± 0.5 seconds (mean ± s.e.m. over subjects).*

The independence from eye-of-origin information

In experiment 3 we investigated whether binocular rivalry and the parsing limit depend on the eye of origin of the competing patterns. Both eyes' patterns were presented simultaneously and swapped between the eyes on every presentation [159], while varying the repetition period between trials (see Figure IV.1C). If the temporal limit were to remain unchanged at around 350 ms, it would indicate that it is dependent on the repetition frequencies of the patterns irrespective of the eye of origin. Alternatively, if the limit halves (i.e. repetition needs to be twice as rapid), the temporal limit is dependent on the eye-specific repetition period of a pattern. The data show temporal parsing limits that were close to the values found in our other experiments (330 ± 34 ms), indicating that the temporal parsing limit depends on pattern repetition frequencies irrespective of the eye of origin and irrespective of intra-eye repetition frequencies.

The independence from contrast polarity

In experiment 4, we used the paradigm of experiment 1, but now with gratings that were reversed in contrast polarity on every presentation (Figure IV.1D), thereby changing the luminance profile of the stimulus, but leaving pattern (i.e. orientation) information unchanged. If binocular rivalry were to depend on the repetition of identical stimuli within a single eye, repetition periods in this experiment should be halved. However, we again found a temporal limit not different from experiment 1 (352 ± 49 ms), demonstrating again a pattern-dependence.

The independence from stimulus energy

In experiment 5 we investigated whether stimulus energy (i.e. luminance) would influence the size of the temporal limit. 'On' and 'blank' times were both varied from 13 ms to 200 ms. We observed that the limit was not dependent on the luminance energy of the stimuli (one-way ANOVA, F(7,46)=0.51; $p > 0.8$; Figure IV.4), and thus the temporal parsing limit is not likely to be due to simple neural persistence or temporal integration.

A comparison of temporal limits

A one-way ANOVA revealed that the measured temporal limits across all experiments were not significantly different (F(5,24)=0.52; $p > 0.75$), and a post-hoc Tukey test showed no significant differences between individual experiments (Figure IV.5, which includes individual subject data). Based on the results of these experiments, we estimate the size of the temporal limit to be 363 ± 11 ms (mean ± s.e.m. over estimations from experiments 1-5).

Dependence on form conflict; independence from motion conflict

These experiments have shown the importance of form information in binocular rivalry relative to low-level information, but have not addressed its importance relative to motion information. The finding that the temporal limit of binocular rivalry is form-based suggests that its ∼350 ms boundary might reflect the temporal grain of the ventral form pathway. In a final experiment we tested this hypothesis.

We constructed a novel binocular rivalry stimulus that allowed for the presence of motion conflict without form conflict (i.e. orientations were always matched). Stimuli consisted of three superimposed gratings of different spatial frequency (Figure IV.6). Motion conflict was introduced by giving both eyes opposite motion directions (see the motion-conflict-only condition in figure IV.6A). Form conflict was introduced by creating a 30-degree orientation conflict between patterns (see the motion-and-form conflict condition in Figure IV.6B).

When both orientation and motion direction were in conflict (blue line, Figure IV.6C), we found that rivalry occurred both for continuous presentation and for repetition periods of 213 ms, but not for the 640 ms repetition periods. This indicates that with form and motion conflict, the 350 ms temporal parsing limit is maintained (*cf.* Experiments 1-5). However, when it was motion information alone that conflicted, rivalry was only observed for continuous presentations (red line, Figure IV.6C). For the 213 ms and 640 ms repetition periods of motion-only conflict, subjects perceived a stationary grating, indicating that the motion information from both eyes was integrated. Unambiguous motion (i.e. the same motion in both eyes) led to clear motion percepts (green line, figure IV.6C), indicating that the absence of motion in these cases is not the result of the flicker in the stimuli or from impoverished motion signals.

IV.3 Discussion

Although generally studied with continuously present stimuli [5], rivalry also occurs when the two eyes' images are not delivered simultaneously [209], or when they are flickered at relatively low rates [209, 288]. Instead of being an oddity of the visual system, this phenomenon may reveal a key constituent of the visual processing. It may also hold the answer to the question: If direct spatial and temporal conflict are not essential for binocular rivalry, what is?

First we determined the upper time limit for which alternating monocular presentation would still give rise to rivalry. Our results reveal two distinct time scales for temporal resolution in the visual system. On the shorter time scale, the brief stimulus presentations of $\sim$50 ms or so are perceived clearly as distinct stimulus pulses, as would normally be expected, given the visual system's temporal resolution. However, another longer time scale of approximately 350 ms is revealed when these stimuli are in interocular conflict, consistent with previous estimations [209]. In this case, we observe binocular rivalry with perceptual alternations between the two eyes' images at a rate much slower than the flicker rate of the stimuli. This alternation rate produces dominance durations typical of binocular rivalry with continuously presented stimuli, as well as the characteristic gamma distribution of dominance times (see Figure IV.3C). When this temporal limit for rivalry was surpassed, each individual stimulus in the alternating monocular presentation was perceived as an independent perceptual event, suggesting that the input stream of visual information had been parsed into separate events, thereby preventing binocular rivalry.

Next, we investigated what upheld the rivalry in the absence of direct conflict. One of the current theories suggests that rivalry is a process of low-level conflict resolution [34]. Therefore, we investigated how several low-level properties of the binocular conflict influenced the parsing limit that bounded the regime of binocular rivalry. First, introducing interocular delays between presentations should have biased perception towards one or the other percept if direct interocular inhibition drives rivalry, because interocular inhibition is strong at short delays, and gradually weakens at longer delays [56].

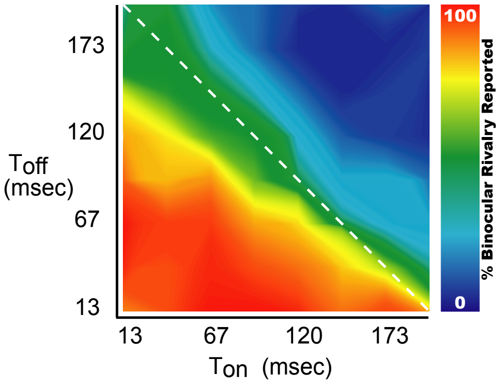


Figure IV.4 | Independence from stimulus energy. *Subjects viewed 6-second trials and indicated whether they perceived binocular rivalry or instead rapid and regular changes in orientation. The 'on'-time (T_{on}) and 'off'-times (T_{off}) were independently varied between 13 and 200 ms, in steps of 26.6 ms. The results depict the percentage of trials per condition categorized as containing binocular rivalry, averaged over subjects. There was no significant dependence of the temporal limit (= $2(T_{on}+T_{off})$) on the size of T_{on}, and therefore on luminance ($p > 0.8$). Instead the transition occurred at a repetition period of $\sim$400 ms (represented by the dashed line), irrespective of the size of T_{on}.*

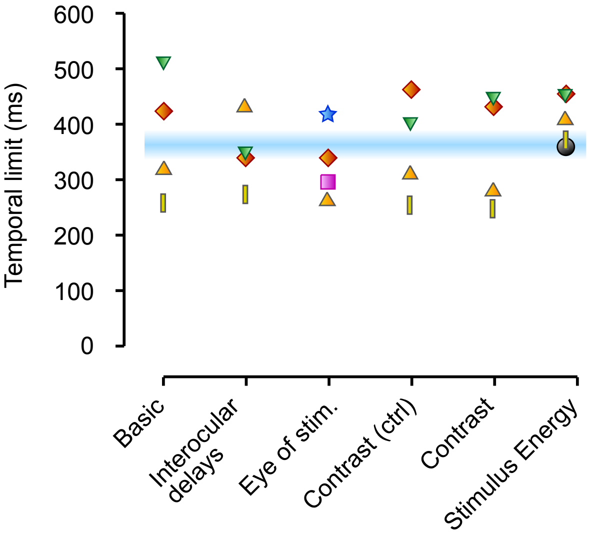


Figure IV.5 | Overview of the obtained temporal limits in experiment 1-5. *Individual subject data are denoted by different symbols. The label "Basic" refers to Experiment 1. Data in the Contrast (ctrl) condition was obtained with stimuli identical to those of experiment 1, which were interspersed with the stimuli of the contrast polarity experiment ("Contrast", experiment 4). The horizontal blue line represents the position of the estimated temporal parsing limit, based on these experiments (363 ± 11 ms, mean ± s.e. over experiments).*

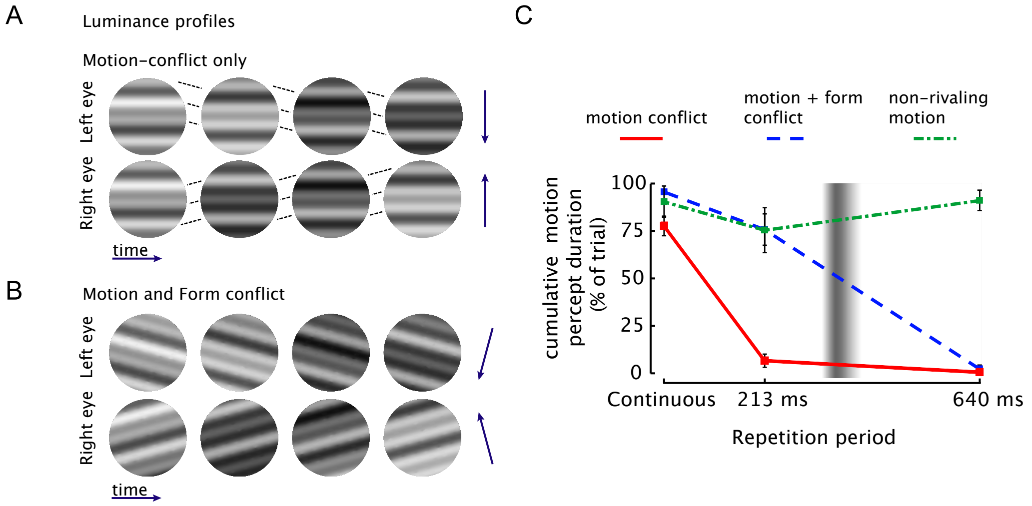


Figure IV.6 | The temporal limit for binocular rivalry is preserved with form conflict. (A) *Stimuli for the motion-conflict-only conditions have identical orientation, but different motion directions in the two eyes (see arrows on the right and dotted lines that trace some of the edges' motion over time).* (B) *Stimuli for the motion-and-form-conflict conditions have both different orientation and motion-direction.* (C) *Rivalry ensues in the form-and-motion conflict condition for repetition periods smaller than ∼350 ms (blue dashed line), causing most of the trial being spent in one of two motion percepts. In the motion-only-conflict condition (red continuous line), rivalry was only observed for continuous presentations, and therefore does not show the 350 ms temporal limit. At repetition periods of 213 and 640 ms, a static pattern was observed. Motion was, however, clearly perceived for all repetition periods when presented in non-rivalrous conditions (green dot-dashed line). Error bars are between-subject standard errors. The vertical Gaussian profile represents the estimated size of the temporal limit (363 ± 11 ms; mean ± s.e.m over experiments) taking data from experiments 1-5.*

However, we found no dependence on interocular delays. Biases, i.e. dichoptic masking, were only instantiated when the temporal parsing limit was passed (that is, when presentations were >∼350 ms apart). Therefore, only when the visual input stream is parsed into separated events may each event individually inhibit conflicting events, leading to dichoptic masking. Otherwise, at shorter intervals, even though still perceived to be intermittent, the stimuli nonetheless give rise to an impression of perceptual continuity while they are bound together into a continuous rivalry process. In these conditions alternations are slow and irregular and span numerous individual pulse cycles, as during conventional binocular rivalry. Apparently, a temporal parsing limit determines whether interocular conflict results in binocular rivalry or dichoptic masking, not direct low-level inhibitory interactions. Consistent with this conclusion is our recent finding (Chapter V) that monocular rivalry, which is necessarily pattern-based, shows the same parsing limit as binocular rivalry.

The temporal parsing limit to rivalry was shown to remain unchanged even when patterns were swapped between the eyes after every presentation (experiment 3). This finding complements the findings of experiment 2 by showing that binding of stimulus information happens irrespective of the eye of origin of the stimulation.

Luminance integration, as a low-level visual process, might be responsible for our results. However, we showed that contrast polarity reversals, which should have weakened stimulus strength and therefore rivalry if luminance integration occurred, had no effect on the temporal limit (experiment 4). Nor did higher stimulus energies lead to longer parsing limits (experiment 5). Luminance integration may therefore be excluded as a possible mechanism underlying temporal parsing. With the temporal parsing limit not dependent on low-level stimulus characteristics, we suggest that pattern information may be of central importance. Such a proposal is bolstered by studies showing that the visual system is especially sensitive to pattern information changes at low temporal frequencies $<\sim$4 Hz, and less so a higher temporal frequencies [140, 135].

If pattern information indeed determines the temporal limit, then the parsing limit is probably a reflection of the temporal grain of the visuals system's ventral form pathway, which is important in form processing [99]. This hypothesis was tested in a final experiment, in which we investigated whether the temporal limit is determined by form (i.e. orientation) information or motion information, aspects processed mainly in the ventral and dorsal pathway respectively. We constructed stimuli that contained only conflicting interocular motion signals, or also additional form conflict. We found that the $\sim$350 ms parsing limit still occurs when form-based conflict was present, but not when motion-only conflict existed. Our findings additionally showed that the limit is not dependent on certain flicker frequencies per se, consistent with the finding that purely temporal patterns, i.e. unpatterned flashes, do not engage in rivalry [207]. These findings show that the temporal parsing limit is not a feature of the motion pathway, and may indeed be a characteristic of the ventral form pathway. Although our findings cannot settle the discussion on whether motion information alone can cause rivalry [59, 110, 8], they do show at least that the temporal parsing limit is not dependent on motion but on form information, and that the rivalry we observe is therefore form-based and not motion-based. Motion conflict, however, did lead to rivalry in our experiments, but only when presentation was continuous, showing that at the very least it has different temporal characteristics from form-based rivalry.

Apart from showing that form conflict is a key determinant of rivalry, our results indicates that motion processing is subservient to form processing. That is, when form rivalry occurs, the motion percepts are bound to the perceived orientation, and eye-based segmentation of motion signals occurs. In the case where the orientations were matched and form-based rivalry was absent, the motion information from both eyes is integrated. Previous studies have shown the governing power of form on motion processing, showing that the perception of a unitary form allows for the global spatial integration of local motion information [164, 248, 220]. Our experiments extend these findings by showing that form information dictates the interocular integration or segmentation of motion. Although sometimes this link between form and motion processing [16, 17], or flicker[58], may be broken this break was not evident with our stimuli.

At what level does rivalry take place with our stimuli? In the binocular rivalry literature, which is based on studies using continuously present stimuli, there has been a lively debate about what mechanisms maintain rivalry, and at what level rivalry takes place. Some have argued that rivalry takes place early in the visual cortex (e.g. Refs. [34, 216, 252]), and that it depends primarily on low-level stimulus characteristics. Meanwhile, others have maintained that rivalry takes place in higher visual areas [159, 149, 150], involving stimulus-based rivalry. Recently it has been proposed that different levels may contribute differently or additively to the process of rivalry [38, 86, 284].

Although the results from experiment 1-5 do not support an inter-ocular account of rivalry [34, 252], at least not for temporally modulated stimuli, they are consistent with studies that implicate early visual areas in the maintenance of rivalry [216, 252], as some of the found independences, e.g. eye-of-origin information, contrast-polarity, and pattern-selectiveness of adaptation, are found in complex cells in the primary visual cortex [27, 119, 95, 194].

Nevertheless, the finding of a temporal limit of around 350 ms shows a rather coarse temporal resolution at the stage where rivalry occurs. Since it is an order of magnitude greater than the critical flicker fusion frequency [100], the 350-ms limit seems rather long for early visual areas. Some have argued that rivalry occurs mainly in the parvocellular (P) system [110], because the P system shows a sustained activity compared to the magnocellular (M) system's transient response [110]. However, the early visual P cells in the retina, lateral geniculate nucleus, and V1 are not nearly as sustained as necessary to bridge blank periods exceeding 300 ms [125, 104]. Therefore the P pathway, at least at early visual levels, does not seem to be responsible for the 350 ms temporal parsing limit. Consequently, at least part of the involved network must be situated at levels higher than V1 (complex) cells [38, 86]. A possible candidate area is the lateral-occipital cortex (LOC) in the ventral form pathway, which is important in shape and object processing [99]. Moreover, the human LOC, but not lower visual areas, shows prolonged activity up until about 400ms after brief stimulation [137], and it is modulated according to the percept during binocular rivalry [191]. These characteristics are consistent with the temporal parsing limit to rivalry described in this report.

What might be the use of such coarse temporal parsing in form processing? We suggest it may be related to the frequency of saccadic eye movements, which occur around 3 times per second. On the one hand, this temporal constraint could mean that binocular conflict needs to be solved rapidly within each fixation, but on the other, the perceptual interpretation would ideally be stable over several fixations as well, in the interests of temporally stable perception. It is within this context of 'temporal coherence' that our finding of a form-based parsing mechanism can be interpreted. We have found that at the time scale of about 350ms binocular conflict resolution is greatly influenced by temporal long-range binding of pattern information that provides a stable perceptual interpretation spanning several stimulus presentation cycles. The binding of pattern information makes a perfect candidate for a mechanism that ensures we perceive a stable world in the face of fixation changes and eye-blinks. The finding that stereo-information as well can be retained over blank periods of about 300 ms [73, 147, 94], suggests that the observed form-based limit is not specific to binocular rivalry but likely extends to binocular fusion processes.

The functioning of this system would be aided if similar orientations are fixated before and after saccades, which seems to occur quite naturally, as saccades are generally small [26, 272] and neighboring patches in visual scenes have very similar orientation content [92]. As a final indication, our data suggest that this mechanism is part of the visual "what" or form pathway, which dovetails well with the finding that during saccades "where" information (from the dorsal pathway) is suppressed whereas "what" information (from the ventral pathway) is retained [123], preserving form information over saccades.

In conclusion we have revealed a pattern-based temporal parsing mechanism, possibly positioned in the ventral form-pathway, that is an essential and heretofore unrecognized mechanism in the formation of visual awareness during the resolution of binocular visual conflict when processing dynamic, inherently-ambiguous, visual scenes.

IV.4 Methods

Apparatus Images were presented on a 22" LaCie electron22blueIV monitor (1600×1200 pix, refreshed at 75 Hz). Subjects were seated at 46 cm from the screen, using a chin-rest to stabilize head position. Experimental procedures were reviewed and approved by the Institutional Review Board.

Stimuli Stimuli in experiments 1-5 were orthogonally oriented (45 deg from vertical) gratings, having a spatial frequency of 0.87 deg/cycle. The stimulus was seen through a circular Gaussian window (sigma = 0.76 deg, cut-off at a diameter of 4.5 deg). Michelson contrast was 1 at the stimulus' center. Background luminance was 20.0 cd/m^2, maximum and minimum luminance (i.e. white and black parts of the grating), were 71.9 cd/m^2, and $\sim$0.0 cd/m^2, respectively. The polarity inversion experiment was performed on a gamma-corrected monitor (background: 6.4 cd/m^2; maximum luminance: 12.8 cd/m^2; minimum luminance: $\sim$0.0 cd/m^2). The stimulus was surrounded by a binocular annulus that was divided into 20 equally-sized parts alternatively made of full contrast (white) and zero contrast (grey). The annulus was 0.1 degrees wide, with a radius of 2.46 deg, and served as a fusion aid. Temporal characteristics of the stimuli are described in the main text and in Figure IV.1. Trials lasted for 60 seconds in experiments 1, 2, and 6; 30 seconds in experiments 3 and 4; and 6 seconds in experiment 5. Experiments 3, 4 and 6 had a small fixation mark. In experiments 1-4, individual events (T_{on}) lasted for 4 frames (∼53 ms), followed by a variable blank period (T_{off}). In experiment 5, T_{on} and T_{off} were both varied. In Experiment 6, stimuli consisted of the sum of three gratings (0.11, 0.44 and 0.87 cycles/deg, the middle frequency was offset by half a cycle). The stimuli were seen through a Gaussian annular window that had a radius of 1.33 degrees and a sigma of 0.34 degrees, which left the surroundings of the fixation mark empty. Stimulus motion was introduced by phase-shifting each grating at every frame by 1/32 cycle. Stimuli were displayed for 8 frames (about 100ms), before being blanked for 0 ms, ∼113, and ∼540 ms, after which the motion sequence resumed where it had left off. This cycle was repeated until the end of a trial, leading to, respectively, continuous presentations, repetition periods of about 213 ms, and 640 ms. Although luminance profiles of the eyes' patterns differed in the motion-conflict-only condition, this does not necessarily lead to rivalry [58], and luminance differences do not prevent patterns from being processed as identical forms when orientations are the same (as shown in experiment 4). The motion-and-form conflict condition was identical to the motion-conflict-only condition, except for a 30 degrees orientation conflict between patterns. The orientations were (near) horizontal to prevent binocular fusion and a resulting depth perception. Subjects indicated the direction of motion, ignoring form rivalry when it occurred without motion rivalry (which did actually not occur).

Procedure In experiments 1-4 and 6, subjects (n=4) presses either of two keys to indicate their dominant percept, and were asked not to press when a fast and regular switching of the stimuli was perceived, or when the two patterns were overlaid without any of the two being stronger. In experiment 5, subject categorized each trial as containing rivalry, or rapid and regular orientation switches. Subjects (n=5) based their categorization on the last 3 seconds of the trial.

The psychophysical data were fitted with: $f(x) = a/(1 + e^{(x-t)/w})$, where a was set to the maximum fraction of "cumulative percept duration/trial duration" per experiment (except in experiment 2, where a was 0.5), w is the width of the curve, and t is the temporal limit, which we report on in this study.

Acknowledgements We would like to thank Tomas Knapen, and Allard Kamphuisen for critical discussions.

Chapter V

Visual rivalry: a balance between binocular ambiguity resolution and monocular interactions

Jeroen J.A. van Boxtel
Tomas Knapen
Casper J. Erkelens
Raymond van Ee

An essential process in the creation of stable visual awareness is the resolution of ambiguities in the input. Generally, a single interpretation prevails. Occasionally, however, reinterpretations repeatedly reach awareness: visual rivalry. We studied three modes of rivalry: monocular, stimulus and binocular rivalry. Contrasting with the general view, we report that a limiting process for inducing visual rivalry is monocular interaction of visual information, which suggests that monocular processes prevent rather than cause rivalry. Importantly, when monocular interaction is bypassed, the differences in perceptual dynamics disappear between the investigated modes of rivalry. These findings constitute evidence for a single dynamical system underlying different modes of visual rivalry, and ambiguity resolution. Our study demonstrates that the dynamics of visual awareness follows from the balance between the volatile nature of a single ambiguity-solving process, which provides the visual system with flexibility, and the stabilizing nature of a monocular interaction process, which prevents frequent reinterpretations.

V.1 Introduction

What shapes our awareness in face of the ambiguities intrinsic in visual information? The mechanisms underlying visual awareness are often studied employing ambiguous stimuli. The merit of such stimuli is that they remain constant, while their multiple perceptual interpretations alternate, revealing the brain's internal processes involved in resolving visual ambiguities.

Interestingly, the alternation of perceptual interpretations, called visual rivalry, may result from different stimulus presentation protocols: i.e. monocular rivalry, binocular rivalry, and stimulus rivalry. Monocular rivalry may be instigated when two incompatible patterns are presented to the same eye. Because the two patterns are presented to the same eye, the rivalry involves pattern-based, not inter-ocular, competition. Binocular rivalry is instigated when two incompatible images are presented to different eyes. Binocular rivalry may therefore involve both inter-ocular and pattern-based competition. During stimulus rivalry, the incompatible images are presented to different eyes, but swapped between the eyes at regular intervals. Competition may therefore involve both inter-ocular and pattern-based processes, as well. For ease of discussion, the visual rivalries resulting from those different stimuli are called different modes of rivalry in the remainder of this chapter, even though they may rely on party shared mechanisms.

Some reports indicate that these different modes of rivalry may be related to each other. For example, a similar amount of orientation conflict instigates rivalry for monocular and binocular rivalry [159], and a pattern simultaneously engaged in different modes of rivalry, alternates in near unison across its spatial extent [213, 19], in disregard of the different protocols with which rivalry is instigated. Moreover, when changing the saliency of one stimulus interpretation in ambiguous stimuli, the percept durations of the unaltered stimulus are changed, leaving dominance durations for the altered stimulus unaffected [152], an effect first established for binocular rivalry [153].

On the other hand, large quantitative differences exist as well among the different modes of rivalry. For instance, the different modes of rivalry seem to depend on different temporal and spatial stimulus characteristics. While binocular rivalry is not strictly bounded to a specific spatial or temporal regime and is easily instigated, the occurrence of monocular rivalry is reported to be bound to low spatial frequencies, and low contrast [158, 173, 264], and the occurrence of stimulus rivalry is reported to be bound to high spatial frequencies and low contrast [143, 38, 159]. Moreover, the different modes of rivalry also differ in alternation dynamics [264, 251]. Binocular rivalry has short (∼2 sec) dominance durations, with relatively short periods spent in transitions from one percept to the next. This behaviour is mimicked by stimulus rivalry [159], but only for certain parameter combinations [143]. Monocular rivalry, on the other hand, shows slow alternation dynamics, and long periods spent in transitions periods. Interestingly, however, faster dynamics may be induced for monocular rivalry when the competing patterns are not presented simultaneously but subsequently [116, 247], suggesting that even this apparent difference in perceptual dynamics across the modes of rivalry is a difference in degree rather than kind.

These similarities and differences in spatial and temporal stimulus characteristics for both the instigation of rivalry, and perceptual behaviour, have long been a centrepiece in the debate as to whether the different modes of rivalry depend on different underlying rivalry processes [143, 38, 251, 144, 15, 47] or a shared rivalry process [19, 149, 138]. The current view is that visual rivalry is a complex, multileveled, and multifaceted process [38], leaving these fundamental differences largely unexplained.

Here, we employed simple grating stimuli that exclude shared higher-level influences, such as context [18], and therefore engage solely the core rivalry resolution processes. We varied the temporal layout of stimulation, because temporal stimulus parameters are important, but little-studied, specifiers for stimulus rivalry [159, 284, 143]. We demonstrate that for the three rivalry-protocols, rivalry will occur when the individual conflicting patterns are repeated within a binocular temporal window of about 350ms, provided that a monocular 50-ms interaction window is bypassed. When these requisites are satisfied, the restrictions on spatial parameters for inducing rivalry become less tight, and perceptual dynamics are very similar among for the different modes of rivalry.

V.2 Results

In a first experiment we investigated the dependence of rivalry occurrence on the inter-swap period (the period between two successive stimulus swaps between the eyes; Figure V.1), and the temporal duty cycle (the presentation time of a stimulus during an ISP; Figure V.1). Stimulus rivalry was induced employing the conventional stimulus rivalry protocol [159], including the 19 Hz flicker, with the addition of a blank period of variable length before the stimulus swap between the eyes. To allow for a fair comparison between rivalry modes, a similar procedure was used for monocular rivalry (see Fig V.1b), but the stimulus was shown either to one eye (Figures V.2, V.3, V.4) or to both eyes in phase (see Fig. V.1b, results in Figure V.5). The grating stimuli had a spatial frequency (2 cyc/deg), stimulus size (1.7 deg) and contrast (50% for most experiments, 100% in Figure V.5) that are thought to compromise monocular [173, 264] as well as stimulus rivalry [143, 47]. Subjects categorized 6-second trials as containing either a plaid, switches, or rivalry [143]. Employing this paradigm, we capitalized on the distinct perceptual experiences [143, 159] to deduce whether stimulus-based percepts (i.e. rivalry) or eye-based percepts (rapid switches) had occurred.

Rivalry percepts in this experiment were very common. At the optimal parameter combinations in this experiment, rivalry was reported in 63% of the trials

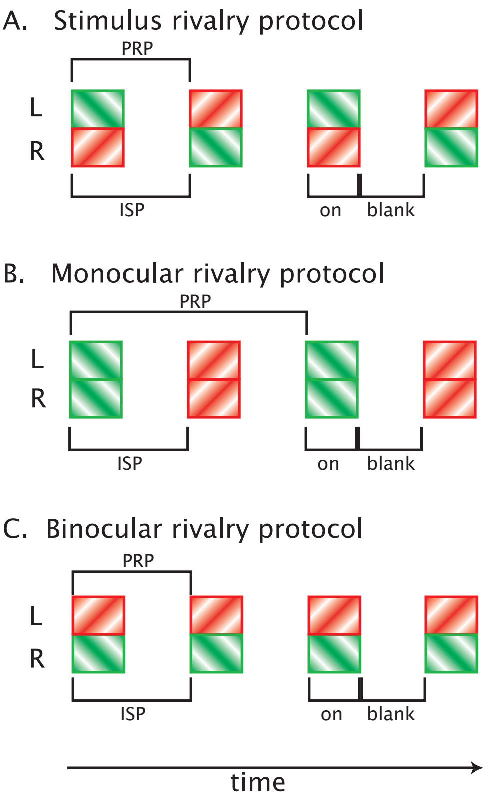


Figure V.1 | Temporal characteristics of the stimulus protocols evoking (a) stimulus, (b) monocular, and (c) binocular rivalry. *To produce stimulus rivalry and monocular rivalry the patterns were swapped between or within the eyes, respectively, at each successive cycle. In Exp. 1, the patterns were additionally flickered at 19Hz during the 'on'-period. L and R, denote left and right eye, respectively. ISP denotes the inter-swap period; PRP denotes the pattern repetition period for identical patterns irrespective of eye of origin. ISP and PRP, together with the temporal duty cycle [on/(on+blank)], define the temporal aspects of stimulation. All stimuli were achromatic.*

for monocular rivalry, and in 85% of the trials for stimulus rivalry. (Exp. 1, Figure V.2). Previously, a near absence of rivalry was reported at the used contrast and spatial frequency [143]. Clearly, the employed temporal parameters have allowed us to bypass the restrictions on contrast and spatial frequency, which were seen as objections against a unified rivalry mechanism.

