DOTTORATO DI RICERCA
TOSCANO IN NEUROSCIENZE

CICLO XXVIII

COORDINATORE Prof. Renato Corradetti

SENSORY MECHANISMS FOR THE PROCESSING OF SPATIAL, TEMPORAL AND NUMERICAL INFORMATION

Settore Scientifico Disciplinare M-PSI/02

Dottorando	**Tutore**
Dott. Fornaciai Michele	Prof. Burr David Charles

Coordinatore
Prof. Corradetti Renato

Anni 2012/2015

“The game of science is, in principle, without end.
He who decides one day that scientific statements
do not call for any further test, and that they
can be regarded as finally verified,
retires from the game.”
- Karl Popper

Acknowledgements

First of all, I would like to thank my advisor prof. David Burr, and prof. Maria Concetta Morrone, from whom I have learnt the most important lessons about science. You have been amazing mentors for me. I would like to thank also Roberto Arrighi, Guido Marco Cicchini and Paola Binda, who guided me step by step and patiently explained me everything whenever I was in doubt. Your contribution and support has been crucial. A special thanks also to all the good labmates of the Pisa Vision Lab: Giovanni Anobile