The results are largely independent of the temporal duty cycle, apart from an apparent decrease of rivalry reports near a duty cycle of 1. This decrease in rivalry reports is consistent with the literature that reported a near absence of rivalry with stimuli with a duty cycle of 1 [143]. This observation suggests that a blank period just before the stimuli are swapped between the eyes will increase rivalry reports. We tested this prediction in a subsequent experiment, looking specifically at spatial frequency and ISP. The results (Figure V.3) show that, indeed, the introduction of a 67-ms blanks before the stimulus-swap enlarges the parameter range for which rivalry is reported as compared to the conventional 19Hz flicker paradigm (Figure V.3), even though the effective contrast was higher for most of the blank-conditions. Furthermore, the occurrence of rivalry with our blanking paradigm was much less dependent on spatial frequency.

Additionally, these experiment showed that stimulus rivalry took place when the ISP was below about 460 ms (red line, Figure V.2b; see methods). For monocular rivalry this limit was about halved to ISPs of 190 ms (red line in Figure V.2a). Strikingly, both these limits translate into a ∼400 ms pattern repetition period, that is, the period in between two successive presentations of an identical stimulus is about 400 ms. (Note, that a monocular 200 ms ISP equals an pattern repetition period of 400 ms, Figure V.1b). A very similar limit at 350 ms has been found for binocular rivalry (Ref. [209] and Chapter IV), which has been shown to be a *pattern-integration limit* (Chapter IV). These results demonstrate that monocular rivalry and stimulus rivalry occur for a similar wide range of spatial and temporal stimulus parameters (ISP<∼350 ms) as binocular rivalry, provided that non-continuous presentations are applied. To examine what caused rivalry to be so common with our blanking paradigm, we determined the dependence of rivalry, plaid and switch reports on both the moment, and the duration of the blank. The ISP was fixed to 260 ms, the 19 Hz flicker was removed, and the temporal duty cycle was varied by inserting blanks just before a pattern swap (top row, Figure V.4a), or right in between two pattern swaps (bottom row, Figure V.4a). Importantly, many of these temporally modulated stimuli ought not to curtail the inter-ocular rivalry mechanism [284, 34], because they meet the requirements for strong interocular inhibitory interactions [143, 284, 288]. That is, they have short blank durations (∼50-80 ms), long 'on' durations (∼210-180 ms), and no 19 Hz flicker. Consequently, we anticipated many 'switch' (i.e. eye-based) reports. However, we found that rivalry reports predominated (Figure V.4b), indicating stimulus-based and not eye-based percepts. The results furthermore suggest that intra-ocular (i.e. monocular) interactions may be important. This suggestion is based on the observation that conditions with blanks in the middle (Figure V.4b, right panels)—which facilitate monocular interactions as consecutive conflicting patterns are temporally abutting—show a large reduction in rivalry reports and an increase in plaid percepts, as compared to stimuli with blanks at the end of the ISP (left panels). Furthermore, consistent with the argument that a monocular interaction process lies at the origin of this effect, purely monocular stimuli led to almost equal amounts of plaid reports as binocular stimuli (Figure V.4c). These data suggest that monocular interactions may prevent rivalry in the stimulus rivalry and monocular rivalry protocol, and conversely that preventing monocular interactions increases rivalry occurrence.

Our experiments so far have determined that monocular, stimulus and binocular rivalry are induced over the same temporal parameter range (all exhibited a pattern-repetition limit of about 350 ms). Furthermore, the restrictions on temporal characteristics (i.e. the 19 Hz flicker), and spatial characteristics (i.e. contrast and spatial frequency) may be lifted when monocular patterns are temporally non-abutting, i.e. when monocular interactions are prevented.

Would this striking similarity between the different modes of rivalry also be found for the perceptual dynamics when monocular interactions are prevented?

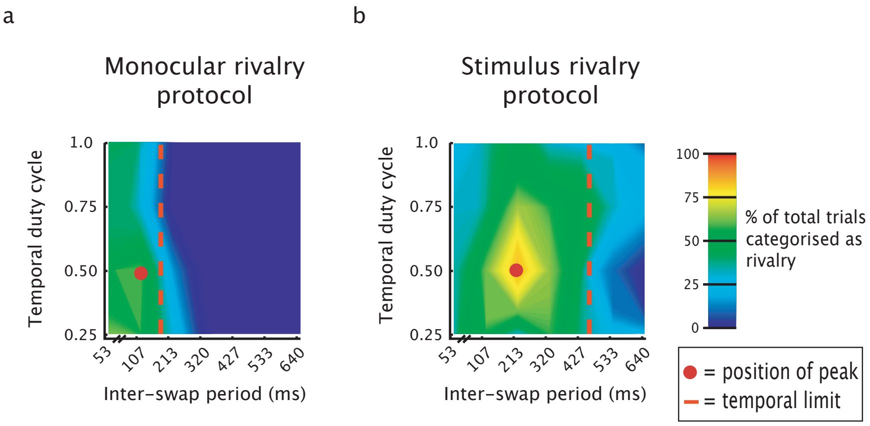


Figure V.2 | Dependence of rivalry reports on the temporal duty cycle and the inter-swap period for stimuli with 2.1 cycles/deg, and 50% contrast. *In both monocular* (a) *and stimulus* (b) *rivalry protocols, a vast proportion of the trials were categorized as producing rivalry (at maximum: 63% in monocular rivalry, 85% in stimulus rivalry), indicating that the temporal characteristics of stimulation lifted the restrictions to rivalry induction imposed by spatial frequency and contrast. The occurrence of rivalry reports is largely independent of the duty cycle, in stead the occurrence seems to be restricted to ISPs of ∼200 ms in monocular rivalry and ∼460 ms in stimulus rivalry (red dashed line), both close to PRPs of ∼400 ms (see Chapter IV), indicating a shared mechanism.*

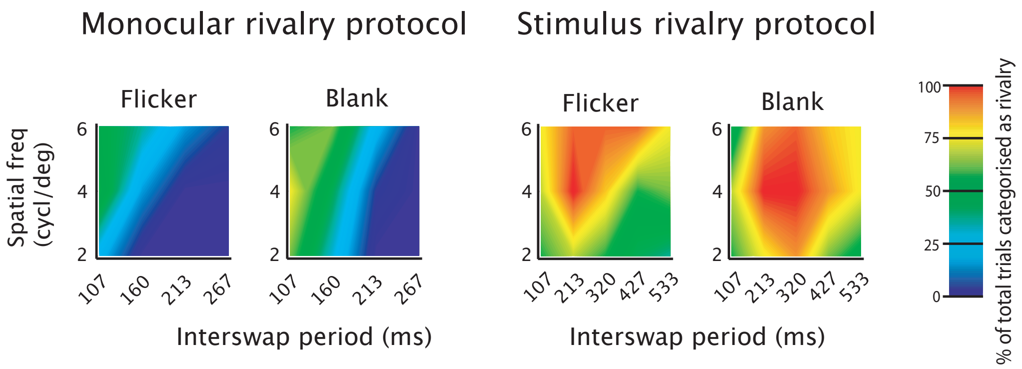


Figure V.3 | Dependence of monocular and stimulus rivalry reports on interswap period (ISP) and spatial frequency. *For conventional monocular and stimulus rivalry, the stimuli were flickered at 19 Hz, while in our blanked stimuli the stimulus was presented continuously except for a 67ms blank before each swap. In both monocular and stimulus rivalry protocols, rivalry was more frequently reported with the blanked stimuli than with conventional stimuli. The incidence of rivalry reports was also less dependent on spatial frequency with the blanked stimuli than with the conventional stimuli. The results show that restrictions on spatial frequency and temporal parameters (19Hz flicker and ISP) for the instigation of rivalry largely vanish when a sufficiently large blank is introduced before a stimulus swap (i.e. monocular images change). Note that the blanked stimuli had an effective contrast that was higher than for the conventional rivalry protocol, for all but the 107 ms ISPs.*

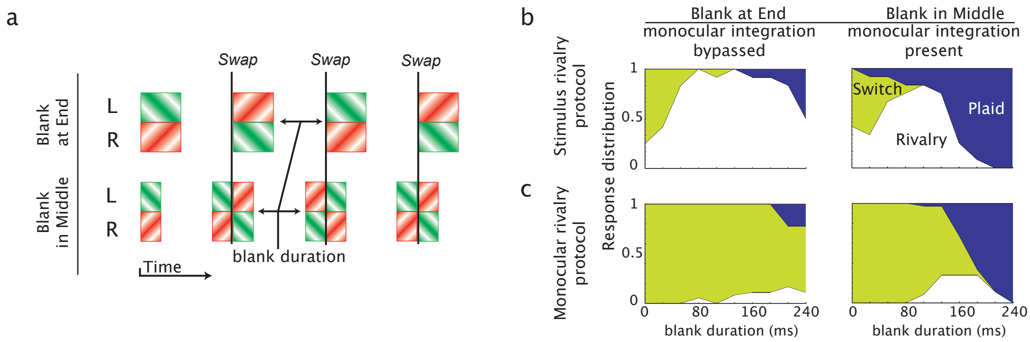


Figure V.4 | Blank duration before a pattern swap determines the occurrence of rivalry. (a) *Stimuli (inter-swap period (ISP): 260 ms) were constructed with a blank just before a swap of the patterns* (top), *or in the middle of the inter-swap period* (bottom). *Note that with identical blank duration the stimuli had identical temporal duty cycles (and did not flicker at 19Hz). The two patterns were temporally separate when the blank was placed at the end of the ISP, however they were temporally abutting when the blank was placed in the middle of the ISP. It is key to note that the latter condition allowed for monocular interactions. Note that the stimuli were achromatic, the color in this image is added for illustrational purposes* (b) *Responses during stimulus rivalry conditions, presented as stacked frequency distributions. Switch = green, rivalry = white, plaid = blue. Rivalry occurred frequently over nearly the full range of blank durations when the blanks were placed at the end (left panel), and was severely decreased when the blanks were placed in the middle (right panel). Plaid reports, however, gained in frequency, suggesting, indeed, that monocular interactions took place between the two patterns.* (c) *Consistent with Exp. 1, rivalry reports at this ISP were rare during monocular rivalry, and switch reports abound. Importantly, when blanks were placed at in the middle of the ISP, plaid reports were almost as frequent as with stimulus rivalry conditions (note the similarity in plaid reports between both the left panels and the right panels in* b *and* c*), underscoring that plaids are mainly the result of monocular interactions. This interaction process may be bypassed by placing blanks at the end of the ISP (note the similarity in plaid reports between the two left and the two right panels in* b *and* c*).*

During 60-second trials, observers were asked to track the dominance of both patterns. Rivalry dynamics were analyzed in terms of average percept durations (defined as the time elapsed between two button presses [therefore including transition periods]), the cumulative amount of exclusive percept duration (a measure that excludes transition periods, the marker of weak alternations), percept duration distributions, and time-until-first-percept. Importantly, we either included or excluded the monocular interaction process by varying the ISP (and therefore blank duration) while keeping the on′ period at ~13 ms (1 frame). With very short blank durations, monocular interaction should produce a plaid in both monocular and stimulus rivalry protocols, and consequently, we expect perceptual behaviour corresponding to classically reported monocular rivalry behaviour. With longer blank durations the perceptual dynamics are expected to become more similar to binocular rivalry, because the monocular interaction (which, by construction, is absent during binocular rivalry) between the conflicting stimuli is eliminated. Note that similar rivalry behaviour among the three modes of rivalry may only be expected up to ISPs of ∼175ms, because this period translates into the 350ms pattern-integration limit in the monocular rivalry protocol (see experiment 1).

Consistent with the classical behaviour of monocular and binocular rivalry [251, 264] we found (Figure V.5A, B) that at near-continuous presentations (ISP = 13 ms) binocular rivalry shows fast and strong alternations (i.e. short percept durations, and small amounts of transition periods), whereas monocular rivalry shows slow and weak alternations. At longer ISPs the difference between the different modes of rivalry disappear. At ISPs between 40 ms to 90 ms, perceptual alternations of all modes of rivalry were strong (Figure V.5b), and fast (Figure V.5a). In fact, up to ∼150ms perceptual dynamics are rather similar (see Figure V.5a). With ISPs between 40 ms to 90 ms, dominance was generally complete, and dominance durations distribution were not significantly different between the different modes of rivalry (Figure V.5c). Finally, the time until the first reported percept showed trends similar to those of the average percept durations (Figure V.5d). From our finding that the perceptual dynamics for all rivalry manifestations become identical at ISPs above ∼40-70 ms, we estimate the monocular interaction window to be ∼50 ms, consistent with integration periods of early visual neurons [108].

Note that the range of nearly identical perceptual dynamics encompassed a large part of the range of ISPs that allows for rivalry without the interference of monocular interactions (e.g. the range 40 ms until ∼175 ms, which are bounds imposed on monocular rivalry by monocular interactions and the pattern-integration limit, respectively).

V.3 Discussion

Monocular, binocular, and stimulus rivalry have been reported to show several differences and similarities, leading some researchers to suggest that they depend largely on different systems and others to suggest that they are identical.

Our data show that when monocular interactions between competing stimuli is prevented, rivalry occurrence is increased. In fact, the rivalry occurrence and the perceptual dynamics in those cases are not different among the different modes of rivalry. In the entire temporal regime that allows for binocular rivalry, that is for pattern repetition periods $<\sim$350ms (Ref. [209], and Chapter IV), monocular and stimulus rivalry could be induced, even for spatial frequencies that were previously thought to prohibit these modes of rivalry (Figure V.2, and V.3). In these regions, the perceptual dynamics of the different modes of rivalry did not differ either (Figure V.5). Our findings weaken the reasons to assume different underlying mechanisms for the different rivalry modes, and instead suggest a common origin. Since monocular rivalry solely involves pattern-based rivalry, the shared mechanism of rivalry is most likely pattern-based.

The perceptual dynamics and ease with which rivalry may be instigated were found to differ greatly between the different modes of rivalry when blank periods were small ($<\sim$50ms). These effects are explained by assuming that monocular interactions take place, which we may envision as a sliding temporal window that temporally integrates neural responses over a 50-ms window. The integration of the two orthogonal achromatic gratings used in our study leads to formation of a plaid with a blurred checkerboard appearance. This stimulus contains few segmentation cues and will therefore only sporadically lead to suppression of one of the component gratings, as shown by our data. Therefore monocular interactions may explain both the slowness and weakness of rivalry for monocular and stimulus rivalry when monocular blanks are $<$50 ms. With the introduction of blanks, the monocular interactions were bypassed, allowing the gratings to compete with each other at a pattern-based level.

Eye-based suppression and inter-ocular rivalry

In the explanation outlined above, a shared pattern-based rivalry process explains the similarities between the different modes of rivalry, while the presence or absence of monocular interactions explains the differences. So far, no reference has been made to the possible influence of inter-ocular rivalry between largely monocular neurons [38]. Inter-ocular rivalry may trump pattern-based rivalry in importance for the specific condition of continuously presented stimuli. The literature provides several lines of evidence supporting inter-ocular rivalry. In the following discussion of these results, we would like to make a distinction between *eye-based suppression*, that is, the reduced sensitivity of monocular neurons to certain stimulus characteristics, and *inter-ocular rivalry*, that is, the direct competitive or inhibitory interactions between pools of monocular neurons that code for different stimulus characteristics. We will argue that, the reported findings are actually support for eye-based suppression and not necessarily for inter-ocular rivalry, which means that the omission of an inter-ocular rivalry stage in our explanation of our results is not in conflict with the existing literature.

Reported evidence for inter-ocular rivalry comes from neuroimaging studies [252, 109, 292, 216] that have reported that early visual areas (or even specifically monocular areas) show a modulation in blood-oxygen-level dependent (BOLD) response tied to the percept. However, because of a low temporal resolution, the data cannot exclude the possibility that the observed waxing and waning of the BOLD response results from feedback or lateral input from binocular neurons. These data are therefore no definite proof of inter-ocular rivalry. In a second line of evidence, it has been shown that a probe stimulus presented to the eye receiving the non-dominant stimulus during rivalry is harder to detect than a probe presented to the eye receiving the dominant stimulus (e.g. [268]). Although clearly

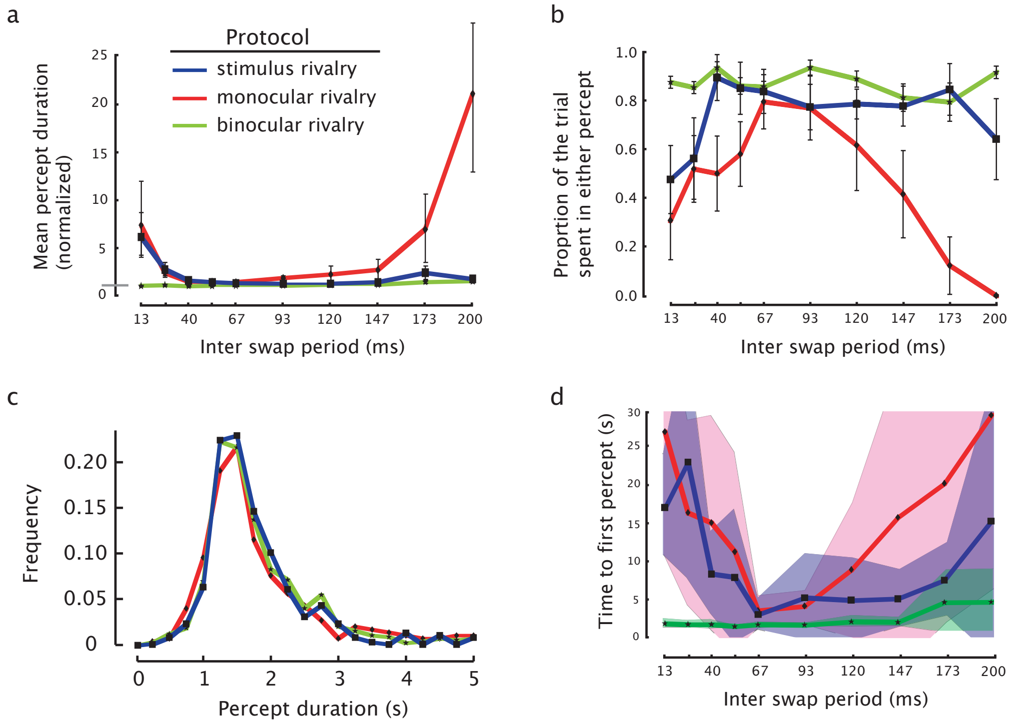


Figure V.5 | Perceptual dynamics of binocular (green), stimulus (blue), and monocular (red) rivalry protocols. Top Panels. *Rapidity and strength of perceptual alternations were probed with the normalized average percept duration* (a) *(see methods), and the total proportion of the trial spent in either percept* (b)*, respectively. At small ISPs, monocular and stimulus rivalry alternations are slow and weak, and binocular rivalry alternations fast and strong. The behaviour of monocular and stimulus rivalry changes with increasing ISP (and therefore blank duration), reaching perceptual dynamics not significantly different from the dynamics of binocular rivalry at ISPs of* $\sim$*40-90 ms, when the monocular interaction process is bypassed. At pattern repetition periods* $>\sim$*150-200 ms monocular rivalry is impeded, consistent with it reaching the pattern integration limit at a PRP of* $\sim$*350 ms (Figure V.1). Error bars are 95% bootstrapped confidence intervals (left) and s.e.m. (right).* Bottom Panels. (c) *Frequency distributions of percept durations shorter than 5 seconds. The distributions are not significantly different form each other (all pairwise Kolmogorov-Smirnov tests:* $p > 0.28$*), nor were the average durations (which were log-transformed to approach normal distributions; all three pair-wise t-tests:* $p > 0.35$*). Data were combined for ISPs of 53, 67, and 93 ms.* (d) *Time until first-reported percept (Tp) depends on the ISP in the same way as the average percept duration does. For small ISPs the Tp is large for both monocular and stimulus rivalry, it decreases for larger ISPs and, for even larger ISPs, it increases again, especially for monocular rivalry, although binocular and stimulus rivalry also show an increase. At ISPs around 70-90ms the differences between the conditions are small. The shaded areas encompass the mean ± 1 standard deviation, based on 8 data points. These combined data in this plot strongly suggest that a single rivalry mechanism underlies the perceptual dynamics for all three modes of rivalry.*

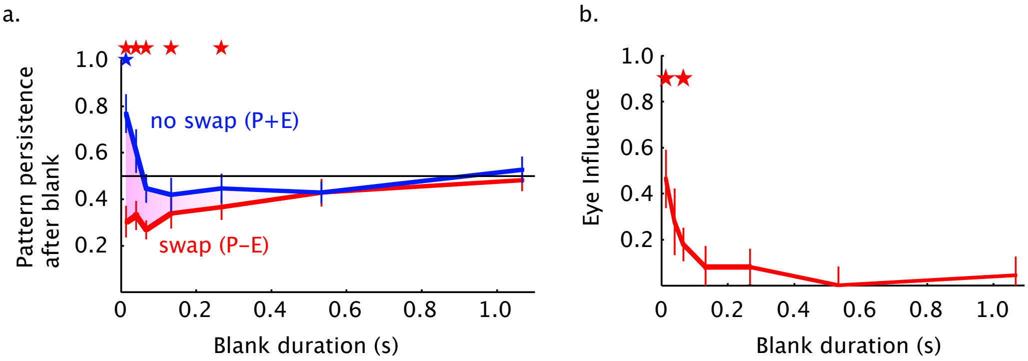


Figure V.6 | The influence of eye-based processes on the persistence of pattern percepts in binocular rivalry after a stimulus swap between the eyes. (a) *The percentage of trials that resulted in a perceptual report consistent with the previously dominant pattern (i.e. pattern-persistence) when the stimuli were swapped between the eyes (red line), or not (blue line), after perceptual dominance was established. Significant differences (t-test) from 0.5 are indicated by stars. Both curves combine eye-based and pattern-based influences. In the non-swap trials, both eye-based and pattern-based processes add to the persistence of a certain pattern percept (indicated by the letters P+E (i.e. Pattern+Eye)). In the swap trials, only pattern-based processes contribute to the persistence of a pattern percept, while eye-based processes inhibit pattern-persistence (indicated by the letters P-E), because they try to keep the percept within the same eye (and as a consequence they favor a switch in pattern). Therefore, the difference between the two curves (i.e. (P+E)-(P-E)), is proportional to the influence of eye-based processes on pattern-persistence. This measure is plotted in* (b). *Significant influences of eye-based processes are only found for blank durations <~70 ms. This finding provides further evidence for the proposal that monocular interactions prevent pattern-based rivalry, but only when conflicting patterns are in temporal proximity. Data points are means ± between-subject s.e..*

suggestive of *eye-based suppression*, these data do not prove that this suppression results from direct *inter-ocular interactions*. Furthermore, these suppressive effects grow with increased stimulus complexity [198], suggesting an important role for feedback from pattern-based processing stages. A third line of evidence showed that when one swaps the eyes' images during a dominance phase, an immediate switch is perceived to the new stimulus in the dominant eye [42]. Although these data suggest that one eye remains suppressed, while the other dominates, even when patterns change eye, there is no proof of the involvement of direct inter-ocular interactions. Indeed, monocular *interactions* can also explain the effect, as we found that two successive monocular conflicting patterns are both perceived when not separated by a blank period (i.e. switch percepts in our experiments). Furthermore, assuming a monocular interaction process predicts that the effect dissipates when a blank duration exceeding the interaction window (~50 ms) is introduced at the time of the swap. We tested this prediction in an additional experiment. Employing the previously reported paradigm [42], we show that, indeed, introducing a blank of sufficient length (here about 70ms; Figure V.6) removes the persistent dominance of the eye that held the previously dominant stimulus (Figure V.6; see methods and figure caption for details). A fourth line of evidence comes from computational studies that have shown that the inclusion of an inter-ocular rivalry stage, followed by a pattern-based rivalry stage ([284], or stages [86]) can reproduce several key findings related to stimulus and binocular rivalry. However, in our view, these models function as well without an inter-ocular inhibition stage, indicating that inter-ocular rivalry may not be essential. Overall, this discussion shows that although inter-ocular rivalry may exist, the existing data do not unequivocally support an inter-ocular rivalry account.

The neural underpinnings of rivalry

The above overview shows that inter-ocular rivalry may not be necessary to explain the current literature. In fact, this chapter and previous investigations provide indications, yet no definite proof, that inter-ocular rivalry is not a major contributor to rivalry. We briefly review the key findings here.

In our experiment 3 (Figure V.4), we found that pattern-based rivalry persisted, even though the individual pattern presentations lasted over 200 ms, which is presumably enough for strong inter-ocular rivalry to operate [284, 288]. Furthermore, we found that monocular rivalry and stimulus rivalry show very similar behaviour in the region with short blank durations (i.e. nearly continuous conditions; Figure V.5). This is unexpected if one assume that inter-ocular rivalry takes place, as stimulus rivalry should allow for inter-ocular inhibition, in addition to pattern-based inhibition, while monocular rivalry should not. Both findings indicate that inter-ocular rivalry was not operational during our experiments. Furthermore, single cell studies have shown that monocular cell activity is not modulated by the presentation of a conflicting pattern in the other eye but binocular cell activity is [21], and monocular cells do not modulate their activity according to the percept, while binocular cells do [149]. These findings indicate that binocular cells, not monocular cells, are involved in rivalry. A recent study [282] expanded on these results by measuring multi-unit activity (MUA), a measure of the output of a neuronal population, and local field potentials (LFPs), an indirect measure of the input into, and local processing within, a neuronal population [162, 163]. It was shown that MUA in V4 correlated with the percept, while the MUA in V1, which contains many monocular neurons, did not. LFPs on the other hand were correlated with the percept in both V1 and V4, but at a delay relative to the MUA in V4. The finding of a delayed correlation to perception specific to the input in V1 (i.e. the LFPs), suggests that the correlation found in V1 is due to input from high visual areas (i.e. feedback). Interestingly, it has been shown that the BOLD response from functional imaging correlates significantly better with LFPs than with other measures from electrophysiology [161]. This finding could indicate that the mentioned functional imaging studies that reported percept-related BOLD-modulations in monocular visual areas [252, 109, 292] did indeed measure feedback signals. Our own data does not allow for specific claims on the cerebral origin of rivalry. Nevertheless, the identified ∼400ms limit (see also Chapter IV) to rivalry suggests as well that higher visual areas are involved in the maintenance of rivalry—though not necessarily its initiation—, as the limit seems to be a too long an interaction window for early visual cortex [125, 137].

This discussion of results shows that although ample evidence exists for eye-based suppression, no unequivocal evidence exists for inter-ocular rivalry. These data do show the plausibility of the hypothesis that rivalry takes place at a pattern level, while subsequent feedback to monocular levels ensures eye-based suppression (this is of course not to say that pattern-based competition should only involve neurons that are activated in perfect balance by the two eyes). This interpretation is consistent with the idea that the site where incompatibility is registered may be different from the site where suppression takes place [35].

V.4 Conclusion

In conclusion, the stimuli that were temporally modulated by the introduction of blanks, proved to be valuable tools in dissecting the visual system and identifying and describing the processes that conduce towards the build-up of awareness during normal vision and rivalry. We have shown that the presence of a blank between monocular stimulus presentations eliminates both qualitative and quantitative differences between binocular, monocular and stimulus rivalry. This is a step beyond showing that there are some qualitative similarities between the modes of rivalry. The implication of our findings could be that the three modes of rivalry, and possibly visual conflict resolution in general, share the same underlying pattern-based rivalry system. This suggestion is particularly interesting because of the distinct possibility that pattern-based rivalry is the only form of rivalry. We have furthermore identified a monocular interaction process that may impede rivalry (making it slower and weaker), which may generally camouflage the possibility that rivalry itself originates from a single system.

These findings accord with the view that the rivalry process is not a self-contained system, but is strongly influenced by “add-on” processes. Apart from the monocular interaction process, other factors, such as adaptation [34], context [18], post-selection adaptation and perceptual trapping [246], may act upon the rivalry mechanism as well, especially for more complex stimuli. As a result, the cascade of processes involved in generating visual awareness is indeed multifaceted [38], however, as shown in this report, this does not need to apply to the rivalry process between simple grating stimuli. Note that this same rivalry mechanism is likely to be multi-levelled, increasing suppression strength with the passing of more and more processing stages, as indicated by neurophysiological

[149, 160, 232], psychophysical [198] and computational findings [86].

The dynamics of conscious perception depends on a balance between the volatile nature of an ambiguity-solving process and a resistance against frequent reinterpretations. Our findings indicate that these key functions may be provided by a single pattern-based rivalry system and a monocular integration process, respectively.

V.5 Methods

Apparatus Images were presented on a gamma-corrected 22″ LaCie electron22blueIV monitor (1600x1200 pix, 75 Hz), using a conventional stereoscope with a septum dividing the screen into two equal parts. Screen distance was 46 cm. A chin-rest stabilized head position. Experimental procedures were reviewed and approved by the Institutional Review Board.

Stimulus All patterns were achromatic orthogonally oriented (45 deg from vertical) gratings containing 2.1 cycles/deg, subtending 1.9 degrees. The stimulus was seen through a circular window. Background luminance was 1.5 cd/m^2, mean stimulus luminance was 10.5 cd/m^2. Michelson contrast was 50% in Exp. 1–3, and 100% in Exp. 4. The stimulus was surrounded by a 2.7 deg binocularly visible annulus (0.06 cd/m^2, 0.2 deg wide), serving as a fusion aid. Note that in Exp. 1, stimuli flickered at 19 Hz (by alternately presenting and blanking the stimulus for two successive frames). This caused stimuli with an ISP of 53 ms, and a temporal duty cycle of 0.5, 0.75, and 1, to be physically identical (the 0.25 duty cycles was obtained by showing the pattern for 1 frame, followed by 3 blank frames). Similarly, for an ISP of 107 ms, temporal duty cycles of 0.25 and 0.5 were identical, just as 0.75 and 1. The monocular rivalry stimuli in Exp. 1–3, were shown to a single eye, the other eye receiving a flickering equally-sized disk of mean luminance.

Procedure The subjects' task in Exp. 1–3 was to categorize each trial as containing either rivalry, rapid and regular orientation switches, or a plaid [143]. Subjects based their categorization on the last 3 seconds of a 6-second trial. In Exp. 1, each of the 5 subjects performed 2 sessions, each with 2 reports per parameter combination for stimulus rivalry conditions, followed by 2 sessions, each with 2 reports per condition for monocular rivalry, and 1 report per condition for stimulus rivalry, in which are all conditions were randomly interleaved. The temporal limit reported in Figure 2, corresponds to the ISP where the amount of rivalry reports halved relative to the maximum. Linear interpolation was used between measured ISPs, collapsing the data for all temporal duty cycles.

In Exp. 2, task and stimuli were identical to those in Exp. 1. Stimuli could include a continuously-present 19 Hz flicker, or a 67-ms blank before the stimulus swap. Spatial frequencies of 2, 4,and 6 cyc/deg were tested. Four subjects participated, all of whom had participated Exp. 1.

In Exp. 3, task and stimuli were identical to those in Exp. 1, except for the removal of the 19 Hz flicker,and the addition of blank of 0 to 240 ms, at the end of the ISP, or in the middle of the ISP. Three subjects participated, all of whom had participated Exp. 1.

In Exp. 4, subjects indicated their dominant percept using either of two keys. Subjects were asked not to press when a fast and regular switching of the two grating patterns was perceived, or when the two patterns were overlaid without any of the two being stronger. Trials lasted 60 sec. All 30 conditions were presented once per session, and were randomly interleaved. Four subjects performed two sessions each. Percept durations were calculated as the time elapsed between two successive button-presses; in case a trial passed without button presses, a single 60-sec transition period was recorded. Before further analysis, each percept duration was normalized by dividing it by the mean percept duration of the binocular rivalry condition with an ISP period of 13 ms. The normalization was done separately for each session. Note that this normalization did not bias our data towards similar outcomes in the three modes of rivalry, as we used only one condition as a basis of normalization, i.e. we did not normalize per ISP or per rivalry conditions separately. Figure V.5c, shows that even non-normalized data shows nearly identical duration distributions, showing that the normalization as such was not responsible for the similarities between the rivalry conditions. The total dominance duration was normalized by the trial duration (Figure V.5). The first 20 sec of a trial were not analyzed for this calculation (the total trial duration was therefore 40 seconds). Similarly, the first 10 percepts (amounting to about 20 seconds) were excluded in the frequency histograms in Figure V.5.

Experiment 5. Each trial commenced with conventional (i.e. non-flickering, etc.) binocular rivalry between two orthogonal gratings. Subjects were verbally cued to wait until either the left of right tilted grating was exclusively dominant. When the observer indicated exclusive perception of the verbally cued grating, a blank of variable length was introduced binocularly after which the rivalling stimuli were presented again for 1 second. In this second presentation the stimuli were either swapped between the eyes relative to the first presentation, or they were presented to the same eyes. Subjects indicated which orientation was most dominant during the second presentation.

Chapter VI

Dichoptic masking and binocular rivalry share common perceptual dynamics

Jeroen J.A. van Boxtel
Raymond van Ee
Casper Erkelens

Two of the strongest tools to manipulate visual awareness of potentially salient stimuli are binocular rivalry and dichoptic masking. Binocular rivalry is induced by presenting incompatible images to the two eyes over prolonged periods of time, leading to an alternating perception of the two images. Dichoptic masking is induced when two images are presented once in rapid succession, leading to the perception of just one of the images. Although these phenomena share some key characteristics, most notably the ability to erase from awareness potentially very salient stimuli, their relationship is poorly understood. We investigated the perceptual dynamics during long-lasting dynamic stimulation leading to binocular rivalry or dichoptic masking. We show that the perceptual dynamics during dichoptic masking conditions meet the classifiers used to classify a process as binocular rivalry, that is: (1) Levelt's 2nd proposition is obeyed; (2) perceptual dominance durations follow a gamma distribution; (3) dominance durations are sequentially independent. We suggest that binocular rivalry and dichoptic masking may be mediated by the same inhibitory mechanisms.

VI.1 Introduction

Binocular rivalry and dichoptic masking provide excellent means to study the formation of awareness, because both tools allow one to control the visibility and awareness of a stimulus by presenting a competing stimulus in close spatial or temporal proximity (reviewed in [38, 56, 169]).

Even though both binocular rivalry and dichoptic masking are well-studied phenomena, their relationship is not well understood. This lack of knowledge exists because the two tools have, to our knowledge, never been studied with the same set of stimuli in a single comparative study. One other reason is that research questions for the two tools have generally differed. Research on binocular rivalry is mainly focused on the spatial determinants controlling stimulus visibility, finding for example that with increasing orientation and spatial frequency differences the rivalry rate increases between competing interpretations (e.g. Refs. [5, 117, 206, 210, 230]). When studied, temporal stimulus modulations in the order of hundreds of milliseconds do not seem to affect binocular rivalry [207, 209, 288]. Dichoptic masking research, on the other hand, is mainly focused on temporal aspects of stimulus visibility, showing, for example, that competition is strong at short stimulus onset asynchronies and decreases with longer asynchronies (e.g. Refs. [56, 55, 136, 171, 188, 228, 259]). Interestingly, there is a distinct research direction within the dichoptic masking field that does study spatial characteristics of stimulation. This line of research uses spatially and temporally overlapping targets and masks (e.g. Refs. [28, 145, 155, 183, 184]), and seems more focused on binocular interactions per se than on object or event perception.

There are hints in the masking literature that suggest a link to binocular rivalry. For example, dichoptic masking studies show that the addition of the third stimulus to the target-mask sequence may lead to the annihilation, or weakening, of mask's efficacy (dichoptic studies: Refs. [188, 259], and monoptic studies: e.g. Refs. [55, 71]). Furthermore, masking strength decreases with increasing number of target-mask cycles (which is implicit some reports [136, 229, 277]), an effect reminiscent of the resurfacing of a percept of the suppressed stimulus after prolonged stimulation during binocular rivalry. In this report, we study the effects of long-lasting stimulation on dichoptic masking, and the possible relations to binocular rivalry it may reveal.

Previously we have shown (Chapter IV) that binocular rivalry may take place with intermittent stimulation (see also Ref. [209], where this condition was termed successive rivalry), even when stimulation in the two eyes is asynchronous (see e.g. Figure VI.1a). While viewing these stimuli, the observer perceives one of the two oriented gratings flickering on and off (being 'off' most of the time), while the other stimulus remains invisible. After a certain time ($\sim$2 seconds) perception switches to the other grating, etc. (this chain of events is schematically drawn in Figure VI.1b, top panel). However, we also showed that with asynchronous stimulation, binocular rivalry cedes to dichoptic masking when the period of pattern repetition exceeded ∼350 ms (Chapter IV; schematically drawn in Figure VI.1b, bottom panel). Taking these minimally different stimuli—differing only in stimulus repetition period—allowed us to investigate whether dichoptic masking and binocular rivalry are outputs of a single dynamical system, or instead are two separate processes. We focused on the perceptual dynamics of binocular rivalry and dichoptic masking employing long-lasting dynamic stimulation described in Figure VI.1.

For binocular rivalry the perceptual dynamics of alternations in dominance are characterized as follows: (1) dominance durations follow a gamma distribution; (2) the dominance durations of the percepts are temporally uncorrelated [84]; (3) Levelt's 2nd proposition is obeyed [85, 154], meaning that changing the contrast of one eye's pattern will change the dominance durations of the pattern in the other eye, but will leave the dominance durations of the contrast-changed pattern unchanged; or put a modified form (which is valid over a larger contrast range): changes in one eye's contrast mainly affect dominance durations in the higher contrast eye [53].

We show that both successive rivalry and dichoptic masking meet the perceptual dynamics criteria set for conventional binocular rivalry. We observe as well that the first reported percept shows signs of dichoptic masking. Depending on stimulus parameters, this initial dichoptic masking may be restricted to the first percept (after which binocular rivalry takes place), or it may be sustained throughout the trial, leading to near-continuous dichoptic masking.

From the fact that the perceptual criteria for binocular rivalry are met by dichoptic masking, we suggest that dichoptic masking and binocular rivalry are mediated by the same inhibitory mechanisms.

VI.2 Methods

Stimulus

Spatial characteristics Stimuli were described in Chapter IV.

Temporal characteristics The grating patterns were flickered on and off for 60 seconds. On-times were 53.3 ms (4 frames), and off-times were either 133.3 ms, or 293.3 ms, leading to repetition periods of 186.7 ms and 346.7 ms, respectively (Figure VI.1). These stimuli produce binocular rivalry and dichoptic masking respectively (Chapter IV). Left and right eye were temporally offset relative to the each other by a stimulus onset asynchrony (SOA). The SOAs ranged from 0 ms (i.e. no lag) to the size of the repetition period, in steps of 26.6 ms (i.e. two frames; Figure VI.1). In plotting and discussing the data we use the shortest of the two SOAs, which we term minSOA (i.e. minSOA = MIN[SOA1,SOA2]). The pattern that preceded the shortest SOA is termed the lead pattern, the pattern following the shortest SOA is termed the lag pattern.

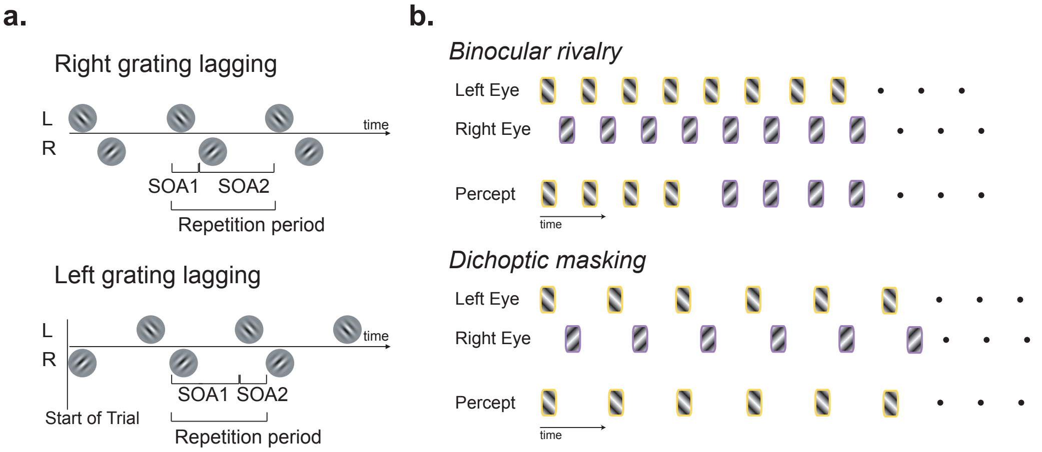


Figure VI.1 | Two example stimuli used in the experiment. (a.) *At the start of the trial, either the grating in the right eye or the orthogonal grating in the left eye was temporally offset (see top panel and bottom panel, respectively). The stimulus onset asynchrony (SOA) between the first and the second presentation was termed SOA1, the other SOA was termed SOA2. The sum of SOA1 and SOA2 is the repetition period. The top panel represents a case where SOA1<SOA2, the bottom panel represents a case when SOA1>SOA2. The effective stimulation is identical in both cases, apart from the start of the trial. In the remainder, we show the dependence on the shortest of the two SOA (i.e. minSOA=MIN[SOA1,SOA2]). We refer to the stimulus that precedes the shortest SOA as the 'lead' stimulus, and the stimulus that follows this SOA as the 'lag' stimulus (the corresponding percepts are called 'lead' and 'lag' percept, see text).* (b.) *The perceptual consequences of stimulation with repetition cycles <350 ms* (top panel) *are binocular rivalry. The percept will consist of bouts of flickering leftward tilted gratings and flickering rightward gratings. (*Bottom panel*) The perceptual consequences of stimulation with repetition periods >350 ms, are dichoptic masking. The percept will consist of long stretches of a single flickering grating, only to be interrupted briefly and infrequently by the competing pattern.*

Procedure

Throughout the trial, subjects indicated their dominant percept using two buttons (left and right arrows for left- and right-tilted gratings, respectively). When both stimuli were perceived to be about equally salient (i.e. a spatial patchwork or a superposition of the two gratings), or when they did not engage in rivalry (i.e. they were rapidly and regularly alternating), subjects did not press any button. Percept durations were recorded as the time between a button press and release. A percept was not used in the analysis when it was truncated by the end of a trial. Trials with different repetition periods were run in separate sessions. Otherwise, presentation order was randomized. Each condition was measured three times per subject.

Subjects

Seven subjects participated, each has normal or corrected-to-normal vision. All but one subject (the first author) were naïve as to the purpose of the experiment. As two subjects did not show signs of dichoptic masking at the repetition period of 347 ms, their data were excluded from the analysis of this repetition period. (For both subjects we have determined that dichoptic masking does take place at still longer repetition periods in a separate experiment.)

Analysis

Perceptual bias The total time spent in the "lag" percept (the percept of the stimulus that followed after the shortest of the two SOAs, see Figure VI.1) was divided by the total time spent in either percept.

Percepts follow a gamma distribution The reported gamma-distribution fits are maximum-likelihood fits, based on all percepts of all subjects, without normalizing. We did not normalize the data in order to show the differences in mean percept durations between minSOA-conditions. When each subject's data was divided by the subject's mean percept duration before pooling the data over all subjects, gamma distributions still provided good fits (all means of squared errors < 0.017).

Drift analysis Before conducting a test of independence of percept durations, we assessed whether there was drift in the data, because drift may cause spurious correlations in the percept durations. We divided each trial in 13 bins, within which the average percept duration was calculated for all percepts that started within that bin. The averages of all trials were then used to calculate average percept durations per bin, per subject. A Bin × Subject ANOVA revealed no significant effect of Bin number ($p>.42$), but a significant effect of Subject ($p<0.001$), post-hoc Tukey analysis showed that one subject had significantly higher durations than all other subjects. The interaction Bin × Subject was not significant ($p> 0.77$). A separate analysis showed that the drift over sessions was also not significant ($p>0.85$).

Independence of percept durations Because of the absence of drift, we could calculate Spearman rank correlations [76] over entire trials. Only trials with more than 5 reported percepts are analyzed, which caused some minSOA results not to be based on all subjects (see Figure VI.2D).

Calculation of bias-index A bias index was designed to describe the development of the average perceptual bias towards one or the other percept during a trial, allowing discrimination between binocular rivalry and dichoptic masking behavior throughout the trial. We calculated a bias-index by assigning a value of

1 to percepts of the lag pattern, 0 to percepts of the lead pattern, and 0.5 to transition periods (i.e. when no buttons were pressed). At several time-points during a trial (taken 1 second apart, Figure VI.4) we calculated the mean over these values for all trials at that point in time. Note that we assigned a value to transition periods, as it seems they represent a separate perceptual state [53]. Because they do not represent a bias in a particular direction we assigned them an intermediate value of 0.5.

VI.3 Results

The perceptual dynamics of the two tested repetition periods are shown in Figure VI.2. To show that binocular rivalry was obtained with the 187 ms repetition cycle, and dichoptic masking with the 347 ms repetition cycle, we first analyzed the overall bias in percept durations (Figure VI.2a). Previous work found that binocular rivalry occured at the shorter repetition period of 187 ms (Ref. [209], and Chapter IV) and dichoptic masking at longer repetition periods (Chapter IV). Therefore, we expected that short repetition periods would not lead to a strong bias toward one or the other stimulus, and hence a bias of around 0.5, and that longer repetition periods would show a strong bias towards one of the two stimuli, and a bias different from 0.5. Both expectations were borne out by the analysis (Figure VI.2a; on the y-axis is shown the total time spent in the percept of the lag grating, divided by the total time spent in either percept).

These analyses, based on the biases calculated over entire trials, have revealed that conventional binocular rivalry took place for stimuli with a repetition cycle of 187 ms (this type of results has been reported for in-phase, SOA=0 ms, and antiphase conditions [209]), whereas strong forward masking occurred with a repetition cycle of 347 ms. Therefore we can use these stimuli to investigate the relationship between binocular rivalry and dichoptic masking with (nearly) identical stimuli.

Is the forward dichoptic masking we observed (Figure VI.2a, right) a different kind of binocular conflict resolution than binocular rivalry (Figure VI.2a, left), or is it a biased version of binocular rivalry? In other words, do these phenomena depend on the same or different neural processes? To investigate the question we looked at the 3 perceptual-dynamics classifiers of binocular rivalry, and assessed whether they applied to dichoptic masking.

Levelt's 2nd proposition

We did not explicitly change the stimulus contrast to which Levelt's 2nd proposition applies, but we did vary the SOAs. As we have shown, this manipulation biased the percept towards one or the other pattern, just as contrast would. We cannot use our minSOA parameter (i.e. Min[SOA1,SOA2]) as a direct analogue of

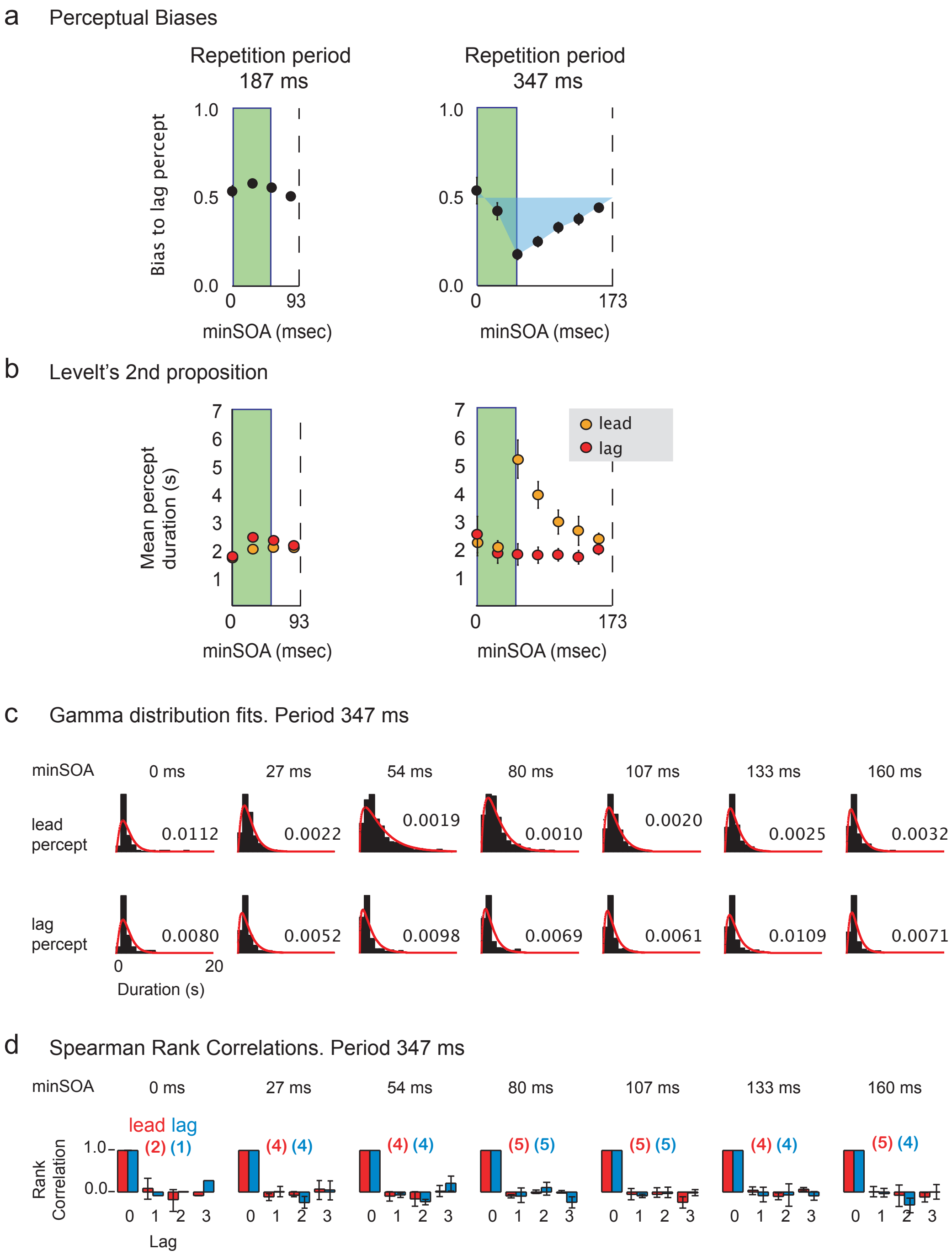

a Perceptual Biases
Repetition period 187 ms
Repetition period 347 ms
Bias to lag percept
1.0
0.5
0.0
0
93
173
minSOA (msec)
b Levelt's 2nd proposition
Mean percept duration (s)
7
6
5
4
3
2
1
lead
lag
c Gamma distribution fits. Period 347 ms
minSOA
0 ms
27 ms
54 ms
80 ms
107 ms
133 ms
160 ms
lead percept
0.0112
0.0022
0.0019
0.0010
0.0020
0.0025
0.0032
lag percept
0.0080
0.0052
0.0098
0.0069
0.0061
0.0109
0.0071
0
20
Duration (s)
d Spearman Rank Correlations. Period 347 ms
Rank Correlation
lead lag
(2) (1)
(4) (4)
(4) (4)
(5) (5)
(5) (5)
(4) (4)
(5) (4)
1.0
0.0
0 1 2 3
Lag

Figure VI.2 | Testing the three classifiers of binocular rivalry on dichoptic masking data. (a) *Perceptual biases were calculated as the total time spent perceiving a 'lag' pattern divided by the total time any of the two patterns was perceived. With repetition periods of 187 ms (left), the value of minSOA (i.e. Min[SOA1,SOA2]) had only a small influence on the perceptual biases, which in all cases were near 0.5. With repetition periods of 347 ms, strong (forward) masking was observed (a bias towards the 'lead' pattern; shaded area). Green-shaded areas denote SOA-conditions with a temporal overlap between the two stimuli. Error bars are s.e.m. over trials.* (b) *The generalized 2nd proposition of Levelt poses that in response to changing the saliency of one of the stimuli, the percept durations of the most dominant pattern should change. The graph for repetition periods of 187 ms shows only marginal changes in dominance durations, which is consistent with the finding that these conditions show hardly any perceptual biases. The graph for repetition periods of 347 ms, however, clearly follows the generalized 2nd proposition of Levelt. Error bars are s.e.m. over trials.* (c) *Gamma distributions (red lines) were well-fit to all percept duration distributions (black frequency distributions), for both lead and lag percepts. Values reported in each graph are the means of squared errors (MSE), a measure of the fit quality.* (d) *Spearman Rank correlations for both lead and lag percepts were calculated for a range of different lags. The correlation is weak at all lag-values, for all conditions. Error bars are between-subject standard errors, the colored numbers indicate the number of subjects over which the mean and s.e. were calculated (see methods).*

contrast, however, because our stimuli had fixed repetition periods, which meant that a change in SOA1 automatically lead to a change in SOA2 (see Figure VI.1), analogous to changing contrast in both eyes. However, we can test the general version of Levelt's 2nd proposition: changes in one eye's contrast mainly affect dominance durations in the eye with the most dominant pattern [53], that is, the lead stimulus (see Figure VI.2a, right).

First, the data pertaining to conditions with repetition cycles of 187 ms show no large change in average dominance durations, and therefore these conditions do not represent a temporal analogue of Levelt's 2nd proposition. This finding is in agreement with the finding that they did not show a perceptual bias either (Figure VI.2a, left).

In the dichoptic masking regime, with repetition periods of 347 ms, mean percept durations did depend on minSOA. Starting, in Figure VI.2b (right panel), at the point where SOA1 and SOA2 are nearly identical (around minSOA = 177 ms), we observe that the percept is unbiased, which would be analogous to equal contrasts in the two eyes. Moving to the left in this graph, we observe that the mean percept duration of the lead stimulus increases whereas that of the lag stimulus does not change, which is in accordance with the generalized 2nd proposition of Levelt.

A deviation from the predicted behavior occurred when the stimuli temporally overlapped (i.e. at the left of the plot). In this case, both patterns had mean dominance durations of about 2 sec. "False fusion" [43, 288], a process causing dichoptic stimuli to fuse with short presentation periods, may have prevented strong inhibitory interactions in these cases, even though rivalry still occurred [209, 288].

Nevertheless, our results show that conditions that lead to dichoptic masking follow the generalized 2nd proposition of Levelt.

Distributions of percept durations

Both lead and lag percept duration distributions followed a gamma distribution (Figure VI.2c), suggesting that the perceptual dynamics followed the same set of rules in all these conditions, as well as during normal binocular rivalry. In cases of strong masking average durations of the lead percept increased while the distribution remained gamma-shaped. The distribution for lag percept durations remained unchanged relative to the conditions in which both patterns exerted about equal inhibitory forces on each other (e.g. at minSOA = 0 ms and 160 ms).

Temporal independence of percept durations

We tested for the independence of the percept durations using the Spearman rank correlation [76]. Both lead and lag percept durations were temporally independent of the following percept of the same class, as evidenced by the small correlations between the different percept durations (Figure VI.2d). The 187 ms period condition also met the perceptual criteria, indicating that these conditions of successive rivalry [209] behaved just as conventional rivalry.

Altogether, the dichoptic masking conditions met the 3 perceptual classifiers of binocular rivalry, suggesting a common neural structure for the found percept-competition reported with both techniques.

The time course of dichoptic masking and binocular rivalry

In order to see how much of the observed biases (in Figure 2a) can be attributed to the first percept (cf. Ref. [122]), we calculated the biases towards either one or the other percept for the first button press. Three interesting observations were made. First, when biases are reported as the number of trials in which the first percept was that of the lag stimulus divided by the total number of trials, both the 187ms and 347 ms repetition periods showed dichoptic masking behavior for the first percept (Figure VI.3a and VI.3b). For the repetition period of 347 ms dichoptic masking was expected, as a bias was already observed in the overall data (Figure VI.2a), but for the repetition period of 187 ms this was not the case (although a small bias seen at minSOAs around 27 ms might have suggested this finding; Figure VI.2a). Interestingly, even though the initial stimulus sequence for the stimuli with 187 ms and 347 ms repetitions cycles was identical (for the minSOA range 0-80 ms), the masking effects of the two repetition periods were in opposite directions. We have no explanation for this results, but it may be related to the observation that the addition of a third stimulus in DM paradigms may reduce backward masking of the mask on the target [188] for some SOAs combinations.

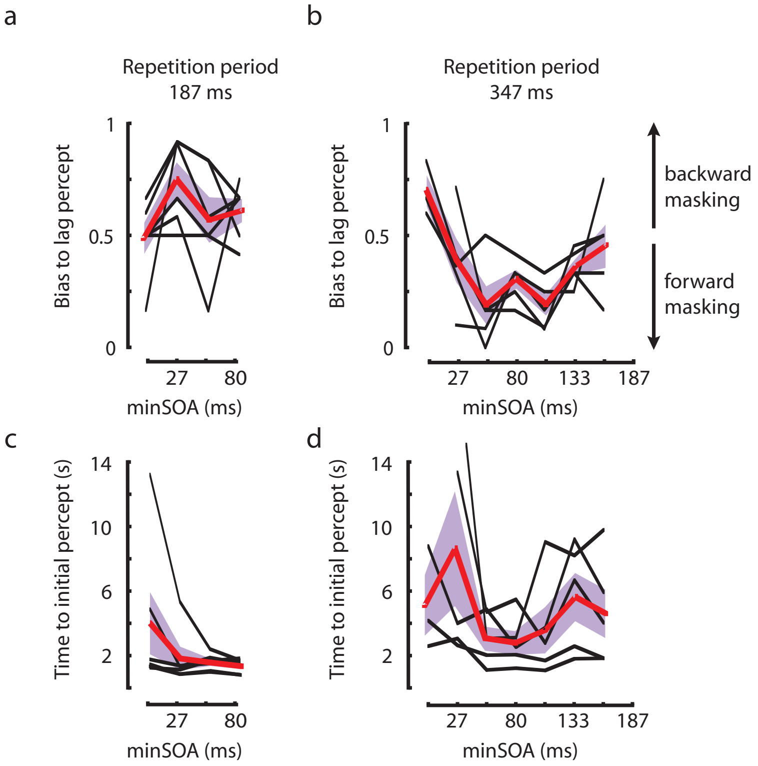


Figure VI.3 | Masking effects on the first perceptual report. (a) *Proportion of the initial percepts being the lag stimulus for stimuli with a repetition period of 187 ms. The bias toward the lag stimulus for short minSOAs shows evidence for backward masking.* (b) *Proportion of the initial percepts being the lag stimulus for stimuli with a repetition period of 347 ms. The bias towards the lead stimulus shows evidence for forward masking.* (c) *The average time until the first percept, for stimuli with a repetition period of 187 ms. Longer durations are observed in conditions with the inhibitory interactions being about equal between the two stimuli (that is at minSOA = 0).* (d) *Same as in c, but for stimuli with a repetition period of 347 ms (the minSOA=0 condition is based on 3 subjects, the other 2 did not report rivalry). Panels c and d show that the first percept generally was formed swiftly, but it takes longer to form when inter-pattern competition is balanced. The thin black lines are individual subject data, the thick line is the mean, the shaded area is the between-subject standard error.*

A second observation was that the first perceptual report was generally made swiftly, around 2-3 seconds after the trial began (Figure VI.3c and VI.3d, for the 187 ms and 347 ms repetition cycles, respectively). However, for both repetition cycles, we observed a tendency towards longer durations for conditions when the inhibitory interactions between the patterns were equal (either because stimuli were near-simultaneous or because SOA1 was nearly equal to SOA2, see the right- and left-end side of each plot). These data suggest that the initiation of binocular rivalry is not always instantaneous (see also Ref. [157]). In fact we found that conditions that lead to masking lead more rapidly to an unambiguous percept than conditions that lead to binocular rivalry.

The third observation followed from the acknowledgement that the bias of the initial percept may be the main cause of the bias observed in dichoptic masking (Figure VI.2a). To analyze how much of the observed perceptual biases based on the full-trial data could be accounted for by the initial percept, we calculated a bias-index (see methods). We compared the time course of the bias-index for the data including only the first percept, and for the data including all percepts. As such, we investigated if the binocular interactions were changing over the course of a trial. The bias-index in conditions with repetition periods of 187 ms and 347 ms are shown in Figure VI.4a and VI.4b respectively.

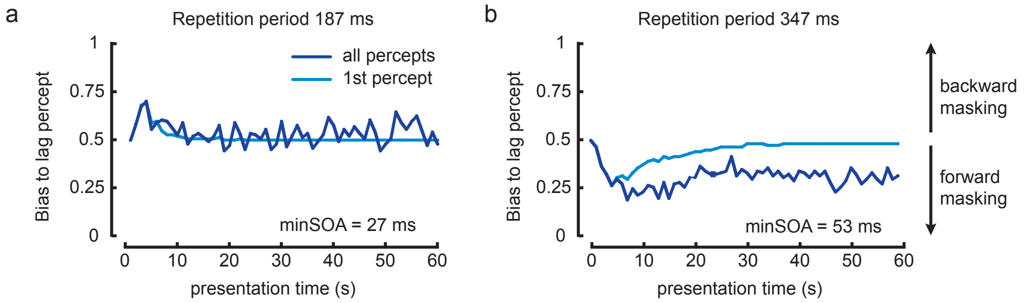


Figure VI.4 | Bias-index during the 60-second trials. (a) *The bias-index for stimuli with a repetition period of 187 ms, and a SOA of ∼27 ms. The bias-index calculated using all percepts (dark blue line) shows an initial deviation towards the lag stimulus (backward masking), but the preference rapidly fades. This bias seems to be fully caused by the first percept (light blue line).* (b) *The bias-index for stimuli with a repetition period of 347 ms, and a SOA of ∼53 ms. The bias-index calculated using all percepts (dark blue line) shows an initial deviation towards the lead stimulus (forward masking), the preference weakens but remains present throughout the trial. The calculated bias-index based solely on the first percept cannot explain the sustained presence of masking, because the first percept does not add to the observed biases after more than 20 seconds from the start of the trial. Therefore, the bias observed in the first percept, Figure VI.3, can explain the bias observed in the total data, Figure VI.2a, only for stimuli with a repetition period of 187ms, not for stimuli with a repetition period of 347ms.*

We show data for the minSOA-conditions with the most extreme biases (see Figure VI.2a, and VI.3). The dark blue line shows the bias-index calculated using all perceptual reports, the light blue line shows the bias-index based solely on the first percept (all later percepts are scored as transition periods, i.e. a value of 0.5). The bias-index based on all perceptual reports for repetition periods of 187 ms showed a short-lasting backward masking, which could however, be fully accounted for by the bias caused by the first percept (the light blue line). The bias-index for repetition periods of 347 ms showed a sustained bias that could not be explained based solely on the bias of the first percept (i.e. the light blue curve regresses toward 0.5, whereas the dark blue curve remains at $\sim$0.25). Therefore, the biases observed in Figure VI.2a are caused predominantly by the first percept for stimuli with repetition periods of 187ms, but not for those with repetition periods of 347ms.

VI.4 General discussion

The relationship between binocular rivalry and dichoptic masking

The relationships between results from the various techniques used in the investigation of the formation of visual awareness are only beginning to be explored (Refs. [56], pp. 272-275, [131, 258]). In this study, we looked at the dependence of the stimulus' visibility on temporal stimulus characteristics, and how this dependence differed during binocular rivalry and repetitive dichoptic masking. Specifically, we tested whether the classifiers related to the perceptual dynamics of binocular rivalry are also met by dichoptic masking.

We show that dichoptic masking satisfies the perceptual dynamics criteria for binocular rivalry (Figure VI.2), namely, (1) the patterns presented to the two eyes are dichoptic (i.e. different in the two eyes); (2) the generalized 2nd proposition of Levelt [53] is met; (3) dominance durations follow a gamma distribution; (4) the dominance durations of the percepts are temporally uncorrelated.

On the basis of these data, we would like to suggest that dichoptic masking effects and binocular rivalry are the resultant of (partly) overlapping neural mechanism. However, we have only looked at the temporal dynamics of perception, and the hypothesis is strengthened if the dependency on spatial stimulus manipulations is also similar. Indeed, it seems that binocular rivalry and dichoptic masking have a similar tuning to orientation (width of 70 deg [106, 155, 206, 230]), a relatively tight tuning to spatial frequency differences between the conflicting stimuli [106, 117, 145, 155, 206, 230] and a relative independence of spatial frequency when both stimuli are large (>4 deg) and have the same spatial frequency [145, 210]. Some further circumstantial evidence comes from a monocular masking study, which indicated that masking and adaptation may have a common origin [93]. Since adaptation is one of the main components thought to underlie binocular rivalry dynamics, these results are suggestive of a link between masking and binocular rivalry. The similarity is larger than one might expect on the basis of the very rare cross-references across the two research fields [5, 56, 184].

Nevertheless, a more extensive comparison is needed to establish whether binocular rivalry and dichoptic masking may indeed depend on the same (or party overlapping) underlying mechanisms. If our suggestion of a single system for dichoptic masking and binocular rivalry turns out to be sustainable, it would provide an interesting test case for modelers, who could test whether their model can explain both dichoptic masking and binocular rivalry data.

Other studies on the relationship between dichoptic masking and binocular rivalry

In a previous study (Ref. [56], pp. 272-275), the relationship between dichoptic masking and binocular rivalry was studied by presenting the observer a dichoptic masking stimulus that was surrounded by an annular binocular rivalry stimulus. It was found that the effectiveness of a dichoptic mask was contingent on whether it was being present in the eye receiving the dominant binocular rivalry stimulus. It was concluded that binocular rivalry takes precedence over dichoptic masking as it determines whether masking occurs or not, a conclusion seemingly at odds with our suggestion of a shared system. However, in that study backward masking conditions were used, which may involve neuronal networks different from those involved in forward masking. Secondly, in our study the same stimuli induced dichoptic masking and binocular rivalry whereas in Ref. [56] separate, and spatially non-overlapping stimuli for binocular rivalry and dichoptic masking were employed. This may be a crucial difference as it has been shown in monocular conditions that masking by a surrounding mask (called meta- and para-contrast masking) is different from masking by conflicting pattern information similar to the target pattern, which in turn is different from masking by noise [56]. These three types of masking are currently also employed in dichoptic masking (e.g. [56, 172, 184, 257], but their mutual relationships have not yet been thoroughly investigated, and only recently their relationship to binocular rivalry received attention (this report, and Refs. [56], pp. 272-275; [170, 258]). So far, the differences in the employed methods and stimuli probably preclude an informative inter-study comparison.

On the timecourse of dichoptic masking and binocular rivalry

We found that the first-reported percept shows systematic biases towards one of the two stimuli, i.e. dichoptic masking (Figure VI.3). This initial bias was found to be the cause of the small bias observed for stimuli with repetition period of 187 ms (Figure VI.4) that would eventually lead to binocular rivalry. The reported biases for the full-trial data of repetition period 347ms can however not be fully explained based on the initial bias, showing that the found near-continuous dichoptic masking was not fully dependent on the initial percept.

To allow for the unbiased perception, i.e. binocular rivalry, with our flickering stimuli, binocular information needs to be integrated over long periods ($\sim$350 ms). Our study does not stand alone in presuming the existence of such long binocular integration periods, as they have been reported before [20, 73, 209]. Presumably, when presentations are separated by more than 350 ms, the involved neurons cannot integrate the information over successive presentations, and transient boosts in activity of these neurons may cause transient cross-inhibition between competing representations resulting in dichoptic masking when SOA1 and SOA2 are unequal.

These particular interactions are not static, however, as we have found that many trials with repetition periods of 187ms start with dichoptic masking. Overall, this dichoptic masking effect turns into binocular rivalry over the first 6 seconds of a trial. We may explain this finding by assuming that the binocular integration window is initially small (<190 ms) but slowly sizes up to its maximum size (∼350 ms). Initially this process allows each individual event to transiently inhibit competing patterns (causing dichoptic masking), but later causes the inhibitory forces for both patterns to become more or less equal (resulting in binocular rivalry). Consistent with this idea, is the finding that a substantial amount of cortical neurons show flicker adaptation ([100], p. 527), meaning that at first they are able to modulate the firing rate according to the flicker modulation of the stimulus, but that after a certain while they stop firing in synchrony with the stimulus, or stop firing altogether. This adaptation has a timescale of 1 to 5 sec ([100], p. 527), close to the timescale we observe for the transition of masking to binocular rivalry for stimuli with repetition periods of 187 ms.

On the usefulness of continuous dichoptic masking

Our report also reveals an interesting tool for future study. Dichoptic masking conditions with repetition periods around 350 ms lead to very long dominance durations. Continuous and complete dichoptic masking, with percept durations in the order of 10 seconds (and up to 30 seconds for some subjects), is possible for repetition periods around 400 ms (data reanalyzed from experiment 2 of Chapter IV). Continuous dichoptic masking may be related to other effects that cause long-lasting perceptual disappearance of potentially salient stimuli, like continuous flash suppression [257], and the dichoptic standing wave of invisibility [170]. The advantage of continuous dichoptic masking is that masking and target stimuli are identical (save for orientation), and need not differ in spatial frequency content, and luminance [257] or presentation duration [170], and thus may provide an additional route for the investigation of stimulus visibility [170] and the origin of visual aftereffects [257].

Chapter VII

Retinotopic and non-retinotopic stimulus encoding in binocular rivalry and the involvement of feedback

Jeroen J.A. van Boxtel
David Alais
Raymond van Ee

Adaptation is one of the key constituents of the perceptual alternation process during binocular rivalry, as it has been shown that preadapting one of the rivaling pairs before rivalry onset biases perception away from the adapted stimulus during rivalry. We investigated the influence of retinotopic and spatiotopic preadaptation on binocular rivalry. We show that for grating stimuli, preadaptation only influences rivalry when adaptation and rivalry locations are retinotopically matched. With more complex house and face stimuli, effects of preadaptation are found for both retinotopic and spatiotopic preadaptation, showing the importance of spatiotopic encoding in binocular rivalry. Are the spatiotopic adaptation influences on rivalry caused by direct spatiotopic stimulus interactions, or instead are they due to altered feedback from the adapted spatiotopic representations to the retinotopic representations that are involved in rivalry? By using rivaling stimuli that minimize rivalry between spatiotopic representations while still engaging these representations in stimulus encoding, we show that at least part of the preadaptation effects with face stimuli depends on feedback information.

VII.1 Introduction

Binocular rivalry is one of the main tools used to study the formation of visual awareness. To instigate binocular rivalry, one places two conflicting but unchanging images in the two eyes. Despite the fact that the retinal input is unchanging, visual awareness changes from one of the eyes' images to the other. This property means that one has a window on the repeated formation of visual awareness without any change to the external stimuli. One idea that has become accepted recently is that rivalry takes place at multiple levels within the visual system, and that rivalry strength increases at successive levels [38, 86, 149, 160, 198]. The specific interactions between these different levels, and the computations that occur within them, are still relatively unclear. The present study seeks to clarify aspects of these processes by comparing retinotopic and spatiotopic effects of adaptation.

What causes the internally induced switches from one rival stimulus to the other? According to the general view, one of the key constituents of the rivalry switching process is visual adaptation (e.g. Ref. [41]). Adaptation is a common neural process that refers to the decrease in sensitivity that occurs among neurons after prolonged stimulation. Critically, adaptation persists after the adapting stimulation ceases. Psychophysically, this leads to effects such as afterimages (of colour or contrast, for example), or to aftereffects such as the motion and tilt aftereffects. In the context of binocular rivalry, adaptation to one of the rival stimuli is thought to decrease the sensitivity of the responsive neural population and cause a shift in inhibitory interactions in favor of the competing neural population, resulting in a perceptual switch to the less adapted stimulus [41, 142, 204, 284].

One of the useful characteristics of adaptation is that it is selective to a particular group of 'tuned' neurons. This specificity can be seen in the variety of forms of adaptation that have been reported. For example, adaptation can occur to image contrast [65, 242], to orientation (e.g. Ref. [45]), to motion [181], to spatial [46] and temporal [227] frequency, and even to object content such face identity [151, 274]. Adaptation therefore is often used as a probe to try to identify the locus of particular visual processes, sometimes known as psychoanatomy [126]. Consequently, the neural loci of adaptation to various visual stimuli are better known than those for binocular rivalry. For this reason, manipulating the visual system's state of adaptation to various stimuli offers great potential for revealing otherwise hidden aspects of the neural processes underlying binocular rivalry and for dissecting the visual hierarchy thought to underlie binocular rivalry.

Apart from stimulus specificity, another interesting aspect of adaptation concerns the spatial frame of reference in which it occurs. Traditionally, the effects of adaptation are measured by adapting locally to a particular stimulus and then presenting a subsequent test image at the same retinal location. However, if the observer makes an eye-movement between adaptation and testing so that their gaze is directed to another location in the outside world, it is possible to separate retinotopic adaptation from spatiotopic adaptation. When the effects of adaptation are measurable at the same retinal location, it is said to be retinotopic adaptation[1]. In contrast, when the effects of adaptation are measurable at the spot in the outside world where the adaptation stimulus was presented, it is said to be spatiotopic adaptation. Both types of adaptation can be demonstrated [185, 203]. In general, the more complex the adapting image, the more spatiotopic adaptation occurs [185].

In this paper, we will compare retinotopic and spatiotopic adaptation effects on binocular rivalry. The usual manner for investigating the influence of adaptation on rivalry has been to preadapt to one of the rival stimuli before the rivalry measurements begin. This is thought to desensitise neurons responsive to that stimulus and bias perception towards the other stimulus during rivalry [40, 196, 265], showing a causal role for adaptation in determining rivalry alternations. In these previous studies, however, retinotopic and spatiotopic adaptation were confounded, and the possible influences of image complexity were not studied. In this report, we will disentangle retinotopic and spatiotopic adaptation processes by separately testing retinotopically-matched and spatiotopically-matched locations. We will also compare the relative importance of these different types of adaptation for both simple grating stimuli and more complex house-face stimuli. In a final experiment, we will report on the effects of feedback of adaptation to lower levels, as feedback is a likely means of communication between the different levels of processing involved in binocular rivalry.

VII.2 Methods

Subjects. 9 subjects participated in Experiment 1, 9 subjects participated in Experiment 2 (6 has also participated in Experiment 1), and 6 subjects participated in Experiment 3 (5 of them participated in Experiment 1, and 5 also participated in Experiment 2).

Apparatus. 3 of the 9 subjects in Experiment 1, and 4 of the 9 subjects in Experiment 2 were shown stimuli on an Iiyama Monitor (1280×1024 pix at 75 Hz), in ambient daylight. The other subjects were shown stimuli on a gamma-corrected 22 LaCie electron22blueIV monitor (1600×1200 pix, refreshed at 75 Hz). No difference was observed between the two groups of subjects, and the data was therefore combined. To create binocular rivalry, we employed a conventional stereoscope with a septum dividing the screen into two equal parts.

Stimuli. In Experiment 1, a circular grating with a diameter of 1.6 degrees, 2.5 cycles/deg was displayed at 100% contrast during adaptation, and gratings had a 30% contrast during rivalry. In Experiment 2, house and face stimuli (see Figure

---

[1] Note that retinotopic adaptation is not identical to retinal adaptation. Retinotopic adaptation takes place relative to the position on the retina, but not necessary in the retina itself.

VII.1c), were viewed through a vertically oriented oval window (long axis $\sim$2.4 deg, short axis $\sim$1.2 deg). House and face stimuli were equated in RMS-contrast (contrast was 45%). In Experiment 3, house-grating rivalry was studied. Figure VII.1c displays the used face stimulus. The grating in Experiment 3 had the same contrast during adaptation and rivalry, which was set per subject to obtain about equal dominance durations for the grating and face stimuli in unadapted situations. The grating was oriented 45 degrees from vertical, and was seen through an oval window with a size identical to the face stimulus. In this experiment, the grating contained 10 cycles ($\sim$4 cycles/deg). All stimuli were displayed on a mean-luminance background, for the LaCie monitor this was 6.4 cd/m^2. A 9$\times$9 degree box, that functioned as a fusion aid, and a spatiotopic reference, surrounded the area in which stimuli could be displayed.

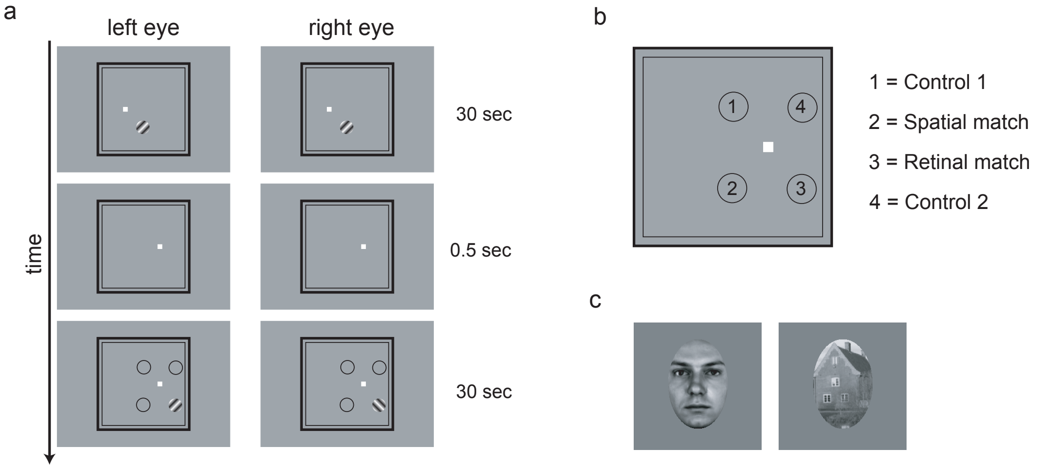


Figure VII.1 | Stimuli and experimental protocol. (a) *A trial started with a 30-second adaptation phase. During this phase, subjects fixated a bright fixation point, while a non-rivalrous stimulus was presented in the periphery. After the adaptation phase, the fixation point jumped to its new location. After 0.5 seconds a rivalrous stimulus appeared for 30 seconds in one of four locations indicated by a circle in this schematic. The circles were not presented in the real stimulus. During the rivalry phase subjects indicated their dominant percept.* (b) *The four possible locations of the rivalry stimulus comprised two control conditions (indicated by 1 and 4, which are called Control 1, and Control 2), one location that was retinotopically-matched to the adaptation condition (location 3), and one location that was spatiotopically-matched to the adaptation conditions (location 2).* (c) *The house and face stimulus used in experiment 2 and 3.*

Protocol. Each trial started with the appearance of a fixation dot, 1.52 deg left of the center of the display. When the subject pressed a button, a period of binocular adaptation to one of the two stimuli was started (see Figure VII.1a). The adaptation period lasted 60 seconds for the first trial, and 30 seconds for all subsequent trials. During this adaptation period the subject was asked to keep fixation on the fixation mark only. At the end of the adaptation period the fixation mark jumped to a position 1.52 deg right of the center of the display, and the subject was required to saccade quickly to the new fixation mark. After 0.5 seconds, the rivalrous stimulus was displayed for 30 seconds, during which period the subject continually indicated the dominance of the two images using two buttons assigned to the two percepts. The rivalrous stimuli could appear at 1 of 4 locations (see Figure VII.1b). One of these was in retinotopic correspondence to the adaptation stimulus, one location was in spatiotopic correspondence, and two locations functioned as control conditions (as illustrated in Figure VII.1b; see Control 1 and Control 2). Image locations were 2.15 degrees from fixation, located at 45 degrees, 135 degrees, 225 degrees, and 315 degrees from vertical upward (see Figure VII.1). Within sessions a single adaptation stimulus was used. House and face adaptation in Experiment 2, and face and grating adaptation in Experiment 3, were balanced within each subject, but adaptation sessions to houses and faces always took place on different days. Within one session each condition was repeated three times.

VII.3 Results

Experiment 1—grating adaptation

During the adaptation phase, subjects viewed a leftward tilted grating (adaptation to the right grating was not tested). After saccading to a new fixation location, a rivalrous stimulus was placed in 1 of 4 spots, matching the adaptation location either in retinal coordinates, spatial coordinates, or neither (see Figure VII.1). Rivalry took place between two gratings and subjects reported their changing percept during the entire 30 seconds. We measured the average dominance duration of both the adapted grating and its rivaling partner grating. To quantify the effect of adaptation, we divided the average percept duration of the adapted grating by the average duration the orthogonal partner grating, normalized by any bias observed in the Control 1 condition shown in Figure VII.1b. Measures smaller than unity indicate that adaptation biased predominance away from the adapted stimulus.

In Figure VII.2a, we plot the adaptation influence for the four tested rivalry-positions. For the grating stimuli used in this experiment, adaptation does not bias perception for either of the control conditions, nor for the spatiotopic test spot. It was only in the retinotopic condition that adaptation led to a significant bias (i.e. a difference from 1) towards the non-adapted grating ($p < 10^{-5}$; when not further specified p-values are derived from two-tailed t-tests). The bias, 0.68 (range among subjects: 0.52-0.82), is of a strength comparable to previous research [40], where biases between 0.5 and $\sim$0.75 were reported.

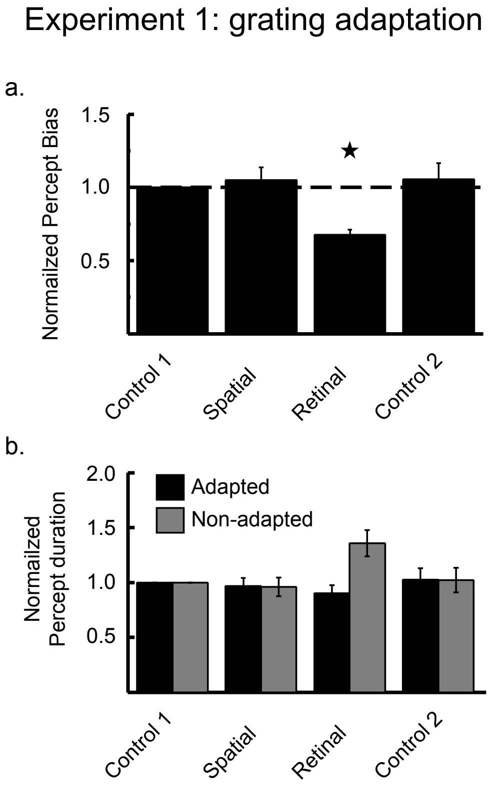


Figure VII.2 | The influence of grating adaptation on grating rivalry. (a) *Normalized percept biases (i.e. average percept duration of the adapted stimulus divided by the average percept duration of the non-adapted stimulus, normalized by the bias of Control 1). Adaptation only influenced the bias for retinotopically-matched locations.* (b) *Average normalized percept durations per percept-type (i.e. adapted and non-adapted stimulus), normalized relative to Control 1. The effect of adaptation is evident as an increase in percept durations for the non-adapted stimulus (the conditions retinotopic non-adapted percept duration is significantly different from 1, $p < 0.01$, all other $p > 0.05$).*

Experiment 2—house-face rivalry

Consistent with what was observed with grating adaptation in Experiment 1, house and face adaptation strongly influenced rivalry at retinotopically-matched locations (averaged over face and house adaptation: $p < 0.0002$; see Figure VII.3). More interestingly, adaptation to house and face also influenced house-face rivalry occurring at the spatiotopically-matched location ($p < 0.002$), although the spatiotopic effect was smaller than the retinotopic effect ($p < 0.01$, paired t-test). Importantly, the Control 2 condition (see Figure VII.1b) did not show any systematic bias. As this control condition was located the same distance from the retinotopic location as the spatiotopic condition, the absence of bias in Control 2 indicates that the observed bias in rivalry competition at the spatiotopic location is not caused by adaptation that may have spread from retinotopic adaptation (for example, due to the large receptive fields of face-selective cells).

We also calculated the time course of the influence of pre-adaptation on house-face rivalry. This is plotted in Figure VII.4 (with trial-time binned into three intervals) as the cumulative dominance duration of the adapted stimulus divided by the cumulative dominance duration of both stimuli. This analysis reveals that retinotopic adaptation has a large initial effect that gradually decreases over the course of the trial, whereas spatiotopic adaptation has a smaller effect but one

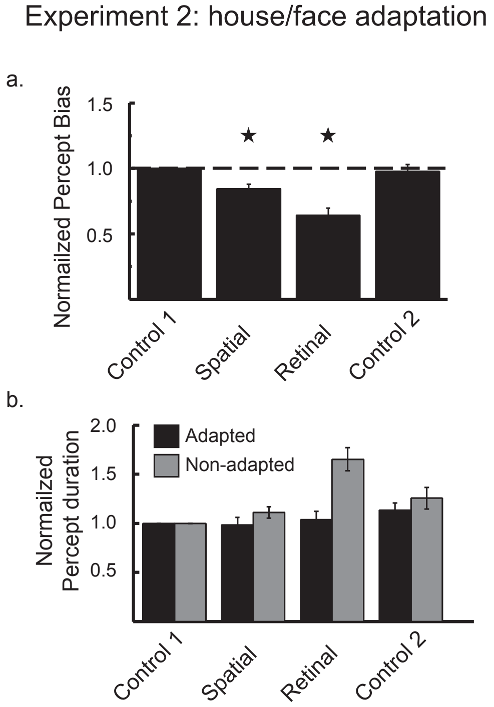


Figure VII.3 | The influence of house or face adaptation on house-face rivalry. (a) *Normalized percept biases (see Figure VII.2a), averaged over house and face adaptation conditions. Adaptation influenced the bias for both retinotopically-matched and spatiotopically-matched locations.* (b) *Average normalized percept durations per percept-type (i.e. adapted and non-adapted stimulus), normalized relative to Control 1. The effect of adaptation is evident as an increase in percept durations for the non-adapted stimulus. This effect is significant for spatiotopically-matched ($p < 0.05$), retinotopically-matched ($p < 0.0003$), and Control 2 ($p < 0.03$; the adapted stimulus shows a nearly significant increase ($p = 0.051$) which is why the ratio of the two leads to no significant change (see* (a)*).*

that is consistently present throughout the trial.

Discussion: Experiments 1 & 2

Experiment 1 confirmed a previous observation [40] that preadaptation to one of the stimuli from a rival pair will bias subsequent rivalry to favour the other, non-adapted stimulus. Using grating stimuli, the comparison between retinotopic and spatiotopic conditions clearly revealed that the effects of grating adaptation on rivalry are limited to the retinotopically-matched location. For rivalry occurring at a spatiotopically-matched location, preadaptation to a grating has no influence on rivalry.

These results contrast with the results obtained in Experiment 2, which examined rivalry between more complex stimuli that were recognisable visual objects—a house and a face. In this case, while we again observed that preadaptation led to a rivalry bias favouring the other stimulus when rivalry was tested at the retinotopically-matched location, we also found a rivalry bias when rivalry was tested at the spatiotopically-matched location. In both cases the direction of the bias was the same—towards the unadapted stimulus. This bias could be due to either a decrease in dominance of the adapted grating, or an increase in the dominance of the non-adapted grating. Plotting the average percept durations of both competing patterns separately for both experiments (see Figures VII.2b and VII.3b), revealed that the bias resulted from an increase in the dominance durations of the non-adapted grating. This result is consistent with the finding that

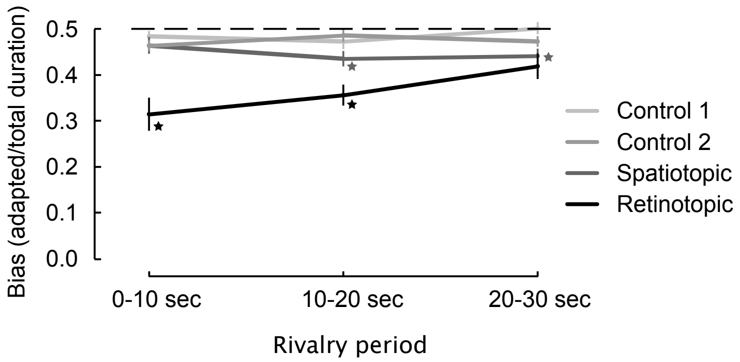


Figure VII.4 | The time course of adaptation influences during house-face rivalry. *Each 30-second trial has been divided into three 10-second periods (x-axis), and average predominance of the adapted stimulus has been calculated for each period. Adaptation has a strong retinotopic influence that gradually decreases over the course of the trial. The influence of adaptation on the spatiotopically matched rivalry location is weaker than for retinotopically matched locations, but is nevertheless sustained throughout the trial. No significant biases are found for the two control conditions. Stars indicate a significantly lower bias than 0.5 (one-tailed t-test with bonferroni correction).*

adaptation reduces the sensitivity of the adapted neuronal population [45], which may be compared to physically reducing the stimulus contrast [40], which in turn is known to increase the dominance duration of a non-adapted stimulus [153]. By disentangling the retinotopic and spatiotopic components of adaptation, we show that this finding is true for both retinotopic (Figure VII.2b) and spatiotopic (Figure VII.3b) adaptation.

Experiment 3—face-grating rivalry

In Experiment 3, we test for a possible influence of feedback on the biased competition. As mentioned in the Introduction of this chapter, rivalry is generally thought to be initiated at early visual levels and subsequently enhanced at higher levels of processing. This formulation does not explicitly invoke feedback from higher to lower levels as an important contributor to the rivalry process, however, several studies have suggested that feedback during rivalry is important [3, 7, 8, 281]. We sought to investigate the importance of feedback signals in establishing the biased competition by adaptation.

In principle, the effects we observed in Experiments 1 and 2 could be explained in one of two ways, one which requires feedback to early processes and one which does not. According to the 'no feedback' account, it would be assumed that low-level rivalry and high-level rivalry exist independently, and that the observed bias in rivalry following spatiotopic house-face adaptation was strictly due to high-level interactions between neuronal pools sensitive to houses and faces respectively. An alternative account involving feedback assumes that even though high-level areas processing visual objects may contribute to rivalry, their role is to feed back to early inhibitory processes which are the primary drivers of rivalry [7, 8]. On this account, the rivalry bias that arises following house-face adaptation would be caused by decreased feedback to early visual levels from the adapted high-level neuronal pool.

In order to investigate the importance of feedback, we instigated rivalry between a face and a grating stimulus and measured separately the effect of preadaptation to the face and the grating. The mismatch in stimulus complexity should preclude high-level rivalry between visual objects, but maintain low-level rivalry due to the abundant local image conflict [7]. We therefore expect that any rivalry bias following face adaptation would be due to an attenuation of feedback signals from extrastriate face-selective neurons. This should be especially so when rivalry is tested at the spatiotopically-matched location, as differences between the stimuli in low-level retinotopic adaptation would have no influence in this condition.

As expected from Experiment 1, adaptation to grating stimuli only biased face-grating rivalry at the retinotopically-matched location (Figure VII.5a). However, for the same rivalry stimuli, adaptation to face stimuli caused a bias for both retinotopic and spatiotopic conditions (Figure VII.5c). These findings indicate that feedback from spatiotopic areas, activated by the face stimulus, makes an important contribution to the rivalry process at earlier visual areas. As observed before in Experiment 1, the retinotopic bias following grating adaptation was caused by an increase in dominance duration of the non-adapted grating (Figure VII.5b). However, following face adaptation, the rivalry bias at the spatiotopically-matched location which we attribute to attenuated feedback was primarily caused by a decrease in the duration of the adapted face stimulus.

VII.4 General Discussion

Accumulating evidence suggests that binocular rivalry is a multi-level process [38], although much remains unknown about the degree of interaction and independence between these different levels. Preadaptation provides a useful tool in this respect for two reasons. First, adaptation is a key factor in determining the dynamics of the rivalry process [40, 41, 146, 204, 284], and second, the neural loci of selective adaptation to a large range of stimuli are far better understood than those of the rivalry process (see Introduction). There is therefore great potential for adaptation protocols to elucidate the neural processes underlying rivalry. In this report, we have shown that preadaptation to simple and complex stimuli affect rivalry dynamics in different ways. Specifically, we have shown that the effects of adaptation to simple grating stimuli on rivalry dynamics are limited to

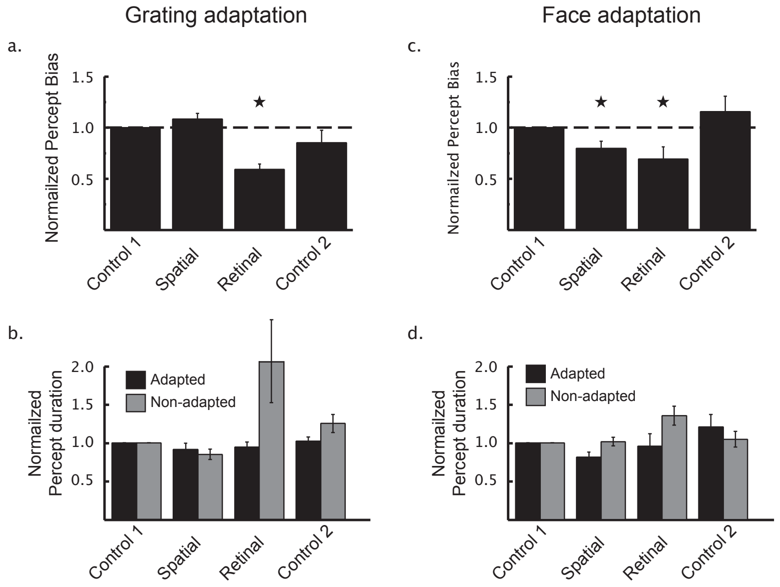


Figure VII.5 | Influences of grating adaptation and face adaptation on face-grating rivalry. (a) *Grating adaptation influences only retinotopically-matched rivalry stimuli (cf. Figure VII.2a).* (b) *The grating adaptation mainly affects the dominance durations of the non-adapted stimuli (Non-adapted: spatial: $p < 0.05$; retinal: $p = 0.055$, ns; control 2: $p < 0.05$; other $p > 0.15$).* (c) *Face adaptation influences rivalry through feedback both for retinotopically- and spatiotopically-matched locations (cf. Figure VII.3a).* (d) *For spatiotopically-matched locations the dominance durations of the adapted stimulus are reduced ($p < 0.02$), while for retinotopically-matched locations the dominance durations of the non-adapted stimulus are lengthened ($p < 0.02$). All other $p > 0.05$.*

retinotopically-matched locations, whereas the effects of adaptation to more complex stimuli such as houses and faces, are observed spatiotopically[2] (in addition to retinotopically) .

We found that retinotopic and, when it occurs, spatiotopic adaptation have qualitatively similar influences on rivalry dynamics. Retinotopic adaptation predominantly influences the dominance durations of the non-adapted stimulus, and we show that this rule also holds for spatiotopic adaptation, indicating that these processes may operate through similar mechanisms. Interestingly, however, the time courses of these adaptation influences on rivalry differ. As shown in Figure VII.4, the effect of retinotopic adaptation was initially strong but dissipated over time, although it remained stronger than the spatiotopic effect even during the final third of the rivalry observation period (i.e. 20-30 seconds after adaptation). In contrast, the spatiotopic effect is smaller but still significant and remains

---

[2] Note that the spatiotopic adaptation effect we show for faces is not necessarily linked to the spatiotopic adaptation previously reported for face-identity [185].

unchanged over the course of the 30-second trial. This suggests that adaptation not only builds up more quickly for complex stimuli (as suggested in Ref. [7]), but may also last longer than adaptation to simple stimuli.

The origin of the spatiotopic adaptation

The spatiotopic effects of adaptation we report have at least two possible origins. First, we may have measured a truly spatiotopic effect, in that the adaptation is linked to the spatial position at which it occurred before an eye-movement was made. A second possibility, however, is that the spatiotopy we observed could be artifactual, being simply a consequence of the large receptive fields in higher visual areas. According to the second hypothesis, the spatiotopic effects we measured during the test period would arise from the same neurons adapted during the preadaptation phase but are simply expressed at a different position in their receptive field. Neurophysiological data show that receptive-field sizes are far larger for cells in higher visual areas responsive to visual objects than they are for orientation-selective cells in V1 [89]. It would therefore appear to be consistent with the second hypothesis that we found spatiotopic adaptation effects were greater for house and face stimuli than for gratings. However, we can discount this possibility on the basis of our Control 2 condition (see Figure VII.1a). In this condition, we measured adaptation effects at a location that was the same distance from the retinotopic location as the spatiotopic test location was. The second hypothesis would therefore predict the same adaptation effects should occur at both locations. However, we only found adaptation effects at the spatiotopic location (Figure VII.3a) and we therefore believe that this represents a truly spatiotopic adaptation effect.

In order to obtain a spatiotopic representation, a decoupling from retinal stimulation needs to occur. Such decoupling requires information from extra-retinal sources. In our experiment, the only source of extra-retinal information comes from eye-movement related signals. One of two main sources of extra-retinal information is the efferent copy, which is a copy of the motor command that instructs the muscle to act. This copy of the motor command is thought to inform sensory areas of upcoming changes in visual input due to self-motion [278, 290]. Neurons in these areas may react by changing the position of their receptive fields (see below). The other main source is proprioceptive information, which provides postural information, e.g. the position of the eye in the head. Proprioceptive information is necessarily more sluggish than the efferent copy, but eye-position signals can still provide useful information in certain conditions such as in depth perception (see e.g. Ref. [25]). Indeed, eye-position signals could have played a role in our experiments because the adaptation and test phases were relatively long, although the efferent copy from the eye movement is likely to be the most potent source informing spatiotopy (see e.g. Ref. [113]).

Eye-movements also give rise to many other effects. For example, neurons in a multitude of cortical areas are known to change their firing patterns and even their receptive field locations (e.g. Refs. [13, 74, 278]) as a consequence of saccadic eye movements. These receptive field changes occur just before the start of a saccade and are important in spatiotopy because they ensure space is remapped from the initial eye position to the new position, so that spatial correspondence is maintained across saccades [74]. If a neuron whose receptive field is remapped in this way is preadapted prior to the eye movement, then its adapted state should be able to be tapped equally well in its new (spatiotopically-matched) receptive-field location. Interestingly, the remapping of receptive fields seems to be much stronger in higher visual areas than in lower areas, and is nearly absent in V1 [195]. This ties in well with our finding that spatiotopic adaptation occurred for complex stimuli such faces and houses, but did not occur for simple grating stimuli.

The influence of feedback

In Experiment 3 we looked more closely at the origin of the spatiotopic effects of adaptation. As noted above, rivalry may take place at multiple levels in the visual system and one of our motives in this study is to find out how these levels interact. The results from Experiment 2 were ambiguous regarding the origin of house-face rivalry: adaptation could have biased rivalry at a high level (where rivalry-related activity changes have been measured [253]) by specifically desensitizing one of the two competing neural populations representing the visual objects, or it could have biased low-level rivalry as a result of decreased feedback from the desensitized high-level area [7][3]. Experiment 3 demonstrates that at least part of the spatiotopic, and likely also retinotopic, influences are caused by feedback from high- to low-visual areas. We have suggested previously that feedback helps in coordinating perceptual coherence over many local rivalry zones [7] and specifically biases perceptual interpretations that are consistent with context [3]. Experiment 3 shows that feedback not only provides an interpretation bias, but that this feedback bias is stable despite the occurrence of eye-movements. This mechanism would provide the visual system with a cost-effective way of processing ambiguous information: instead of trying to solve ambiguities anew after each eye-movement, it could quickly determine whether the previous interpretation could be maintained before searching for alternatives.

In line with previous findings showing influences of feedback from spatial context (see e.g. Ref. [38]), we find that the feedback imposed by temporal context (i.e. adaptation) specifically influences the time of suppression of the adapted stimulus (Figure VII.5d). This adaptation effect should be most evident (as it is in our data) when testing at a spatiotopically-matched location, as this would

---

[3] The final logical alternative—that rivalry is entirely early and local—can be excluded because of observations such as high-coherence and deep suppression of rivaling complex images [7, 8, 198] and the spatiotopic adaptation we report above.

be free from non-feedback related effects such as retinotopic adaptation. However, feedback cannot be the only factor involved in the rivalry process because it operates mainly on the dominance durations of the adapted stimulus, while the adaptation bias mainly affected the dominance of the non-adapted stimulus (see Experiments 1 and 2). It appears therefore that the rivalry system can operate independently of any feedback influence when stimuli are simple, presumably as inhibitory interactions between competing pools of early neurons. In addition, when the stimuli are complex visual objects or when global context is provided, feedback from higher-level areas can influence these low-level interactions. However, because it affects mainly the adapted stimulus rather than the unadapted one, feedback appears to operate through a different mechanism (see also Ref. [38]).

Chapter VIII

Discussion: Implications for the visual architecture

VIII.1 Preamble to the discussion

The preceding chapters were concerned with two relatively different, but widely studied, topics in the vision sciences: motion processing and binocular rivalry. Although well-studied, the architecture of the part of the visual system that underlies these processes is still not completely understood. The studies in this thesis were specifically aimed to unravel the architecture of neuronal networks that support motion perception and binocular conflict resolution. The advancement in the knowledge of the neuronal architecture will certainly help to understand how our visual system works.

Both the chapters on motion perception and binocular rivalry focussed on proposed divisions in the neural architecture that were invoked to explain apparently dichotomized data-sets. In the motion studies, the proposed division in fast and slow speed-sensitive systems was scrutinized. In the binocular rivalry studies, the focus was on the division in eye-based versus pattern-based explanations, and the difference between dichoptic masking and binocular rivalry. In all cases, evidence was presented in this thesis that showed that a single system or mechanism may suffice to explain the previously-presented data. The finding that a single system account suffices, need not necessarily imply that the visual system follows this most parsimonious account. Nevertheless, I suggest that we may use the data in this thesis to support the use of a single-system account as a "null-hypothesis" for future experiments on speed perception, and find conditions that require changes or additions to this view.

In the following, I will briefly review the findings described in this thesis, and I will present descriptive models of how I view the architecture of the visual system underlying the perception of motion and binocular rivalry. These models are based on the findings presented in this thesis in combination with data reported in the literature.

VIII.2 Motion processing

Motion is described by a direction and a speed of displacement. As discussed in Chapter II and Chapter III, one has come to the general agreement that motion direction is represented in a continuous manner in the brain [178, 286, 12, 175]. However, such a general agreement is absent in research concerning speed perception. It has been suggested that at least two speed-tuned systems exist: one for high speeds and one for low speeds [75, 130, 236]. Nevertheless, neurophysiological findings suggest that speed could be represented in a continuous fashion as well [156, 182].

One of the lines of research that suggested a division in fast and slow speed system involves motion aftereffects (MAE). Previous psychophysical research on MAEs showed that: (1) slow motions led to MAEs on static test displays, while fast motions led to MAEs on dynamic displays [236, 261], (2) slow and fast MAEs could be elicited simultaneously with an appropriate test stimulus [236], showing some form of independence, (3) two slow or two fast motions did not lead to two separate MAEs, but instead to one integrated MAE [236, 261]. On the basis of these data, it was proposed that fast and slow motion systems exist. It was furthermore proposed that within each system all direction information are integrated, but between systems not, and that the different systems are read out with different test stimuli.

The hypothesis of a bipartite division would make the following predictions:

- the occurrence of the MAE after adaptation to slow or fast motion is strictly dependent on the type of test stimulus,
- a maximum of two different speed-dependent MAEs can be induced with a single adaptation stimulus,
- when three different refresh frequencies for the test stimulus are used, one should obtain two, and not three, peaks of maximum-duration MAEs in the speed domain. These peaks should be positioned at the point of maximum sensitivity of the two motion systems, and
- independently of the speed differences, discontinuous transitions from slow to fast MAEs should exist, when the motions fall in the ranges activating different motion systems,
- only a mechanism consisting of two independent motion systems can produce discontinuous MAE data.

It turns out that these assumptions were not supported by the findings in Chapter II. We found that a medium-speed adaptor motion (with a speed falling in the fast motion range) can induce a MAE on both static and dynamic test patterns, depending on the accompanying adaptor motion. We showed that at least three MAEs could be elicited after adaptation to a single stimulus. Further, we showed that the maximum MAE duration was found at different adaptor speeds for different test patterns. The position of the peak MAE-duration depended on the temporal frequency content of the test, and not strictly on the maximum sensitivity of the putative motion systems. Furthermore, continuous transitions from 'slow' to 'fast' MAEs were found when adaptation took place to motions that lay in the 'fast' and the 'slow' domain, but were close in speed content. These experimental findings speak against a division of the motion system into a subsystem sensitive to fast speeds, and another subsystem sensitive to slow speeds. With a computational model (Chapter II), we found moreover that a single continuous system could account for our findings as well as those previously reported in the literature, suggesting that a single continuous system underlies our local speed processing. In Chapter III, it was shown that data from 'detection-in-noise' paradigms could as well be described with a single-system model. Contrary to previously published suggestions [75, 130], we showed that coherent-motion detection in noise depends in a continuous fashion on the overlap between the neural representations of noise and signal speed, and not on the existence of independent fast and slow motion systems.

The distinction between local and global motion processing in psychophysical approaches

Motion processing is generally divided into two distinct processes: local and global motion processing [193]. Local motion processing considers a relatively small patch of the visual field (<1 degree of visual angle), while global motion processing analyzes motion over a larger area of the visual field (∼5 degrees of visual angle, or more).

In the introduction of this thesis, Chapter II was introduced as being an investigation on local motion processing. This denomination is probably not entirely justified. In principle, local motion signals were unambiguous in this study, because the employed motion stimuli were not noisy, and motion information was identical over the stimulated visual field. However, false matches between successive motion frames will lead to a noisy motion signals, and furthermore, it seems likely that the visual system will always try to integrate motion signals over large areas to reduce the impact of these noisy signals. Moreover, we were investigating speed coding, which is likely performed in the global-motion processing area MT [193, 224, 244]. (The local motion processing area V1 does not seem to be tuned to speed as such, but to temporal frequency [215, 250, 81].) These points suggest that the global motion system was involved in the processing of the stimuli used in Chapter II.

Similarly, the global motion system builds on signals from the local motion system, making it difficult, if not impossible, to investigate the global motion system in isolation with psychophysical approaches.

This comparison highlights the difficulties in studying the local and global motion systems independently with psychophysical means. Therefore, even though the term local-motion processing has been used to introduce the findings in Chapter II, and the term global-motion processing for Chapter III, one should not interpret these denominations literally.

The single speed system account, and multi-level motion processing

Motion is processed at several levels in the visual system, and each of these levels makes its own contribution to our perception. Therefore, even though the single-speed-system account is tenable, a one-level account of motion perception is probably not.

In fact, the MAE literature, apart from providing data suggesting a division in the speed dimension, also suggests the involvement of multiple stages in the formation of the MAE. The phenomenon that reveals the different stages is called MAE storage. Generally, the MAE is probed with the presentation of a test stimulus just after the adaptation period has ended. However, when a test stimulus follows a long blank period (in the order of the normal MAE duration), the MAE is almost as long as when the test stimulus would be provided directly following the adaptation [262], i.e. the MAE is stored. When the blank period is replaced with a dynamic test pattern, a MAE will not only be perceived on this dynamics display, but also on a subsequently shown static test pattern. In fact, the MAE duration on the static test pattern is hardly influenced by the intervening dynamic pattern. Conversely, when a static test pattern replaces the blank, a MAE will be perceived on this static display, but not on a subsequent dynamic test [262].

How to explain this asymmetry in MAE storage? It is thought that static test patterns activate low-level mechanisms, while dynamic test patterns are thought to activate higher-level mechanisms [201]. We may assume that the static test pattern activates neurons within the primary visual area (V1; and indirectly the higher visual Medial Temporal area, MT), while the dynamic test predominantly activates neurons within MT. This assumption could explain the MAE-storage asymmetry in the following way. Adaptation due to prolonged motion viewing affects (i.e. adapts) both V1 and MT neurons. When, following an adaptation period, a MAE is induced with a dynamic test stimulus, only the adaptation in MT is read out. When a static test pattern is subsequently presented, the adaptation in V1 is read out, which would be only slightly reduced due to the preceding presentation of the dynamic pattern. The two MAEs—the dynamic and static one—will sum up to more than just the static one measured directly following adaptation. When the order of the test stimuli is reversed, the static test reads out the adaptation in V1, but also a large part of the MT adaptation. A subsequent presentation of a dynamic test pattern therefore reads out a decreased adaptation state in MT, and thus causes a much reduced MAE. In this case, the two MAEs sum up to only slightly more than the static MAE alone.

The two-stage configuration is consistent with several other psychophysical findings as well. For instance, MAEs on the static test patterns are primarily tuned to temporal frequency [211, 291], which the low-level visual area V1 is most sensitive to [215], while those on dynamic tests are tuned to speed [22], which MT is most sensitive to [193, 214, 217]. Furthermore, the MAE transfers between the eyes with dynamic tests [202] (indictating a read out at the binocular level, such as MT), while it does only partially transfer for static stimuli [287, 190] (indicating a read-out at the monocular level, such as V1, and possibly a small read-out at the binocular MT-level). It may also explain the absence of inter-ocular transfer of MAEs with patternless flickering stimuli [98]. Such flickering stimuli do not contain speed information, but temporal frequency information. This is therefore more consistent with V1 than MT specialization, and thus consistent with a read-out at the monocular level and not at the binocular level. It is interesting to see that the transition between 'slow' and 'fast' MAEs occurs at about 20Hz [236], which is about the temporal frequency at which we start to see continuous motion (which is why films are often shot at 24Hz). Perhaps, this finding indicates that the balance of the adaptation read-out shifts from V1 to MT.

Note that the multi-level interpretation and the possible involvement of V1, or other low-level visual stages, in motion perception, does not negate the finding that speed (which is not represented in V1) is represented in a continuous fashion. Overall the data in this thesis are consistent with the architecture described in the introduction of this thesis. That is, the visual system consists of several levels, and several brain areas. Within each area, the neurons have a similar task (e.g. speed or temporal frequency processing), and those neurons that have an even more similar task (e.g. the processing of nearly identical motions) are positioned in each other's vicinity within each brain area, forming a continuous sheet of, for example, speed sensitive cells.

VIII.3 Binocular rivalry

As mentioned in previous chapters, there exist two main theories of binocular rivalry. One theory assumes that binocular rivalry depends on inhibitory interactions between monocular neurons (i.e. interocular rivalry), the other assumes that rivalry depends on inhibitory interactions between pattern representations (i.e. pattern-based rivalry). Evidence for both theories will be briefly discussed, followed by a discussion of the additions of this thesis to the research field of binocular rivalry. I will end with a discussion of a descriptive model that explains the results presented in this thesis, and those in the literature, in terms of pattern-based rivalry, and eye-based suppression.

Why interocular rivalry?

Studies using functional Magnetic Resonance Imaging, reported that in early visual areas [216] (and even specifically monocular areas [252, 109, 292]) modulations in blood-oxygen-level dependent (BOLD) response—a measure for brain activity—are tied to the percept of the subject, suggesting that monocular low-level areas are involved in the maintenance of binocular rivalry.

A second line of evidence is based on a frequently employed psychophysical measure of the strength of rivalry: depth of suppression. Depth of suppression is measured by presenting a probe stimulus during rivalry in either the dominant or the suppressed eye. The ratio of the probe's detection (or identification) thresholds of the suppressed over the dominant eye is the suppression depth. The probe is generally harder to detect (or identify) in the suppressed compared to the dominant eye, and the depth of suppression is generally about 0.6 (e.g. Refs. [199, 82, 83]). Importantly, the suppression depth seems to be independent of the pattern similarity between the conflicting stimuli and the probe, suggesting that suppression depth is independent of pattern information, and instead primarily dependent on eye-of-origin information.

A third piece of evidences is that when the eyes' images are swapped immediately after the establishment of perceptual dominance, the percept generally 'sticks' to the dominant eye, and not to the dominant pattern, i.e. a stimulus swap is perceived. Again, this finding is indicative of suppression based on eye-information, and not on the pattern-information [42].

Other evidence includes the finding that the course of rivalry is strongly influenced by low-level stimulus characteristics such as contrast and luminance that do not change pattern-information [153, 280, 129].

Lastly, the area of exclusive dominance—the maximum stimulus size that supports rivalry without resulting in incomplete rivalry—increases with viewing eccentricity not unlike the cortical magnification factor characteristic of V1 [39]. Consistently, this finding is dependent on the retinal stimulus size and not the perceived stimulus size [37]. Again these data are indicative of a low-level base of rivalry interactions.

Why pattern-based rivalry?

As briefly reviewed in the introduction of this thesis, pattern-based accounts of rivalry are diverse. In its most extreme form, these accounts state that binocular rivalry is not a direct cause of inhibitory interactions between stimulus representations, but instead that it results from a reinterpretation of visual information by non-visual brain areas that, subsequently, act upon stimulus representations to make them more or less visible (see Ref. [150], and also [167, 168, 243, 223]). An alternative idea is that rivalry is caused by an inhibitory interaction between pattern-sensitive neurons [138, 198].

Direct evidence that rivalry does not depend on inhibitory inter-ocular interactions was provided by Logothetis *et al.* [159]. They showed that when the eyes' images are swapped between the eyes at a rate of 3 Hz, seemingly normal rivalry ensues between the patterns (called stimulus rivalry), and that the percept does not follow a single eye's image (which is rapidly alternating).

Further evidence for pattern-based influences on rivalry comes from studies that show inter-ocular grouping during rivalry [72, 138, 134, 246, 275]. These studies show that when pattern information is distributed between the two eyes, the percept often switches between two complete patterns, and not between two eye-images, even though the pattern information originates from different eyes.

Finally, other forms of rivalry that do not involve inter-ocular conflict—examples are monocular rivalry [173, 264, 54], the Necker cube [197], and Structure from motion [271]—may show perceptual dynamics similar to binocular rivalry [152, 76, 52].

Amalgam theory of rivalry

The wealth of information in favor for both theories makes the adherence to the extreme ends of the spectrum of theories untenable. Furthermore, the evidence in favor of either theory is repeatedly shown to be dependent on specific parameter settings. For example, (1) stimulus rivalry was reported to be confined to a small parameter space [143], but as I have shown in Chapter V this parameter space may be increased with some specific temporal stimulus characteristics. (2) Inter-ocular grouping during rivalry may be partly explained by the observers' change in criteria for what is considered to be complete dominance [144]. (3) Although the percept sticks to the eye and not the stimulus when the eyes' images are swapped during a dominance phase, this effect is dependent on the monocular interaction of the two competing stimuli (see ChapterV). And (4) even though depth of suppression is not influenced by stimulus parameters, the duration of dominance after the presentation of a probe is affected by pattern-information for some stimulus parameters [270].

Given the strong evidence for, and the limitations of, both eye-based and pattern-based processes, researchers nowadays increasingly favor an amalgam theory of rivalry that includes both processes (e.g. Refs. [38, 284, 86, 8]). In the amalgam theory of rivalry, binocular rivalry is initiated by inter-ocular conflict and operationalized by interocular inhibition between monocular neurons. Pattern information contributes to rivalry by influencing the strength of inhibitory interactions between monocular neurons or by increasing the perceptual suppression by means of direct inter-pattern conflict [38, 8, 198].

The additions of this thesis

This thesis has provided additional data that may help to determine where in the brain, and in what way, binocular rivalry is initiated and maintained. Most of the findings reported in this thesis build on the finding that binocular rivalry may proceed even when the two competing stimuli are presented without temporal overlap [209]. The repetition of the competing stimuli may be delayed up to ∼350 ms before rivalry breaks down [209]. It has not yet been determined which stimulus characteristics underlie this phenomenon. In Chapter IV, it was shown that low-level stimulus characteristics, such as luminance, the inter-ocular delay, contrast polarity and eye-of-origin, and even motion, are not determining whether rivalry occurs or not. Instead, we have argued that the continued presence of rivalry in the the absence of direct spatio-temporal conflict, is dependent on pattern-information.

The conclusion was subsequently strengthened by the finding, presented in Chapter V, that monocular and stimulus rivalry—both dependent on pattern conflict—are also limited by a 350-ms pattern-based limit. This chapter also showed that monocular and stimulus rivalry are curtailed when the competing stimuli are presented within a ∼50 ms monocular interaction-window. Chapter V, therefore, reconfirmed the pattern-information influences on rivalry, and showed the importance of two temporal limits on rivalry: one narrow ∼50 ms monocular window, and one wide ∼350 ms binocular window.

Chapter VI provided further support for the theory of pattern-based rivalry by more extensively studying the independence of binocular rivalry of inter-ocular delays. Furthermore, this chapter suggested that binocular rivalry and dichoptic masking, both very potent tools in manipulating the visibility of potentially salient stimuli, may depend on similarly-operating or even identical neural substrates.

Together, these results primarily add weight to the pattern-based account of rivalry, at the expense of eye-based accounts of rivalry.

Finally, Chapter VII investigated the communication between high and low-level rivalry mechanisms. We reported that feedback strongly influenced the course of rivalry, and that feedback allows for the spatiotopic influences of adaptation on binocular rivalry. This chapter shows that we may investigate the communication between different levels in the visual architecture involved in binocular rivalry by employing adaptation protocols.

A multi-level pattern-based account of rivalry

Based on the findings reported in this thesis, and the current literature, I will unfold a theory of rivalry that is multi-leveled and essentially pattern-based. According to this theory, rivalry takes place at a binocular pattern-sensitive level, while suppression takes place at a monocular, eye-based, level. Suppression, in this theory, is the result of regulatory feedback from the binocular rivalry-stage.

Rivalry at a binocular pattern-sensitive level

Following a reductionist approach, I have come to the believe that interocular interactions are not key to binocular rivalry. Following this approach we could ask: What are the minimal conditions to obtain rivalry without affecting the basic perceptual properties of the phenomenon? In order to answer this question we may take a classical binocular rivalry stimulus and remove a feature from this stimulus. If this removal is not affecting the occurrence nor the perceptual dynamics of rivalry, we may conclude that the removed element was not necessary for rivalry.

For example, a classical binocular rivalry stimulus involves the presentation of an 'S' in one eye, and an 'A' on the corresponding spot in the other eye [279]. This stimulus contains both rivaling meanings—the letter 'S' versus the letter 'A'—and local orientation conflict. Another classical rivalry stimulus consists of a set of lines in one eye and an orthogonal set of lines in the other eye [54]. This stimulus contains only orientation conflict. Both rivalry set-ups lead to binocular rivalry, and the conclusion one may draw is that stimulus conflict is necessary to initiate rivalry, while conflicting object meaning is not.

Following a similar procedure, we may test whether it is indeed necessary to have inter-ocular conflict to induce rivalry. The simple answer would be 'no', as when we place two conflicting stimuli in a *single* eye, the stimuli fluctuate in perceptual salience, under some conditions even leading to complete perceptual suppression and dominance [54]. However, this type of rivalry is far less compelling than the classical form of binocular rivalry, and the alternations between percepts are generally much slower [251], suggesting that it may involve mechanisms different from those involved in binocular rivalry. The stimulus rivalry paradigm [159] provided stronger evidence against the necessary involvement of inter-ocular conflict in rivalry, because the rivalry for this type of stimulus showed rapid alternations and strong perceptual suppression [159]. However, this type of rivalry was subsequently shown to occur in a rather limited range of parameter settings [143], and moreover, inter-ocular influences remain possible, because at any time the two eyes' views were in conflict.

Chapter V provides final evidence against the need for interocular interactions to obtain the fast dynamics and strong perceptual suppression of binocular rivalry. A completely monocular stimulation protocol allowed for the instigation of rivalry over a wide range of stimulus parameters, while at the same time showing dynamics typical for binocular rivalry. In fact, when monocular interactions between the two rivalrous stimuli were prevented, by introducing blank periods, the differences between binocular rivalry, stimulus rivalry and monocular rivalry were statistically non-significant.

These data strongly suggest that inter-ocular conflict is not necessary for (binocular) rivalry, and thus that an eye-independent (i.e. pattern-based) rivalry mechanism is operational. This view is strengthened by research showing that low-level stimulus characteristics, such as luminance, contrast, and interocular delays, have a very limited influence on whether rivalry will occur or not (Refs. [33, 157, 209] and Chapters IV, V and VI), as long as low-level conflict is present [59, 33]. This thesis, furthermore, showed that rivalry involves a pattern-integration window of about 350 ms. We showed that when the presentation of both conflicting patterns is repeated within this window, rivalry may proceed, independent of, for example, the eye of origin of stimulation.

Suppression at a monocular level

The position of the suppression-process at early stages within the visual system is supported by the finding that manipulations of low-level stimulus characteristics, such as contrast, luminance and eye-of-origin information, may shorten and even break the perceptual suppression (e.g. Refs. [153, 33, 210, 42, 269]). At the same time, suppression also influences information processing at low-level visual areas. Recent findings—which contradict earlier conclusions [36, 148, 208, 266, 63, 141]—show that perceptual suppression influences the build-up of low-level visual aftereffect and afterimages [257, 96, 44]. These data place the suppression-process at early visual stages, before, or at the level of adaptation and after-image formation (see also Ref. [265]).

Even more convincing evidence for the position of the suppression-stage at early monocular levels, is the observation that the perceptual suppression during conventional binocular rivalry is clearly linked to a single eye. Eye-based suppression is perceptually apparent through the complete phenomenological disappearance of one of the two competing stimuli. Indeed, experimental measurements, probing stimulus detection thresholds during perceptual dominance and suppression, have shown that a probe is harder to detect in the suppressed eye than in the dominant eye (e.g. Refs. [199, 82, 83]). This finding is independent on the similarity in pattern information between the probe and the rivaling stimuli, and therefore suggests that all the information originating from the suppressed eye was weakened. Since the conflict only involves two simple line stimuli, while a rich set of probes is suppressed, the establishment of complete monocular suppression requires a spread of suppression though the entire monocular layer of the suppressed eye (both in spatial dimensions, and non-spatial, e.g. orientation, dimensions). Indeed, possible footprints of this spread, i.e. traveling waves of suppression, are observed in spatial coordinates [283].

With conventional binocular rivalry stimuli, shutting off an eye (i.e. a monocular channel) leads to the perception of a single stimulus. On the other hand, in the stimulus rivalry protocol such a process would lead to the percept of a rapid and regular alternation in orientations. Indeed, this type of percept is reported with the stimulus rivalry protocol without blanks (Exp 1, Chapter V). The blocking of a monocular channel would also predict that the percept will 'stick' in the dominant eye when the eyes' images are swapped upon the establishment of perceptual dominance (Exp. 5, Chapter V, and Ref. [42]). The entrapment within a monocular channel may be prevented, and rivalry reports increased, when a blank periods of about 50 ms was introduced before the stimulus swap (Chapter V). An intriguing possibility is that this effect was obtained because the introduced blanks curtailed the spread of the suppression over the monocular layer by breaking the spatio-temporal proximity of the monocularly-interacting stimuli. Circumstantial evidence supporting the suggestion that blanks prevent the spread of suppression is the observation that traveling waves and piecemeal rivalry are less frequent in the stimulus rivalry paradigm, which includes a 19-Hz flicker, than during conventional binocular rivalry [159].

How pattern-based rivalry could cause monocular suppression

Taken together, the data suggest that the stage where suppression takes place differs from the stage where rivalry takes place (see also Ref. [35]). *Suppression* appears to be a low-level, and eye-based process, whereas *rivalry* appears to take place at a binocular pattern-based level.

How could this arrangement be operating? First, upon stimulation with rivalrous stimuli, both monocular and binocular levels within the visual system are activated (see "Excitation", in Figure VIII.1). Subsequently, the rivalry stage determines which stimulus will be dominant and which will be suppressed, and feeds back this output into a lower monocular stage, where it causes suppression ("Suppresion", in Figure VIII.1). One of the ways in which the feedback from the rivalry stage to the suppression stage could take place is through pattern-specific inhibitory feedback (labeled 1, in Figure VIII.1). The pattern-invariant suppression that is generally found in binocular rivalry may be explained by a subsequent non-specific spread of inhibition (or a decrease in stabilizing cooperation [87, 88]) among neurons through a large part of the monocular layer (labeled 2 in Figure VIII.1). An alternative route for feedback-based suppression could involve decreased positive feedback from the suppressed binocular neuron to neurons with a similar pattern-preference in the monocular layer, which weakens this patterns monocular representation.

With the current descriptive model we cannot be certain whether these forms of feedback will cause the types of results presented in the last few sections. To that end, computational models will need to be developed. However, from this discussion it should be clear that the presented theory is a genuine option.

Binocular rivalry, stimulus rivalry, and monocular rivalry in a single framework

The current arrangement places the phenomena of binocular rivalry, stimulus rivalry and monocular rivalry in a single framework. In this framework the rivalry stage, which is binocular, is the same for the three modes of rivalry. The

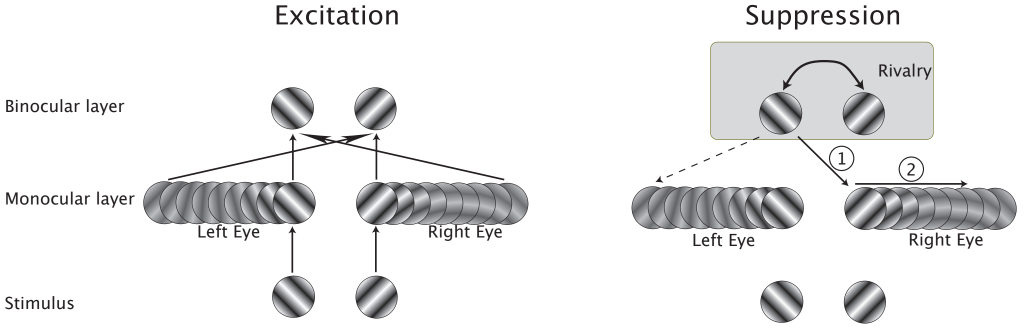


Figure VIII.1 | A model for binocular rivalry. (Left pannel) *Excitation paths. The stimulus excites monocular neurons that are sensitive to the presented stimulus. These monocular cells subsequently excite binocular cells based on pattern information.* (Right panel) *Suppression paths. The rivalry stage determines the dominant stimulus, which then inhibits, through feedback, monocular neurons that have different orientation preferences (labeled 1). The inhibition, or a decrease in neural cooperation [87, 88], subsequently spreads through the monocular layer, causing a pattern independent, but monocular, suppression (labeled 2).*

perceptual differences among these modes of rivalry are caused by monocular interactions between the conflicting stimuli. Consequently, when the monocular interactions are prevented, the three modes of rivalry will show similar perceptual dynamics (Chapter V).

This arrangement produces several key characteristics of rivalry. First, the pattern-specific feedback followed by pattern-invariant spread of suppression through the monocular layer, can explain the finding that the depth of suppression is invariant to the pattern of the probe that is used (e.g. Refs. [199, 82, 83]). Second, the pattern-invariant suppression which is linked to a single eye, explains why perceptual suppression is strong for binocular rivalry conditions, and weak for classical monocular rivalry stimuli (e.g. Ref. [251]). This result follows, because in the former case, the eye-based suppression will suppress only the to-be-suppressed stimulus, while in the latter case both the 'dominant' and 'suppressed' stimulus will be suppressed. One prediction that may be derived from this observation is that suppression depth will be shallower for monocular and stimulus rivalry compared to binocular rivalry. If we assume, as suggested above, that the introduction of a blank period indeed halts the spread of monocular suppression, we may make an additional prediction, namely that suppression depth with flickering binocular stimuli should be less than with conventional binocular rivalry, and that the occurrence of traveling waves should be rare.

Pattern-based rivalry, but how?

The suggestion that rivalry is not based on inter-ocular conflict has often automatically led to the idea that rivalry is pattern-based (or stimulus-based). I have so far used this term quite loosely as well. But what is “pattern” to the visual system?

Following the reductionist approach exemplified above, we may conclude that orientation is “pattern” to the visual system, as when we remove orientation conflict from the stimulus by equating the orientations between the eyes, or by replacing them with the space-averaged luminance, rivalry ceases. According to this approach, orientation is not the only form of “pattern”. Color conflict alone [127, 138], and motion conflict alone (e.g. Ref. [8]), without the presence of orientation conflict, induce rivalry as well.

Since the stimulus characteristics such a color, motion and orientation are quite elementary to the visual system, without much overlap in terms of their physical descriptions, it seems to me that the rivalry induced by these stimulus characteristics depends on different neural substrates.

Still, one needs to be careful not to draw conclusions too fast. Motion stimuli that move in different directions in the two eyes, almost inevitably cause a local mismatch in element positions between the eyes, which may cause rivalry [158]. In a stimulus that was specifically designed to have locally matching elements, while still having conflicting motions between the eyes, motion rivalry was not observed [59]. The finding indicates that “pure” motion conflict does not lead to rivalry.

A similar confound generally exist when investigating orientation conflict, which often involves conflicts in luminance between the eyes. Here as well, it has been shown that when local orientation conflict is prevented, but luminance conflict is not, rivalry ceases [59], suggesting the importance of orientation conflict and not luminance conflict in binocular rivalry.

A potentially important but not very well investigated characteristic that may induce rivalry is contrast-polarity differences (or color-contrast polarity differences). In some stimulus configurations, opposite contrast-polarities in the two eyes leads to binocular rivalry [166, 59], while in others it does not (Refs. [118, 166], and Chapter IV). The exact spatial and temporal arrangement seems to be an important factor in determining whether rivalry will be induced or not. The influence of contrast-polarity differences, however, still needs to be thoroughly investigated.

“Pattern” will likely turn out to be a too general description of what really causes rivalry. From the discussion above, it seems that only orientation and color may be termed “pattern”. However, with careful investigations, other characteristics may turn out to be “pattern” to the visual system as well.

Higher-level influences

Higher level stimulus information influences rivalry [38]. For example, meaningful pictures enjoy an advantage over non/less meaningful pictures during rivalry [77], and meaningful words and upright faces break suppression faster than do non-meaningful words and inverted faces [124]. The influence of gestalt properties such as good continuation and common fate [6, 4, 246, 275] constitute further evidence, just as the finding that observers may be trained to influence the course of binocular rivalry through voluntary control [141].

These high-level influences on the course of rivalry are probably governed in part by feedback information from higher-level to lower-level visual areas (Chapter VII, and Refs. [3, 7]). Exactly where this feedback enters the rivalry circuit—at the level of the rivalry, or at the level of suppression—is not known. Most data shows that feedback specifically targets the dominant stimulus in rivalry [38, 241], suggesting that feedback grabs on at or after the level of rivalry. Recent results, however, have shown that the suppressed state may also be influenced by feedback information [124], suggesting that feedback grabs on before the level of rivalry. Adaptation techniques may be very useful in determining how feedback operates because for adaptation one has been more successful at describing at what level it occurs than for rivalry. For example, using a preadaptation technique [40] we showed that feedback information is involved in instantiating spatiotopic influences of adaptation during rivalry (Chapter VII). Possibly the preadaptation technique may shed light on where feedback information enters the rivalry circuit.

VIII.4 So does it take two to tango?

In the Chapters concerning motion perception, it was argued that speed processing may involve only one system, but motion processing as a whole most likely involves several processing levels. Similarly, in the Chapters concerning binocular rivalry it was argued that the rivalry process is one and the same for different kinds of rivalry, but that even in the case of orientation conflict the formation of a percept involves several levels. Motion perception and binocular conflict resolution are complex processes, and this thesis has provided insights into how the visual system is arranged to cope with these complex processes.

For the visual system as for dancing the tango, I guess, some movements may be made by one of the dancers alone, and some movements need to be made by both dancers together. Nevertheless, one should keep in mind that for most movements the tango requires the dancers to merge into a single entity. Therefore, only when we consider the dancers as a unity, we may fully appreciate the dance as a whole (i.e. motion processing, or binocular conflict resolution) , and the movements (direction/speed processing, rivalry/suppression) as its parts.

References

The bold number in front of each reference is the number with which it is cited in the main text. The bracketed numbers at the end of each reference refer to the page numbers on which the citations have been made.

1. **Adelson, E.H.** & **Bergen, J.R.** (**1985**) *Spatiotemporal energy models for the perception of motion.* J Opt Soc Am A, 2(2):284–299 [*3*]
2. **Adelson, E.H.** & **Movshon, J.A.** (**1982**) *Phenomenal coherence of moving visual patterns.* Nature, 300(5892):523–525 [*7, 32*]
3. **Alais, D.** & **Blake, R.** (**1998**) *Interactions between global motion and local binocular rivalry.* Vision Res, 38(5):637–644 [*10, 102, 106, 122*]
4. **Alais, D.** & **Blake, R.** (**1999**) *Grouping visual features during binocular rivalry.* Vision Res, 39(26):4341–4353 [*122*]
5. **Alais, D.** & **Blake, R.** (Eds.), *Binocular Rivalry* (Cambridge, MA: MIT Press, **2005**) [*9, 57, 82, 92*]
6. **Alais, D.**, **Blake, R.**, & **Lee, S.H.** (**1998**) *Visual features that vary together over time group together over space.* Nat Neurosci, 1(2):160–4 [*122*]
7. **Alais, D.** & **Melcher, D.** (**2007**) *Strength and coherence of binocular rivalry depends on shared stimulus complexity.* Vision Res, 47(2):269–279 [*10, 102, 103, 105, 106, 122*]
8. **Alais, D.** & **Parker, A.** (**2006**) *Independent binocular rivalry processes for motion and form.* Neuron, 52(5):911–920 [*9, 60, 102, 103, 106, 115, 121*]
9. **Alais, D.**, **Verstraten, F.A.**, & **Burr, D.C.** (**2005**) *The motion aftereffect of transparent motion: two temporal channels account for perceived direction.* Vision Res, 45(4):403–412 [*28*]
10. **Alais, D.**, *et al.* (**2000**) *On binocular alternation.* Perception, 29(12):1437–1445 [*9*]
11. **Albright, T.D.** (**1984**) *Direction and orientation selectivity of neurons in visual area MT of the macaque.* J Neurophysiol, 52(6):1106–1130 [*17*]
12. **Albright, T.D.**, **Desimone, R.**, & **Gross, C.G.** (**1984**) *Columnar organization of directionally selective cells in visual area MT of the macaque.* J Neurophysiol, 51(1):16–31 [*29, 39, 110*]
13. **Andersen, R.A.** (**1997**) *Multimodal integration for the representation of space in the posterior parietal cortex.* Philos Trans R Soc Lond B Biol Sci, 352(1360):1421–1428 [*106*]
14. **Anderson, S.** & **Burr, D.** (**1985**) *Spatial and temporal selectivity of the human motion*

detection system. Vision Res, 25(8):1147–1154 [*28*]

15. **Andrews, T.J.** (**2001**) *Binocular rivalry and visual awareness.* Trends Cogn Sci, 5(10):407–409 [*67*]
16. **Andrews, T.J.** & **Blakemore, C.** (**1999**) *Form and motion have independent access to consciousness.* Nat Neurosci, 2(5):405–6 [*61*]
17. **Andrews, T.J.** & **Blakemore, C.** (**2002**) *Integration of motion information during binocular rivalry.* Vision Res, 42(3):301–9 [*61*]
18. **Andrews, T.J.** & **Lotto, R.B.** (**2004**) *Fusion and rivalry are dependent on the perceptual meaning of visual stimuli.* Curr Biol, 14(5):418–23 [*67, 77*]
19. **Andrews, T.J.** & **Purves, D.** (**1997**) *Similarities in normal and binocularly rivalrous viewing.* Proc Natl Acad Sci U S A, 94(18):9905–9908 [*66, 67*]
20. **Andrews, T.J.**, *et al.* (**1996**) *Temporal events in cyclopean vision.* Proc Natl Acad Sci U S A, 93(8):3689–92 [*52, 94*]
21. **Andrews, T.**, **Sengpiel, F.**, & **Blakemore, C.**, *From contour to object-face rivalry: multiple neural mechanisms resolve perceptual ambiguity.* In: *Binocular Rivalry* (D. Alais & R. Blake, eds.) (Cambridge, MA: MIT Press, **2005**) [*76*]
22. **Ashida, H.** & **Osaka, N.** (**1995**) *Motion aftereffect with flickering test stimuli depends on adapting velocity.* Vision Res, 35:1825–1934 [*5, 113*]
23. **Azzopardi, P.** & **Cowey, A.** (**2001**) *Motion discrimination in cortically blind patients.* Brain, 124(Pt 1):30–46 [*28, 29*]
24. **Azzopardi, P.**, *et al.* (**2003**) *Response latencies of neurons in visual areas MT and MST of monkeys with striate cortex lesions.* Neuropsychologia, 41(13):1738–1756 [*28*]
25. **Backus, B.T.**, *et al.* (**1999**) *Horizontal and vertical disparity, eye position, and stereoscopic slant perception.* Vision Res, 39(6):1143–1170 [*105*]
26. **Bahill, A.T.**, **Adler, D.**, & **Stark, L.** (**1975**) *Most naturally occurring human saccades have magnitudes of 15 degrees or less.* Invest Ophthalmol, 14(6):468–469 [*62*]
27. **Bair, W.** (**2005**) *Visual receptive field organization.* Curr Opin Neurobiol, 15(4):459–464 [*61*]
28. **Baker, D.H.**, **Meese, T.S.**, & **Summers, R.J.** (**2007**) *Psychophysical evidence for two routes to suppression before binocular summation of signals in human vision.* Neuroscience, 146(1):435–448 [*82*]
29. **Barbur, J.**, *et al.* (**1993**) *Conscious visual perception without V1.* Brain, 116:1293–302 [*5, 14, 29*]
30. **Barlow, H.** & **Hill, R.** (**1963**) *Evidence for a physiological explanation of the waterfall phenomenon and figural after-effects.* Nature, 200:1345–1347 [*5*]
31. **Barlow, H.**, **Hill, R.**, & **Levick, W.** (**1964**) *Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit.* J Physiol, 173:377–407 [*1*]
32. **Barlow, H.** & **Levick, W.** (**1965**) *The mechanism of directionally selective units in rabbit's retina.* J Physiol, 178(3):477–504 [*1, 3*]
33. **Blake, R.** (**1977**) *Threshold conditions for binocular rivalry.* J Exp Psychol Hum Percept Perform, 3(2):251–257 [*118*]
34. **Blake, R.** (**1989**) *A neural theory of binocular rivalry.* Psychol Rev, 96(1):145–167 [*8, 57, 61, 69, 77*]
35. **Blake, R.** (**2001**) *A primer on binocular rivalry, including current controversies.* Brain Mind, 2:5–38 [*9, 77, 119*]

36. **Blake, R.** & **Fox, R.** (**1974**) *Adaptation to invisible gratings and the site of binocular rivalry suppression*. Nature, 249(456):488–490 [*118*]
37. **Blake, R.**, **Fox, R.**, & **Westendorf, D.** (**1974**) *Letter: Visual size constancy occurs after binocular rivalry*. Vision Res, 14(7):585–586 [*114*]
38. **Blake, R.** & **Logothetis, N.K.** (**2002**) *Visual competition*. Nat Rev Neurosci, 3(1):13–21 [*9, 10, 61, 66, 67, 73, 77, 82, 96, 103, 106, 107, 115, 121, 122*]
39. **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 [*114*]
40. **Blake, R.** & **Overton, R.** (**1979**) *The site of binocular rivalry suppression*. Perception, 8(2):143–152 [*97, 99, 101, 102, 103, 122*]
41. **Blake, R.**, **Sobel, K.V.**, & **Gilroy, L.A.** (**2003**) *Visual motion retards alternations between conflicting perceptual interpretations*. Neuron, 39(5):869–878 [*96, 103*]
42. **Blake, R.**, **Westendorf, D.H.**, & **Overton, R.** (**1980**) *What is suppressed during binocular rivalry?* Perception, 9(2):223–231 [*75, 114, 118, 119*]
43. **Blake, R.**, **Yang, Y.**, & **Westendorf, D.** (**1991**) *Discriminating binocular fusion from false fusion*. Invest Ophthalmol Vis Sci, 32(10):2821–2825 [*88*]
44. **Blake, R.**, *et al.* (**2006**) *Strength of early visual adaptation depends on visual awareness*. Proc Natl Acad Sci U S A, 103(12):4783–4788 [*118*]
45. **Blakemore, C.** & **Campbell, F.W.** (**1969**) *On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images*. J Physiol, 203(1):237–60 [*96, 102*]
46. **Blakemore, C.**, **Nachmias, J.**, & **Sutton, P.** (**1970**) *The perceived spatial frequency shift: evidence for frequency-selective neurones in the human brain*. J Physiol, 210(3):727–50 [*96*]
47. **Bonneh, Y.**, **Sagi, D.**, & **Karni, A.** (**2001**) *A transition between eye and object rivalry determined by stimulus coherence*. Vision Res, 41(8):981–9 [*67*]
48. **Born, R.T.** & **Bradley, D.C.** (**2005**) *Structure and function of visual area MT*. Annu Rev Neurosci, 28:157–189 [*3, 32*]
49. **Bosking, W.H.**, *et al.* (**1997**) *Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex*. J Neurosci, 17(6):2112–2127 [*40*]
50. **Braddick, O.** (**1974**) *A short-range process in apparent motion*. Vision Res, 14(7):519–527 [*5, 14*]
51. **Braddick, O.J.**, **Wishart, K.A.**, & **Curran, W.** (**2002**) *Directional performance in motion transparency*. Vision Res, 42(10):1237–1248 [*32*]
52. **Brascamp, J.W.**, *et al.* (**2005**) *Distributions of alternation rates in various forms of bistable perception*. J Vis, 5(4):287–298 [*115*]
53. **Brascamp, J.W.**, *et al.* (**2006**) *The time course of binocular rivalry reveals a fundamental role of noise*. J Vis, 6(11):1244–1256 [*83, 86, 88, 92*]
54. **Breese, B.** (**1899**) *On inhibition*. Psychol Monogr, 3:1–65 [*10, 115, 117*]
55. **Breitmeyer, B.**, **Rudd, M.**, & **Dunn, K.** (**1981**) *Metacontrast investigations of sustained-transient channel inhibitory interactions*. J Exp Psychol Hum Percept Perform, 7(4):770–779 [*82*]
56. **Breitmeyer, B.** & **Ogmen, H.**, *Visual Masking: Time slices through conscious and unconscious vision*, *Oxford Psychology Series*, vol. 41. 2nd edn. (New York: Oxford University Press, **2006**) [*9, 53, 54, 57, 82, 92, 93*]
57. **Burr, D.**, **Fiorentini, A.**, & **Morrone, C.** (**1998**) *Reaction time to motion onset of*

luminance and chromatic gratings is determined by perceived speed. Vision Res, 38(23):3681–3690 [*5, 14, 15, 27, 28, 29, 32, 33, 48, 49*]

58. **Carlson, T.A.** & **He, S.** (**2000**) *Visible binocular beats from invisible monocular stimuli during binocular rivalry.* Curr Biol, 10(17):1055–1058 [*61, 63*]
59. **Carlson, T.A.** & **He, S.** (**2004**) *Competing global representations fail to initiate binocular rivalry.* Neuron, 43(6):907–914 [*60, 118, 121*]
60. **Cavanagh, P.** & **Mather, G.** (**1989**) *Motion: the long and short of it.* Spat Vis, 4(2-3):103–129 [*5, 14*]
61. **Chawla, D.**, *et al.* (**1999**) *Speed-dependent responses in V5: A replication study.* Neuroimage, 9(5):508–515 [*28, 29*]
62. **Chklovskii, D.B.** & **Koulakov, A.A.** (**2004**) *Maps in the brain: what can we learn from them?* Annu Rev Neurosci, 27:369–392 [*1, 32*]
63. **Craik, K.** (**1940**) *Origin of visual after-images.* Nature, 145:512 [*118*]
64. **Crick, F.** & **Koch, C.** (**1992**) *The problem of consciousness.* Sci Am, 267(3):152–159 [*9*]
65. **Crowder, N.A.**, *et al.* (**2006**) *Relationship between contrast adaptation and orientation tuning in V1 and V2 of cat visual cortex.* J Neurophysiol, 95(1):271–283 [*96*]
66. **Curran, W.** & **Benton, C.P.** (**2003**) *Speed tuning of direction repulsion describes an inverted U-function.* Vision Res, 43(17):1847–1853 [*47*]
67. **Curran, W.** & **Benton, C.** (**2006**) *Test stimulus characteristics determine the perceived speed of the dynamic motion aftereffect.* Vision Res, 46(19):3284–3290 [*5*]
68. **Dakin, S.C.** & **Mareschal, I.** (**2000**) *The role of relative motion computation in 'direction repulsion'.* Vision Res, 40(7):833–841 [*32*]
69. **De Valois, R.L.** & **Cottaris, N.P.** (**1998**) *Inputs to directionally selective simple cells in macaque striate cortex.* Proc Natl Acad Sci U S A, 95(24):14488–14493 [*28*]
70. **DeAngelis, G.C.** & **Newsome, W.T.** (**1999**) *Organization of disparity-selective neurons in macaque area MT.* J Neurosci, 19(4):1398–1415 [*39*]
71. **Dember, W.N.** & **Purcell, D.G.** (**1967**) *Recovery of masked visual targets by inhibition of the masking stimulus.* Science, 157(794):1335–1336 [*82*]
72. **Diaz-Caneja, E.** (**1928**) *Sur l'alternance binoculaire.* Annales d'Oculistique, pp. 721–731 [*9, 115*]
73. **Dodwell, P.C.** & **Engel, G.R.** (**1963**) *A theory of binocular fusion.* Nature, 198:39–40 [*52, 62, 94*]
74. **Duhamel, J.**, **Colby, C.**, & **Goldberg, M.** (**1992**) *The updating of the representation of visual space in parietal cortex by intended eye movements.* Science, 255(5040):90–92 [*106*]
75. **Edwards, M.**, **Badcock, D.R.**, & **Smith, A.T.** (**1998**) *Independent speed-tuned global-motion systems.* Vision Res, 38(11):1573–1580 [*5, 8, 14, 28, 29, 31, 33, 34, 35, 36, 37, 38, 40, 45, 110, 111*]
76. **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 [*85, 89, 115*]
77. **Engel, E.** (**1956**) *The rôle content in binocular resolution.* Am J Psychol, 69(1):87–91 [*121*]
78. **Feldman, J.** & **Ballard, D.** (**1982**) *Connectionist models and their properties.* Cognitive Psychol, 6:205–254 [*1*]

79. **Felleman, D.** & **Van Essen, D.** (**1991**) *Distributed hierarchical processing in the primate cerebral cortex.* Cereb Cortex, 1(1):1–47 [*2*]
80. **ffytche, D.H.**, **Guy, C.N.**, & **Zeki, S.** (**1995**) *The parallel visual motion inputs into areas V1 and V5 of human cerebral cortex.* Brain, 118 (Pt 6):1375–1394 [*28*]
81. **Foster, K.H.**, *et al.* (**1985**) *Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey.* J Physiol, 365:331–363 [*111*]
82. **Fox, R.** & **Check, R.** (**1968**) *Detection of motion during binocular rivalry suppression.* J Exp Psychol, 78(3):388–395 [*8, 114, 118, 120*]
83. **Fox, R.** & **Check, R.** (**1972**) *Independence between binocular rivalry suppression duration and magnitude of suppression.* J Exp Psychol, 93(2):283–289 [*8, 114, 118, 120*]
84. **Fox, R.** & **Herrmann, J.** (**1967**) *Stochastic properties of binocular rivalry alternations.* Percept Psychophys, 2:432–436 [*83*]
85. **Fox, R.** & **Rasche, F.** (**1969**) *Binocular rivalry and reciprocal inhibition.* Percept Psychophys, 5:215–217 [*83*]
86. **Freeman, A.W.** (**2005**) *Multistage model for binocular rivalry.* J Neurophysiol, 94(6):4412–4420 [*61, 76, 78, 96, 115*]
87. **Fries, P.**, *et al.* (**1997**) *Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry.* Proc Natl Acad Sci U S A, 94(23):12699–12704 [*119, 120*]
88. **Fries, P.**, *et al.* (**2002**) *Oscillatory neuronal synchronization in primary visual cortex as a correlate of stimulus selection.* J Neurosci, 22(9):3739–3754 [*119, 120*]
89. **Gattass, R.**, *et al.* (**2005**) *Cortical visual areas in monkeys: location, topography, connections, columns, plasticity and cortical dynamics.* Philos Trans R Soc Lond B Biol Sci, 360(1456):709–731 [*105*]
90. **Gegenfurtner, K.** & **Hawken, M.** (**1995**) *Temporal and chromatic properties of motion mechanisms.* Vision Res, 35(11):1547–1563 [*5, 14, 15, 28, 29, 32, 33, 48, 49*]
91. **Gegenfurtner, K.** & **Hawken, M.** (**1996**) *Interaction of motion and color in the visual pathways.* Trends Neurosci, 19(9):394–401 [*5, 14, 15, 28, 29*]
92. **Geisler, W.S.**, *et al.* (**2001**) *Edge co-occurrence in natural images predicts contour grouping performance.* Vision Res, 41(6):711–724 [*62*]
93. **Georgeson, M.A.** & **Georgeson, J.M.** (**1987**) *Facilitation and masking of briefly presented gratings: time-course and contrast dependence.* Vision Res, 27(3):369–379 [*92*]
94. **Gheorghiu, E.** & **Erkelens, C.J.** (**2005**) *Differences in perceived depth for temporally correlated and uncorrelated dynamic random-dot stereograms.* Vision Res, 45(12):1603–1614 [*62*]
95. **Gilbert, C.D.** (**1977**) *Laminar differences in receptive field properties of cells in cat primary visual cortex.* J Physiol, 268(2):391–421 [*61*]
96. **Gilroy, L.A.** & **Blake, R.** (**2005**) *The interaction between binocular rivalry and negative afterimages.* Curr Biol, 15(19):1740–1744 [*118*]
97. **Gorea, A.**, **Papathomas, T.**, & **Kovacs, I.** (**1993**) *Two motion systems with common and separate pathways for color and luminance.* Proc Natl Acad Sci U S A, 90(23):11197–11201 [*5, 14, 28, 29*]

98. **Green, M.**, **Chilcoat, M.**, & **Stromeyer, C.F.** (**1983**) *Rapid motion aftereffect seen within uniform flickering test fields.* Nature, 304(5921):61–62 [*113*]
99. **Grill-Spector, K.** (**2003**) *The neural basis of object perception.* Curr Opin Neurobiol, 13(2):159–166 [*60, 61*]
100. **Van de Grind, W.**, **Grüsser, O.**, & **Lunkenheimer, H.**, *Temporal transfer properties of the afferent visual system. physiological, neurophysiological and theoretical investigations.* In: *Handbook of Sensory Physiology (Vol. VII/3A)* (R. Jung, ed.), pp. 431–574 (Berlin: Springer Verlag, **1973**) [*52, 61, 94*]
101. **van de Grind, W.**, *et al.* (**2001**) *Slow and fast visual motion channels have independent binocular-rivalry stages.* Proc R Soc Lond B Biol Sci, 268(1465):437–443 [*5, 14, 15, 19, 27, 28, 29*]
102. **Grinvald, A.**, *et al.* (**1986**) *Functional architecture of cortex revealed by optical imaging of intrinsic signals.* Nature, 324(6095):361–364 [*1, 5*]
103. **Grunewald, A.** & **Lankheet, M.J.** (**1996**) *Orthogonal motion after-effect illusion predicted by a model of cortical motion processing.* Nature, 384(6607):358–360 [*17, 19*]
104. **Gur, M.** & **Snodderly, D.M.** (**1997**) *A dissociation between brain activity and perception: chromatically opponent cortical neurons signal chromatic flicker that is not perceived.* Vision Res, 37(4):377–382 [*61*]
105. **Hammett, S.T.**, *et al.* (**2005**) *A ratio model of perceived speed in the human visual system.* Proc Biol Sci, 272(1579):2351–2356 [*28*]
106. **Harrad, R.A.** & **Hess, R.F.** (**1992**) *Binocular integration of contrast information in amblyopia.* Vision Res, 32(11):2135–2150 [*92*]
107. **Hawken, M.J.**, **Gegenfurtner, K.R.**, & **Tang, C.** (**1994**) *Contrast dependence of colour and luminance motion mechanisms in human vision.* Nature, 367(6460):268–270 [*5, 14, 15, 27, 28, 29, 33, 34, 49*]
108. **Hawken, M.J.**, **Shapley, R.M.**, & **Grosof, D.H.** (**1996**) *Temporal-frequency selectivity in monkey visual cortex.* Vis Neurosci, 13(3):477–492 [*72*]
109. **Haynes, J.D.**, **Deichmann, R.**, & **Rees, G.** (**2005**) *Eye-specific effects of binocular rivalry in the human lateral geniculate nucleus.* Nature, 438(7067):496–499 [*73, 77, 114*]
110. **He, S.**, *et al.*, *Parallel pathways and temporal dynamics in binocular rivalry.* In: *Binocular Rivalry* (D. Alais & R. Blake, eds.), pp. 81–100 (MIT Press, **2005**) [*60, 61*]
111. **Heeger, D.J.**, **Simoncelli, E.P.**, & **Movshon, J.A.** (**1996**) *Computational models of cortical visual processing.* Proc Natl Acad Sci U S A, 93(2):623–627 [*48*]
112. **Heeger, D.J.**, *et al.* (**1999**) *Motion opponency in visual cortex.* J Neurosci, 19(16):7162–7174 [*5, 16, 17, 48*]
113. **von Helmholtz, H.**, *Handbuch der Physiologischen Optik* (Hamburg: Voss, **1867**) [*8, 105*]
114. **Hering, K.**, *Outlines of a theory of the light sense* (Cambridge, Mass.: Harvard University Press, **1964**) [*8*]
115. **Hiris, E.** & **Blake, R.** (**1996**) *Direction repulsion in motion transparency.* Vis Neurosci, 13(1):187–197 [*32*]
116. **Holcombe, A.O.** & **Cavanagh, P.** (**2001**) *Early binding of feature pairs for visual perception.* Nat Neurosci, 4(2):127–128 [*67*]

117. **Hollins, M.** (**1980**) *The effect of contrast on the completeness of binocular rivalry suppression*. Percept Psychophys, 27(6):550–556 [*82, 92*]

118. **Howard, I.P.** (**1995**) *Depth from binocular rivalry without spatial disparity*. Perception, 24(1):67–74 [*121*]

119. **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 [*61*]

120. **Hubel, D.** & **Wiesel, T.** (**1963**) *Shape and arrangement of columns in cat's striate cortex*. J Physiol, 165:559–568 [*1, 5*]

121. **Hubel, D.** & **Wiesel, T.** (**1968**) *Receptive fields and functional architecture of monkey striate cortex*. J Physiol, 195(1):215–243 [*1, 5, 32*]

122. **Hupé, 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 [*89*]

123. **Irwin, D.E.** & **Brockmole, J.R.** (**2004**) *Suppressing where but not what: the effect of saccades on dorsal- and ventral-stream visual processing*. Psychol Sci, 15(7):467–473 [*62*]

124. **Jiang, Y.**, **Costello, P.**, & **He, S.** (**2007**) *Processing of invisible stimuli: advantage of upright faces and recognizable words in overcoming interocular suppression*. Psychol Sci, 18(4):349–355 [*122*]

125. **Jiang, Y.**, **Zhou, K.**, & **He, S.** (**2007**) *Human visual cortex responds to invisible chromatic flicker*. Nat Neurosci, 10(5):657–662 [*61, 77*]

126. **Julesz, B.**, *Foundations of cyclopean perception* (Chicago: University of Chicago Press, **1971**) [*96*]

127. **Kakizaki, S.** (**1950**) *Effects of preceding conditions upon binocular rivalry I*. Jpn J Psychol, 20:24–32 [*121*]

128. **Kanwisher, N.**, **McDermott, J.**, & **Chun, M.** (**1997**) *The fusiform face area: a module in human extrastriate cortex specialized for face perception*. J Neurosci, 17(11):4302–4311 [*1*]

129. **Kaplan, I.** & **Metlay, W.** (**1964**) *Light intensity and binocular rivalry*. J Exp Psychol, 67:22–26 [*9, 114*]

130. **Khuu, S.K.** & **Badcock, D.R.** (**2002**) *Global speed processing: evidence for local averaging within, but not across two speed ranges*. Vision Res, 42(28):3031–3042 [*5, 8, 14, 28, 29, 31, 33, 34, 35, 37, 38, 40, 45, 47, 110, 111*]

131. **Kim, C.Y.** & **Blake, R.** (**2005**) *Psychophysical magic: rendering the visible 'invisible'*. Trends Cogn Sci, 9(8):381–388 [*92*]

132. **Kim, J.** & **Wilson, H.R.** (**1993**) *Dependence of plaid motion coherence on component grating directions*. Vision Res, 33(17):2479–2489 [*27, 28*]

133. **Kim, J.** & **Wilson, H.R.** (**1996**) *Direction repulsion between components in motion transparency*. Vision Res, 36(8):1177–1187 [*32, 48*]

134. **Knapen, T.**, *et al.* (**2007**) *Stimulus flicker alters interocular grouping during binocular rivalry*. Vision Res, 47(1):1–7 [*115*]

135. **Koenderink, J.** & **van Doorn, A.** (**1979**) *Spatiotemporal contrast detection threshold surface is bimodal*. Opt Lett, 4:32–34 [*60*]

136. **Kolers, P.A.** & **Rosner, B.S.** (**1960**) *On visual masking (metacontrast): dichoptic observation*. Am J Psychol, 73:2–21 [*82*]

137. **Kourtzi, Z.** & **Huberle, E.** (**2005**) *Spatiotemporal characteristics of form analysis in the human visual cortex revealed by rapid event-related fMRI adaptation*. Neuroim-

age, 28(2):440–452 [*61, 77*]

138. **Kovács, I.**, *et al.* (**1996**) *When the brain changes its mind: interocular grouping during binocular rivalry.* Proc Natl Acad Sci U S A, 93(26):15508–15511 [*9, 10, 67, 114, 115, 121*]
139. **Krekelberg, B.** & **Albright, T.D.** (**2005**) *Motion mechanisms in macaque MT.* J Neurophysiol, 93(5):2908–2921 [*18*]
140. **Kulikowski, J.J.** & **Tolhurst, D.J.** (**1973**) *Psychophysical evidence for sustained and transient detectors in human vision.* J Physiol, 232(1):149–162 [*60*]
141. **Lack, L.C.**, *Selective Attention and the Control of Binocular Rivalry* (The Hague, The Netherlands: Mouton, **1978**) [*118, 122*]
142. **Laing, C.R.** & **Chow, C.C.** (**2002**) *A spiking neuron model for binocular rivalry.* J Comput Neurosci, 12(1):39–53 [*96*]
143. **Lee, S.H.** & **Blake, R.** (**1999**) *Rival ideas about binocular rivalry.* Vision Res, 39(8):1447–1454 [*66, 67, 68, 69, 78, 115, 117*]
144. **Lee, S.H.** & **Blake, R.** (**2004**) *A fresh look at interocular grouping during binocular rivalry.* Vision Res, 44(10):983–991 [*67, 115*]
145. **Legge, G.E.** (**1979**) *Spatial frequency masking in human vision: binocular interactions.* J Opt Soc Am, 69(6):838–847 [*82, 92*]
146. **Lehky, S.R.** (**1988**) *An astable multivibrator model of binocular rivalry.* Perception, 17(2):215–228 [*103*]
147. **Lehmkuhle, S.** & **Fox, R.** (**1980**) *Effect of depth separation on metacontrast masking.* J Exp Psychol Hum Percept Perform, 6(4):605–621 [*52, 62*]
148. **Lehmkuhle, S.W.** & **Fox, R.** (**1975**) *Effect of binocular rivalry suppression on the motion aftereffect.* Vision Res, 15(7):855–859 [*118*]
149. **Leopold, D.A.** & **Logothetis, N.K.** (**1996**) *Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry.* Nature, 379(6565):549–553 [*61, 67, 76, 78, 96*]
150. **Leopold, D.A.** & **Logothetis, N.K.** (**1999**) *Multistable phenomena: changing views in perception.* Trends Cogn Sci, 3(7):254–264 [*9, 61, 114*]
151. **Leopold, D.A.**, *et al.* (**2001**) *Prototype-referenced shape encoding revealed by high-level aftereffects.* Nat Neurosci, 4(1):89–94 [*96*]
152. **Leopold, D.**, *et al.*, *Binocular rivalry and the illusion of monocular vision.* In: *Binocular Rivalry* (D. Alais & R. Blake, eds.) (Cambridge, MA: MIT Press, **2005**) [*66, 115*]
153. **Levelt, W.J.** (**1965**) *On binocular rivalry.* Tech. rep., Institute for perception RVO-TNO [*8, 9, 66, 102, 114, 118*]
154. **Levelt, W.J.** (**1967**) *Note on the distribution of dominance times in binocular rivalry.* Br J Psychol, 58(1):143–145 [*83*]
155. **Levi, D.M.**, **Harwerth, R.S.**, & **Smith, E.L.** (**1979**) *Humans deprived of normal binocular vision have binocular interactions tuned to size and orientation.* Science, 206(4420):852–854 [*82, 92*]
156. **Liu, J.** & **Newsome, W.T.** (**2003**) *Functional organization of speed tuned neurons in visual area MT.* J Neurophysiol, 89(1):246–256 [*15, 16, 17, 29, 32, 39, 110*]
157. **Liu, L.**, **Tyler, C.W.**, & **Schor, C.M.** (**1992**) *Failure of rivalry at low contrast: evidence of a suprathreshold binocular summation process.* Vision Res, 32(8):1471–1479 [*91, 118*]

158. **Logothetis, N.K.** (**1998**) *Single units and conscious vision.* Philos Trans R Soc Lond B Biol Sci, 353(1377):1801–1818 [*8, 9, 66, 121*]

159. **Logothetis, N.K.**, **Leopold, D.A.**, & **Sheinberg, D.L.** (**1996**) *What is rivalling during binocular rivalry?* Nature, 380(6575):621–624 [*9, 10, 55, 61, 66, 67, 115, 117, 119*]

160. **Logothetis, N.K.** & **Schall, J.D.** (**1989**) *Neuronal correlates of subjective visual perception.* Science, 245(4919):761–763 [*78, 96*]

161. **Logothetis, N.K.**, *et al.* (**2001**) *Neurophysiological investigation of the basis of the fMRI signal.* Nature, 412(6843):150–157 [*76*]

162. **Logothetis, N.** (**2002**) *The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal.* Philos Trans R Soc Lond B Biol Sci, 357(1424):1003–1037 [*76*]

163. **Logothetis, N.** (**2003**) *The underpinnings of the bold functional magnetic resonance imaging signal.* J Neurosci, 23(10):3963–3971 [*76*]

164. **Lorenceau, J.** & **Alais, D.** (**2001**) *Form constraints in motion binding.* Nat Neurosci, 4(7):745–751 [*61*]

165. **Lu, Z.L.** & **Sperling, G.** (**2001**) *Three-systems theory of human visual motion perception: review and update.* J Opt Soc Am A Opt Image Sci Vis, 18(9):2331–2370 [*5, 15*]

166. **Ludwig, I.**, **Pieper, W.**, & **Lachnit, H.** (**2007**) *Temporal integration of monocular images separated in time: stereopsis, stereoacuity, and binocular luster.* Percept Psychophys, 69(1):92–102 [*121*]

167. **Lumer, E.D.**, **Friston, K.J.**, & **Rees, G.** (**1998**) *Neural correlates of perceptual rivalry in the human brain.* Science, 280(5371):1930–1934 [*9, 114*]

168. **Lumer, E.D.** & **Rees, G.** (**1999**) *Covariation of activity in visual and prefrontal cortex associated with subjective visual perception.* Proc Natl Acad Sci USA, 96:1669–1673 [*9, 114*]

169. **Macknik, S.L.** (**2006**) *Visual masking approaches to visual awareness.* Prog Brain Res, 155:177–215 [*82*]

170. **Macknik, S.L.** & **Martinez-Conde, S.** (**2004**) *Dichoptic visual masking reveals that early binocular neurons exhibit weak interocular suppression: implications for binocular vision and visual awareness.* J Cogn Neurosci, 16(6):1049–1059 [*93, 94*]

171. **Macknik, S.L.**, **Martinez-Conde, S.**, & **Haglund, M.M.** (**2000**) *The role of spatiotemporal edges in visibility and visual masking.* Proc Natl Acad Sci U S A, 97(13):7556–7560 [*82*]

172. **Macknik, S.** & **Livingstone, M.** (**1998**) *Neuronal correlates of visibility and invisibility in the primate visual system.* Nat Neurosci, 1(2):144–149 [*93*]

173. **Maier, A.**, **Logothetis, N.K.**, & **Leopold, D.A.** (**2005**) *Global competition dictates local suppression in pattern rivalry.* J Vis, 5(9):668–677 [*66, 67, 115*]

174. **Malach, R.**, *et al.* (**1997**) *Organization of intrinsic connections in owl monkey area MT.* Cereb Cortex, 7(4):386–393 [*39, 40*]

175. **Malonek, D.**, **Tootell, R.B.**, & **Grinvald, A.** (**1994**) *Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT.* Proc Biol Sci, 258(1352):109–119 [*32, 39, 110*]

176. **Marshak, W.** & **Sekuler, R.** (**1979**) *Mutual repulsion between moving visual targets.* Science, 205(4413):1399–1401 [*32*]

177. **Masland, R.** (**2001**) *The fundamental plan of the retina.* Nat Neurosci, 4(9):877–886 [*1*]
178. **Mather, G.** (**1980**) *The movement aftereffect and a distribution-shift model for coding the direction of visual movement.* Perception, 9(4):379–392 [*5, 6, 29, 32, 110*]
179. **Mather, G.** & **Harris, J.**, *Theoretical models of the motion aftereffect.* In: *The motion aftereffect* (G. Mather, F. Verstraten, & S. Anstis, eds.), pp. 157–185 (MIT Press, **1998**) [*29*]
180. **Mather, G.** & **Moulden, B.** (**1980**) *A simultaneous shift in apparent direction: further evidence for a "distribution-shift" model of direction coding.* Q J Exp Psychol, 32(2):325–333 [*5, 6*]
181. **Mather, G.**, **Verstraten, F.**, & **Anstis, S.** (Eds.), *The motion aftereffect* (MIT Press, **1998**) [*4, 96*]
182. **Maunsell, J.** & **Van Essen, D.** (**1983**) *Functional properties of neurons in middle temporal visual area of the macaque monkey. I. selectivity for stimulus direction, speed, and orientation.* J Neurophysiol, 49(5):1127–1147 [*15, 16, 17, 19, 29, 37, 39, 48, 110*]
183. **McKee, S.P.**, *et al.* (**1994**) *Stereo matching precedes dichoptic masking.* Vision Res, 34(8):1047–1060 [*82*]
184. **Meese, T.S.** & **Hess, R.F.** (**2004**) *Low spatial frequencies are suppressively masked across spatial scale, orientation, field position, and eye of origin.* J Vis, 4(10):843–859 [*82, 92, 93*]
185. **Melcher, D.** (**2005**) *Spatiotopic transfer of visual-form adaptation across saccadic eye movements.* Curr Biol, 15(19):1745–1748 [*97, 104*]
186. **Metha, A.B.** & **Mullen, K.T.** (**1997**) *Red-green and achromatic temporal filters: a ratio model predicts contrast-dependent speed perception.* J Opt Soc Am A Opt Image Sci Vis, 14(5):984–996 [*28, 33, 49*]
187. **Metha, A.** & **Mullen, K.** (**1998**) *Failure of direction discrimination at detection threshold for both fast and slow chromatic motion.* J Opt Soc Am A Opt Image Sci Vis, 15(12):2945–2950 [*28, 33, 49*]
188. **Michaels, C.** & **Turvey, M.** (**1979**) *Central sources of visual masking: indexing structures supporting seeing at a single, brief glance.* Psychol Res., 41(1):2–61 [*9, 82, 90*]
189. **Morand, S.**, *et al.* (**2000**) *Electrophysiological evidence for fast visual processing through the human koniocellular pathway when stimuli move.* Cereb Cortex, 10(8):817–825 [*28*]
190. **Moulden, B.** (**1980**) *After-effects and the integration of patterns of neural activity within a channel.* Philos Trans R Soc Lond B Biol Sci, 290(1038):39–55 [*113*]
191. **Moutoussis, K.**, *et al.* (**2005**) *A binocular rivalry study of motion perception in the human brain.* Vision Res, 45(17):2231–2243 [*61*]
192. **Movshon, J.A.** & **Newsome, W.T.** (**1996**) *Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys.* J Neurosci, 16(23):7733–7741 [*40*]
193. **Movshon, J.**, *et al.* (**1985**) *The analysis of moving visual patterns.* Pont Acad Scien Scripta Varia, 54:117–151 [*5, 7, 32, 47, 111, 113*]
194. **Müller, J.R.**, *et al.* (**1999**) *Rapid adaptation in visual cortex to the structure of images.* Science, 285(5432):1405–1408 [*61*]
195. **Nakamura, K.** & **Colby, C.L.** (**2002**) *Updating of the visual representation in mon-*

key striate and extrastriate cortex during saccades. Proc Natl Acad Sci U S A, 99(6):4026–4031 [*106*]

196. **Nawrot, M.** & **Blake, R.** (**1989**) *Neural integration of information specifying structure from stereopsis and motion.* Science, 244(4905):716–718 [*97*]
197. **Necker, L.** (**1832**) *Observations on some remarkable phenomena seen in Switzerland: And an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid.* London Edinburgh Philos Mag J, 3:329–337 [*115*]
198. **Nguyen, V.A.**, **Freeman, A.W.**, & **Alais, D.** (**2003**) *Increasing depth of binocular rivalry suppression along two visual pathways.* Vision Res, 43(19):2003–2008 [*8, 75, 78, 96, 106, 114, 115*]
199. **Nguyen, V.A.**, **Freeman, A.W.**, & **Wenderoth, P.** (**2001**) *The depth and selectivity of suppression in binocular rivalry.* Percept Psychophys, 63(2):348–360 [*114, 118, 120*]
200. **Nguyen-Tri, D.** & **Faubert, J.** (**2002**) *The perceived speed of drifting chromatic gratings is mechanism-dependent.* Vision Res, 42(17):2073–2079 [*28*]
201. **Nishida, S.**, **Ashida, H.**, & **Sato, T.** (**1994**) *Complete interocular transfer of motion aftereffect with flickering test.* Vision Res, 34(20):2707–2716 [*112*]
202. **Nishida, S.** & **Sato, T.** (**1995**) *Motion aftereffect with flickering test patterns reveals higher stages of motion processing.* Vision Res, 35(4):477–490 [*5, 14, 113*]
203. **Nishida, S.**, *et al.* (**2003**) *Gaze modulation of visual aftereffects.* Vision Res, 43(6):639–649 [*97*]
204. **Noest, A.J.**, *et al.* (**2007**) *Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model.* J Vis, 7(8):10–10 [*96, 103*]
205. **Ooi, T.L.** & **He, Z.J.** (**2003**) *A distributed intercortical processing of binocular rivalry: psychophysical evidence.* Perception, 32(2):155–166 [*9*]
206. **O'Shea, R.P.**, *Effects of orientation and spatial frequency on monocular and binocular rivalry.* In: *Proceedings of the Fourth International Conference on Neural Information Processing* (N. Kasabov, R. Kozma, K. Ko, R. O'Shea, G. Coghill, & T. Gedeon, eds.), pp. 67–70 (Singapore: Springer-Verlag, **1998**) [*82, 92*]
207. **O'Shea, R.P.** & **Blake, R.** (**1986**) *Dichoptic temporal frequency differences do not lead to binocular rivalry.* Percept Psychophys, 39(1):59–63 [*60, 82*]
208. **O'Shea, R.P.** & **Crassini, B.** (**1981**) *Interocular transfer of the motion after-effect is not reduced by binocular rivalry.* Vision Res, 21(6):801–804 [*118*]
209. **O'Shea, R.P.** & **Crassini, B.** (**1984**) *Binocular rivalry occurs without simultaneous presentation of rival stimuli.* Percept Psychophys, 36(3):266–276 [*9, 10, 51, 52, 57, 69, 72, 82, 86, 89, 94, 116, 118*]
210. **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 [*82, 92, 118*]
211. **Pantle, A.** (**1974**) *Motion aftereffect magnitude as a measure of the spatio-termporal response properties of direction-selective analyzers.* Vision Res, 14:1229–1236 [*5, 113*]
212. **Pantle, A.** & **Picciano, L.** (**1976**) *A multistable movement display: evidence for two separate motion systems in human vision.* Science, 193(4252):500–502 [*5, 14*]
213. **Pearson, J.** & **Clifford, C.W.** (**2005**) *When your brain decides what you see: grouping across monocular, binocular, and stimulus rivalry.* Psychol Sci, 16(7):516–519 [*66*]

214. **Perrone, J.A.** & **Thiele, A.** (**2001**) *Speed skills: measuring the visual speed analyzing properties of primate MT neurons.* Nat Neurosci, 4(5):526–532 [*29, 39, 113*]

215. **Perrone, J.A.** & **Thiele, A.** (**2002**) *A model of speed tuning in MT neurons.* Vision Res, 42(8):1035–1051 [*28, 111, 113*]

216. **Polonsky, A.**, *et al.* (**2000**) *Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry.* Nat Neurosci, 3(11):1153–1159 [*61, 73, 114*]

217. **Priebe, N.J.**, **Cassanello, C.R.**, & **Lisberger, S.G.** (**2003**) *The neural representation of speed in macaque area MT/V5.* J Neurosci, 23(13):5650–5661 [*39, 113*]

218. **Purves, D.**, *et al.*, *Neuroscience* (Sunderland, MA: Sinauer Associates, Inc. Publishers, **1997**) [*1*]

219. **Qian, N.**, **Andersen, R.A.**, & **Adelson, E.H.** (**1994**) *Transparent motion perception as detection of unbalanced motion signals. i. psychophysics.* J Neurosci, 14(12):7357–7366 [*5, 7*]

220. **Ramachandran, V.S.**, *et al.* (**1998**) *Object recognition can drive motion perception.* Nature, 395(6705):852–853 [*61*]

221. **Reichardt, W.**, *Autocorrelation, a principle for the evaluation of sensory information by the central nervous system.* In: *Sensory Communication* (W. Rosenblith, ed.), pp. 303–317 (Cambridge, MA: MIT Press, **1961**) [*3*]

222. **Reid, R.C.**, **Soodak, R.E.**, & **Shapley, R.M.** (**1991**) *Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex.* J Neurophysiol, 66(2):505–529 [*32*]

223. **Ricci, C.** & **Blundo, C.** (**1990**) *Perception of ambiguous figures after focal brain lesions.* Neuropsychologia, 28(11):1163–1173 [*9, 114*]

224. **Rodman, H.R.** & **Albright, T.D.** (**1989**) *Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT).* Exp Brain Res, 75(1):53–64 [*32, 47, 111*]

225. **Ruppertsberg, A.I.**, **Wuerger, S.M.**, & **Bertamini, M.** (**2003**) *The chromatic input to global motion perception.* Vis Neurosci, 20(4):421–428 [*28*]

226. **van Santen, J.P.** & **Sperling, G.** (**1985**) *Elaborated reichardt detectors.* J Opt Soc Am A, 2(2):300–321 [*3*]

227. **Saul, A.B.** & **Cynader, M.S.** (**1989**) *Adaptation in single units in visual cortex: the tuning of aftereffects in the temporal domain.* Vis Neurosci, 2(6):609–620 [*96*]

228. **Schiller, P.H.** (**1965**) *Monoptic and dichoptic visual masking by patterns and flashes.* J Exp Psychol, 69:193–199 [*82*]

229. **Schiller, P.H.** & **Smith, M.C.** (**1966**) *Detection in metacontrast.* J Exp Psychol, 71(1):32–39 [*82*]

230. **Schor, C.M.** (**1977**) *Visual stimuli for strabismic suppression.* Perception, 6(5):583–593 [*82, 92*]

231. **Schrater, P.** & **Simoncelli, E.** (**1998**) *Local velocity representation: evidence from motion adaptation.* Vision Res, 38(24):3899–3912 [*5, 48*]

232. **Sheinberg, D.L.** & **Logothetis, N.K.** (**1997**) *The role of temporal cortical areas in perceptual organization.* Proc Natl Acad Sci U S A, 94(7):3408–3413 [*78*]

233. **Shou, T.**, *et al.* (**1995**) *Direction biases of X and Y type retinal ganglion cells in the cat.* J Neurophysiol, 73(4):1414–1421 [*1, 3*]

234. **Simoncelli, E.P.** & **Heeger, D.J.** (**1998**) *A model of neuronal responses in visual area mt.* Vision Res, 38(5):743–761 [*32, 48*]

235. **Sincich, L.C.**, *et al.* (**2004**) *Bypassing V1: a direct geniculate input to area MT.* Nat Neurosci, 7(10):1123–1128 [*28*]

236. **van der Smagt, M.**, **Verstraten, F.**, & **van de Grind, W.** (**1999**) *A new transparent motion aftereffect.* Nat Neurosci, 2(7):595–596 [*5, 14, 15, 19, 20, 22, 23, 25, 27, 29, 32, 33, 34, 49, 110, 113*]

237. **Smith, A.T.**, **Snowden, R.J.**, & **Milne, A.B.** (**1994**) *Is global motion really based on spatial integration of local motion signals?* Vision Res, 34(18):2425–2430 [*40*]

238. **Smith, A.**, **Scott-Samuel, N.**, & **Singh, K.** (**2000**) *Global motion adaptation.* Vision Res, 40(9):1069–1075 [*8*]

239. **Smith, M.A.**, **Majaj, N.J.**, & **Movshon, J.A.** (**2005**) *Dynamics of motion signaling by neurons in macaque area MT.* Nat Neurosci, 8(2):220–228 [*32*]

240. **Snowden, R.J.**, *et al.* (**1991**) *The response of area MT and V1 neurons to transparent motion.* J Neurosci, 11(9):2768–2785 [*17*]

241. **Sobel, K.V.** & **Blake, R.** (**2002**) *How context influences predominance during binocular rivalry.* Perception, 31(7):813–824 [*122*]

242. **Solomon, S.**, *et al.* (**2004**) *Profound contrast adaptation early in the visual pathway.* Neuron, 42(1):155–62 [*1, 96*]

243. **Sterzer, P.** & **Kleinschmidt, A.** (**2007**) *A neural basis for inference in perceptual ambiguity.* Proc Natl Acad Sci U S A, 104(1):323–328 [*9, 114*]

244. **Stoner, G.R.** & **Albright, T.D.**, *Visual motion integration: a neurophysiological and psychophysical perspective.* In: *Visual detection of motion* (A. Smith & R. Snowden, eds.) (London: Academic, **1994**) [*32, 40, 47, 111*]

245. **Sutherland, N.** (**1961**) *Figural aftereffects and apparent size.* Q J Exp Psychol, 13:222–228 [*6*]

246. **Suzuki, S.** & **Grabowecky, M.** (**2002**) *Evidence for perceptual "trapping" and adaptation in multistable binocular rivalry.* Neuron, 36(1):143–157 [*77, 115, 122*]

247. **Suzuki, S.** & **Grabowecky, M.** (**2002**) *Overlapping features can be parsed on the basis of rapid temporal cues that produce stable emergent percepts.* Vision Res, 42(24):2669–2692 [*67*]

248. **Tadin, D.**, *et al.* (**2002**) *What constitutes an efficient reference frame for vision?* Nat Neurosci, 5(10):1010–1015 [*61*]

249. **Thompson, P.** (**1982**) *Perceived rate of movement depends on contrast.* Vision Res, 22(3):377–380 [*28*]

250. **Tolhurst, D.J.** & **Movshon, J.A.** (**1975**) *Spatial and temporal contrast sensitivity of striate cortical neurones.* Nature, 257(5528):674–675 [*111*]

251. **Tong, F.** (**2001**) *Competing theories of binocular rivalry: A possible resolution.* Brain Mind, 2:35–83 [*9, 66, 67, 72, 117, 120*]

252. **Tong, F.** & **Engel, S.A.** (**2001**) *Interocular rivalry revealed in the human cortical blind-spot representation.* Nature, 411(6834):195–199 [*61, 73, 77, 114*]

253. **Tong, F.**, *et al.* (**1998**) *Binocular rivalry and visual awareness in human extrastriate cortex.* Neuron, 21(4):753–759 [*106*]

254. **Tootell, R.B.**, *et al.* (**1997**) *Functional analysis of V3A and related areas in human visual cortex.* J Neurosci, 17(18):7060–7078 [*29*]

255. **Tsao, D.Y.**, *et al.* (**2003**) *Stereopsis activates V3A and caudal intraparietal areas in*

macaques and humans. Neuron, 39(3):555–568 [*29*]

256. **Tsao, D.Y.**, *et al.* (**2006**) *A cortical region consisting entirely of face-selective cells.* Science, 311(5761):670–674 [*1*]
257. **Tsuchiya, N.** & **Koch, C.** (**2005**) *Continuous flash suppression reduces negative afterimages.* Nat Neurosci, 8(8):1096–1101 [*93, 94, 118*]
258. **Tsuchiya, N.**, *et al.* (**2006**) *Depth of interocular suppression associated with continuous flash suppression, flash suppression, and binocular rivalry.* J Vis, 6(10):1068–1078 [*92, 93*]
259. **Turvey, M.T.** (**1973**) *On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli.* Psychol Rev, 80(1):1–52 [*82*]
260. **Vaina, L.M.** (**1998**) *Complex motion perception and its deficits.* Curr Opin Neurobiol, 8(4):494–502 [*32, 47*]
261. **Verstraten, F.**, **van der Smagt, M.**, & **van de Grind, W.** (**1998**) *Aftereffect of high-speed motion.* Perception, 27(9):1055–1066 [*5, 14, 19, 22, 25, 26, 27, 29, 49, 110*]
262. **Verstraten, F.**, *et al.* (**1996**) *Recovery from adaptation for dynamic and static motion aftereffects: evidence for two mechanisms.* Vision Res, 36(3):421–424 [*112*]
263. **Verstraten, F.**, *et al.* (**1999**) *Integration after adaptation to transparent motion: static and dynamic test patterns result in different aftereffect directions.* Vision Res, 39(4):803–810 [*5, 33, 49*]
264. **Wade, N.J.** (**1975**) *Monocular and binocular rivalry between contours.* Perception, 4(1):85–95 [*66, 67, 72, 115*]
265. **Wade, N.J.** & **de Weert, C.M.** (**1986**) *Aftereffects in binocular rivalry.* Perception, 15(4):419–434 [*97, 118*]
266. **Wade, N.** & **Wenderoth, P.** (**1978**) *The influence of colour and contour rivalry on the magnitude of the tilt after-effect.* Vision Res, 18(7):827–35 [*118*]
267. **Wainwright, M.** & **Cavanagh, P.** (**1997**) *Static and flicker motion aftereffects in a velocity space.* Invest. Ophthalmol. Vis. Sci., 38(S374) [*18, 22*]
268. **Wales, R.** & **Fox, R.** (**1970**) *Increment detection thresholds during binocular rivalry suppression.* Percept Psychophys, 8:90–94 [*73*]
269. **Walker, P.** (**1978**) *Binocular rivalry: Central or peripheral selective processes?* Psychol Bull, 85:376–389 [*8, 118*]
270. **Walker, P.** & **Powell, D.J.** (**1979**) *The sensitivity of binocular rivalry to changes in the nondominant stimulus.* Vision Res, 19(3):247–249 [*115*]
271. **Wallach, H.** & **O'Connell, D.** (**1953**) *The kinetic depth effect.* J Exp Psychol, 45:205–217 [*115*]
272. **von Wartburg, R.**, *et al.* (**2007**) *Size matters: saccades during scene perception.* Perception, 36(3):355–365 [*62*]
273. **Watson, A.B.** & **Pelli, D.G.** (**1983**) *Quest: a bayesian adaptive psychometric method.* Percept Psychophys, 33(2):113–120 [*36*]
274. **Webster, M.A.** & **MacLin, O.H.** (**1999**) *Figural aftereffects in the perception of faces.* Psychon Bull Rev, 6(4):647–653 [*96*]
275. **de Weert, C.M.**, **Snoeren, P.R.**, & **Koning, A.** (**2005**) *Interactions between binocular rivalry and gestalt formation.* Vision Res, 45(19):2571–2579 [*115, 122*]
276. **Welch, L.** (**1989**) *The perception of moving plaids reveals two motion-processing*

stages. Nature, 337(6209):734–736 [*32*]

277. **Werner, H.** (**1935**) *Studies on contour: I. Qualitative analyses.* Am J of Psychol, 47:40–64 [*82*]
278. **Wexler, M.** & **van Boxtel, J.J.** (**2005**) *Depth perception by the active observer*. Trends Cogn Sci, 9(9):431–438 [*105, 106*]
279. **Wheatstone, C.** (**1938**) *Contributions to the physiology of vision: I. on some remarkable, and hitherto unobserved, phenomena of binocular vision.* Philos Trans R Soc Lond, 128:371–394 [*117*]
280. **Whittle, P.** (**1965**) *Binocular rivalry and the contrast at contours*. Q J Exp Psychol, 17(3):217–226 [*9, 114*]
281. **Wiesenfelder, H.** & **Blake, R.** (**1990**) *The neural site of binocular rivalry relative to the analysis of motion in the human visual system*. J Neurosci, 10(12):3880–3888 [*102*]
282. **Wilke, M.**, **Logothetis, N.K.**, & **Leopold, D.A.** (**2006**) *Local field potential reflects perceptual suppression in monkey visual cortex*. Proc Natl Acad Sci U S A, 103(46):17507–17512 [*76*]
283. **Wilson, H.R.**, **Blake, R.**, & **Lee, S.H.** (**2001**) *Dynamics of travelling waves in visual perception*. Nature, 412(6850):907–910 [*118*]
284. **Wilson, H.** (**2003**) *Computational evidence for a rivalry hierarchy in vision*. Proc Natl Acad Sci U S A, 100(24):14499–14503 [*61, 67, 69, 76, 96, 103, 115*]
285. **Wilson, H.**, **Ferrera, V.**, & **Yo, C.** (**1992**) *A psychophysically motivated model for two-dimensional motion perception*. Vis Neurosci, 9(1):79–97 [*6*]
286. **Wilson, H.** & **Kim, J.** (**1994**) *A model for motion coherence and transparency*. Vis Neurosci, 11(6):1205–1220 [*6, 18, 110*]
287. **Wohlgemuth, A.** (**1911**) *On the after-effect of seen movement*. Brit J of Psychol (Supp.), 1:1–117 [*113*]
288. **Wolfe, J.M.** (**1983**) *Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli*. Perception, 12(4):447–456 [*51, 57, 69, 76, 82, 88, 89*]
289. **Wolfe, J.M.** (**1984**) *Reversing ocular dominance and suppression in a single flash*. Vision Res, 24(5):471–478 [*9*]
290. **Wolpert, D.** & **Flanagan, J.** (**2001**) *Motor prediction*. Curr Biol, 11(18):R729–732 [*105*]
291. **Wright, M.** & **Johnston, A.** (**1985**) *Invariant tuning of motion aftereffect*. Vision Res, 25:1947–1955 [*5, 113*]
292. **Wunderlich, K.**, **Schneider, K.A.**, & **Kastner, S.** (**2005**) *Neural correlates of binocular rivalry in the human lateral geniculate nucleus*. Nat Neurosci, 8(11):1595–1602 [*73, 77, 114*]
293. **Wurtz, R.H.** (**1969**) *Visual receptive fields of striate cortex neurons in awake monkeys*. J Neurophysiol, 32(5):727–742 [*32*]
294. **Yuille, A.L.** & **Grzywacz, N.M.** (**1988**) *A computational theory for the perception of coherent visual motion*. Nature, 333(6168):71–74 [*32*]
295. **Zihl, J.**, **von Cramon, D.**, & **Mai, N.** (**1983**) *Selective disturbance of movement vision after bilateral brain damage*. Brain, 106 (Pt 2):313–340 [*5, 14, 29*]

Nederlandse Samenvatting

Onze dagelijkse visuele perceptie wordt ondersteund door een ingewikkelde set van interacties tussen hersengebieden. Vaak hebben de individuele hersengebieden een specifieke taak. Tussen en ook binnen deze hersengebieden vindt veel communicatie plaats. Het totaal van verbanden tussen, en binnen, de verschillende hersengebieden noem ik de achitectuur van de hersenen. Hoewel er veel bekend is over de hersen architectuur uit histologisch onderzoek, is er over de functionele architectuur (over *hoe* de interacties werken) nog vrij veel onbekend. In het eerste deel van dit proefschrift heb ik gekeken naar de functionele architectuur die aan de basis ligt van onze perceptie van snelheden van objecten. In het tweede deel heb ik gekeken naar de menselijke waarneming in situaties waarin een groot conflict bestaat tussen de twee oogbeelden (een situatie waarbij "binoculaire rivaliteit" ontstaat).

Bewegingen bestaan uit twee componenten: een richting en een snelheid. Het is algemeen aanvaard dat de waarneming van bewegingsrichting wordt ondersteund door een enkel systeem in de hersenen (dat is, een collectie hecht samenwerkende hersengebieden). Dit systeem is gevoelig voor alle bewegingsrichtingen. Voor bewegingssnelheid is dat echter niet het geval. Voor snelheidswaarneming is vaak geopperd dat het ondersteund wordt door twee (of meer) vrij onafhankelijke systemen: een voor lage snelheden, en een voor hoge snelheden. In dit proefschrift heb ik voorgesteld dat de tot nu bekende data ook een eenvoudigere interpretatie toelaat, namelijk dat er een enkel systeem bestaat dat het hele zichtbare snelheidsbereik beslaat. Ik heb dit voorstel getoetst door middel van enkele experimenten en wiskundige modelstudies. De resultaten wijzen uit dat een enkel snelheidssysteem inderdaad de in dit proefschrift gepresenteerde data, en dat van eerder werk, kan verklaren.

Binoculaire rivaliteit ontstaat wanneer onze twee ogen conflicterende beelden ontvangen. In plaats van een middeling van de twee beelden, zien we een voortdurende afwisseling van de twee beelden. De twee beelden uit verschillende ogen vechten als het ware om gezien te worden, vandaar de naam binoculaire rivaliteit. Hoe ruimtelijke aspecten van de rivaliserende beelden het rivaliteitsproces beïnvloeden is vaak onderzocht. Echter, de temporele aspecten zijn nog maar nauwelijks onderzocht. In dit proefschrift heb ik een eerste stap gezet in het systematisch onderzoeken van de invloed van die temporele aspecten. Opmerkelijk genoeg blijkt binocular rivaliteit vrij ongevoelig te zijn voor het tijdelijk weghalen van beide beelden; de twee beelden rivaliseren zelfs wanneer ze nooit beide tegelijk aanwezig zijn. Wel moeten ze herhaaldelijk aangeboden worden binnen 350 milliseconde, anders stopt de rivaliteit. Deze temporele grens aan binoculaire rivaliteit is afhankelijk van de aanwezigheid van conflicterende patrooninformatie, maar is onafhankelijk van de aanwezigheid van conflicterende bewegingsinformatie. Dezelfde temporele grens wordt ook gevonden wanneer de rivaliserende patronen in een en het zelfde oog worden geplaatst; er is dus geen conflict tussen de twee ogen nodig voor binoculaire rivaliteit. Deze reeks aan experimenten dragen bij aan de discussie over waar en hoe in de hersenen binoculaire conflicten worden opgelost. Er is een theorie die zegt dat binoculaire rivaliteit voornamelijk een competitie is tussen de twee ogen, terwijl een andere theorie zegt dat binoculaire rivaliteit een competitie is tussen de twee beelden, onafhankelijk van de ogen waarin ze gepresenteerd worden. De data in dit proefschrift zijn voornamelijk ondersteuning voor de tweede theorie.

Tezamen werpen deze resultaten een nieuwe blik op de functionele architectuur van het visuele systeem. Ten eerste heeft dit proefschrift laten zien dat voor onze snelheidswaarneming een enkel systeem afdoende is. Deze conclusie gaat in tegen theorieën die stellen dat er voor een goed functionerende architectuur twee systemen nodig zijn. Ten tweede is aangetoond dat binoculaire rivaliteit niet afhankelijk is van competitie tussen de ogen, maar in plaats daarvan afhankelijk is van een langzame competitie tussen patronen op een niveau waar oog-afhankelijke informatie grotendeels verloren is.

Ringkasan

Persepsi dari penglihatan kita sehari-hari dihasilkan oleh interaksi yang cukup rumit antar beberapa bagian otak. Setiap bagian otak ini seringkali mempunyai tugas yang khusus. Komunikasi yang aktif terjadi antar dan di dalam bagian-bagian otak tersebut. Kumpulan dari bagian-bagian otak ini dan interaksi yang terjadi didalamnya, saya namakan tatanan arsitektur otak. Walaupun telah banyak yang diketahui tentang tatanan arsitektur otak dari ilmu jaringan tubuh, bagaimana tatanan arsitektur otak itu sendiri berfungsi (dengan kata lain bagaimana interaksi di dalam otak bekerja) masih belum diketahui. Pada bagian pertama tesis ini, saya paparkan penelitian saya tentang bagaimana tatanan arsitektur otak kita berfungsi dalam membentuk persepsi penglihatan kita terhadap kecepatan suatu objek yang bergerak. Pada bagian kedua, saya meneliti bagaimana persepsi kita terbentuk dalam situasi dimana kedua mata kita masing-masing melihat pandangan yang berlainan satu sama lainnya (yang disebut juga rivalitas antar dua mata).

Gerakan dibentuk oleh dua komponen: arah dan kecepatan. Secara umum diketahui bahwa persepsi kita terhadap arah dibentuk oleh sistim tunggal di dalam otak yang sebenarnya dapat mengenali keseluruhan arah pergerakan. Akan tetapi hal ini tidak terbukti untuk persepsi terhadap kecepatan. Persepsi terhadap kecepatan seringkali dikatakan dibentuk oleh dua sistim atau lebih: satu untuk kecepatan rendah/lambat dan yang lain untuk kecepatan tinggi. Dalam tesis ini, saya mengajukan hipotesis dimana data yang dipaparkan oleh penelitian-penelitian terdahulu, dapat juga dijelaskan secara lebih sederhana, yaitu dengan adanya sistim tunggal yang dapat mencakup seluruh kecepatan yang berbeda. Hipotesis ini saya uji lebih jauh dengan mengadakan beberapa percobaan dan menggunakan bebagai model matematik. Hasil dari uji-coba dan model matematik tersebut ternyata berhasil memperlihatkan bahwa sebuah sistim tunggal dapat menjelaskan data yang dihasilkan dari uji-coba dalam tesis ini, dan juga data yang telah dipaparkan oleh penelitian-penelitian sebelumnya.

Ketika rivalitas antar dua mata yang menjadi fokus pada bagian kedua tesis ini terjadi, kita tidak melihat penglihatan rata-rata dari kedua pandangan yang berbeda tersebut, melainkan pergantian penglihatan antara satu pandangan dengan pandangan lainnya. Hal ini sepertinya terjadi karena adanya persaingan antara kedua pandangan tersebut untuk dilihat, oleh karena itu dina-

makan rivalitas antar dua mata. Telah banyak diketahui bahwa bentuk dari suatu stimulus/pandangan dapat mempengaruhi proses rivalitas antar dua mata tersebut. Tetapi, pengaruh waktu timbulnya stimulus-stimulus tersebut terhadap proses rivalitas antar dua mata, belum pernah diteliti. Oleh karena itu, saya mengambil langkah awal untuk meneliti secara sistematis pengaruh dari konfigurasi waktu. Hasil dari penelitian ini ternyata memperlihatkan bahwa proses rivalitas antar dua mata tidak bergantung pada konfigurasi waktu; dua pandangan/stimulus tetap bersaing walaupun mereka tidak terlihat dalam waktu yang bersamaan. Akan tetapi, kedua stimulus/pandangan tersebut harus tetap diperlihatkan secara berulang-ulang dengan waktu tunda diantaranya tidak lebih kecil dari 350 milidetik, untuk tetap menghasilkan rivalitas antar dua mata. Limitasi dari konfigurasi waktu ini bergantung pada adanya perbedaan bentuk kedua stimulus, tetapi tidak bergantung pada perbedaan gerakan dalam stimulus-stimulus itu sendiri. Lebih jauh lagi, ketika kedua stimulus/pandangan tesebut dipresentasikan hanya pada satu mata, limitasi konfigurasi waktu tersebut tetap sama. Oleh karena itu presentasi stimulus yang berbeda pada dua mata tidak diperlukan untuk mengundang terbentuknya rivalitas antar dua mata. Beberapa penelitian yang dipaparkan dalam tesis ini telah menambah wawasan dalam diskusi tentang bagaimana dan dimana terbentuk rivalitas antar dua mata pada otak. Salah satu teori tentang terbentuknya rivalitas antar dua mata menerangkan bahwa kompetisi terjadi antar dua mata, sedangkan teori lainnya menyarankan adanya kompetisi antar dua pandangan/stimulus yang tidak bergantung pada mata mana yang melihatnya. Data yang saya paparkan dalam tesis ini ternyata lebih mengacu pada teori kedua.

Secara keseluruhan, penelitian dalam tesis ini telah membuka wawasan baru tentang bagaimana tatanan arsitektur visual sistem kita berfungsi. Pada bagian pertama diperlihatkan bahwa keberadaan sistim yang tunggal dalam otak adalah cukup untuk melandasi persepsi penglihatan kita terhadap kecepatan suatu gerakan. Kesimpulan ini menentang teori umum yang menyatakan bahwa arsitektur otak harus memiliki dua sistim untuk dapat berfungsi secara optimal. Pada bagian kedua diperlihatkan bahwa rivalitas antar dua mata tidak bergantung pada kompetisi antar dua mata, melainkan bergantung pada perbedaan bentuk antar kedua stimulus/pandangan yang diproses di bagian arsitektur otak, yang mana tidak lagi mengetahui dari mata mana informasi tersebut berasal.

Dankwoord/Acknowledgments

Well, there we are. The last few pages of this thesis (although they may be the first you look at...). It is about time to thank all the people that contributed to this work in so many ways.

Let me start with my supervisor/advisor. Dear Casper, thank you for all your good advise, and for having given me the freedom to conduct research on whatever topic I thought was interesting at the time. Thank you for your minute readings of my manuscripts, my sloppy writing has improved significantly. I also very much appreciate that you always supported me to go to international conferences, where I indeed have learned a lot.

I also like to thank Raymond, my co-promotor. Although I started out on a topic not really in your field of study, I soon drifted towards binocular rivalry research. I would also like to thank you for helping me improve my writing skills and for the many interesting discussions on binocular rivalry.

I should also thank Mark Wexler here, who has gone through great pains teaching me the scientific essentials. Thank you, Mark, I've learned a lot from you. And my thanks also go to Jacques Droulez, Sasha Panfilov, and Paulien Hogeweg, as well. You had taken the roughest edges off before my PhD endeavor started.

I'd also like to thank my other colleagues for the really pleasant atmosphere: Myrthe, Dagmar, Gijs, Loes, Rob, Allard, Stefan, Wouter, Alex, Bernhard, Robbert, Maarten, Bram, Bala, Michelle, Mirella, Astrid, Jan and Ans, Pieter and Hans, and also Chris Paffen for the fun time in Australia. I'd like to thank Tomas Knapen and Jan Brascamp for your many interesting, and inspiring discussions, and Tomas as well for the technical support he was always ready to give. Of course, my thanks extend to the neurobiology group, for their interest and critical views on my work during the NEST meetings.

David Alais, thank you for your hospitality and providing a great work environment in Sydney, it has been a productive and fun time.

My deep gratitude goes to my dear international family in Utrecht and beyond: Sana, Elsa, Vicky, Lorenzo, Jerome, Fernando, April, Meindert, Luca, Timo, Gianluca, Elena, Eyk, Leo, Maarten, Davide, Hugo, Hugo, Roka, Růžena, Dodi, Jeroen Aalderink, Roel Kalkman, and Camille Morvan. You provided the necessary distraction from work, but also the energy to focus again.

Jan and Henny, Theo and Wil, thank you for helping out so many times during the last four years. April and I could always count on your support.

Meer dan zomaar een dank je wel gaat naar mijn ouders, Henk en Det, en de rest van de familie. Jullie stonden altijd klaar om te helpen, ook al was ik zelf vaak afwezig, vooral tegen het einde van mijn proefschrift. Dank jullie wel. Terima kasih banyak sekali untuk keluarga saya di Indonesia. Walaupun jauh di mata, tetapi dekat di hati.

April, we are together very much since the start of my PhD. You have always been there to support me, and to criticize me when it was necessary. You served as a subject and as a critical reviewer. Indeed, you have made important contributions to so many parts of this thesis. I thank you for your endurance and your love.

And special thanks to everyone I forgot.

Publications

Journal Articles

- van Boxtel JJ, Wexler M, & Droulez J. *Perception of plane orientation from self-generated and passively observed optic flow.* J Vis. 2003, 3(5):318-32.
- Wexler M, & van Boxtel JJ. *Depth perception by the active observer.* Trends Cogn. Sci. 2005, 9(9):431-8.
- van Boxtel, JJ, van Ee R, & Erkelens, CJ. *A single system explains human speed perception.* J. Cogn. Neurosc. 2006, 18(11):1808-1819
- van Boxtel JJ, & Erkelens, CJ. *A single motion system suffices for global-motion perception.* Vision Res. 2006, 46(28):4634-45
- van Boxtel JJ, van Ee, R., & Erkelens, CJ. *Dichoptic masking and binocular rivalry share common perceptual dynamics* J Vis. 2007, 7(14):1-11
- Knapen, T.H.J., Kanai, R. , Brascamp, J.W., van Boxtel J.J.A., & van Ee, R. *Distance in feature space determines exclusivity in visual rivalry* Vision Res. 2007, 47(26):3269-3275
- van Boxtel, J.J.A., Alais, D., Erkelens, C. J., & van Ee, R. *The role of temporally course form processing in binocular rivalry* (PLoS ONE, in press)
- van Boxtel, J.J.A., Alais, D., & van Ee, R. *Retinotopic and non-retinotopic stimulus encoding in binocular rivalry and the involvement of feedback* (J Vis., accepted pending revisions)
- van Boxtel, J.J.A., Knapen, T., C.J. Erkelens, & van Ee, R. *Visual rivalry: a balance between binocular ambiguity resolution and monocular interactions* (submitted)
- van Ee, R., van Boxtel, J.J.A., Parker, A., & Alais, D. *Attentional control of ambiguous visual perception is enhanced by congruent auditory signals* (submitted)

Conference Abstracts

- J.J.A. van Boxtel & A.V. Panfilov (2001). *Spiral Breakup in an Array of Coupled Cells. The Role of Gap Junctional Conductance.* XXXIVth International Congress of Physiological Sciences, Christchurch, New Zealand
- J.J.A. van Boxtel & M. Wexler (2003) *The consequences of head motor signals for the perception of 3-D surfaces.* European Conference on Visual Perception (ECVP 2003) Perception 32: 62-62 Suppl.
- J.J.A. van Boxtel & C.J. Erkelens (2004) *A single population of velocity detectors can account for the adaptation differences to fast and slow motion.* (ECVP 2004) Perception 33: 158-158 Suppl.
- J.J.A.van Boxtel & C.J.Erkelens (2005) *Global-motion perception is governed by a single motion system.* (ECVP 2005) Perception 34: 228-228 Suppl. S 2005
- J.J.A. van Boxtel, Kamphuisen, A. P., van Ee, R., & Erkelens, C. J. (2006). *The occurrence of binocular rivalry and dichoptic masking depends on temporal aspects of stimulation.* [Abstract VSS]. Journal of Vision, 6(6), 850a.
- J.J.A. van Boxtel, Knapen, T., van Ee, R., & Erkelens, C. J. (2006). *Exploring the parameter space of stimulus rivalry.* Perception 35: 32-32 Suppl.
- J.J.A. van Boxtel, Knapen, T., van Ee, R., & Erkelens, C. J. (2007). *Identical rivalry dynamics for monocular, stimulus and binocular rivalry.* [Abstract VSS].
- R. van Ee, Parker, A.L., van Boxtel, J.J.A., & Alais, D.M. (2007). *Voluntary control over visual rivalry can be enhanced by sound that is consistent with scene interpretation.* ECVP 2007
- D. Alais, Parker, A.L., van Boxtel, J.J.A., Paffen, C. & van Ee, R. (2007). *Attending to auditory signals slows binocular rivalry.* ECVP 2007
- J.J.A. van Boxtel, van Ee, R. & Erkelens, C.J.. (2007). *Binocular rivalry for temporally modulated stimuli depends on a form-based parsing mechanism.* ECVP 2007
- R van Ee, Parker, A., van Boxtel, J.J.A. & Alais, D. *Voluntary control over visual rivalry is enhanced by sound that is consistent with scene interpretation.* [Neuroscience 2007]
- T. Knapen, van Boxtel J., & van Ee, R. *Neural circuitry of segregation and binding in perceptual rivalry.* [Neuroscience 2007]
- J.J.A. van Boxtel, van Ee, R. & Erkelens, C.J. *Binocular rivalry depends on temporally coarse form processing* [Neuroscience 2007]

Curriculum Vitae

The author of this thesis, Jeroen van Boxtel—born in Gouda, the Netherlands—studied Biology at the Utrecht University in the Netherlands. He specialized in Neuroethiology and Theoretical/Computational Biology. He worked on the projects "Structure-from-motion perception in the active observer", together with Mark Wexler, and Jacques Droulez at the Collège de France, Paris, and "Spiral breakup in a computational model of the heart" with Sasha Panfilov in Utrecht. He obtained his propedeuse (cum laude) in 1998, and his master's degree (cum laude) in 2002. The following year, 2003, he obtained a master's degree in Cognitive Sciences at the Université Pierre et Marie Curie, Paris VI, in Paris, France.

He subsequently moved back to Utrecht to start his PhD-studies at the Helmholtz Institute, where he followed the graduate school, and was a member of the PhD-council. He worked under supervision of Prof. Casper Erkelens, and in collaboration with Dr. Raymond van Ee and Dr. David Alais. He worked on the topics of motion processing and binocular conflict resolution. The results of this last endeavor are presented in the thesis that lies in front of you.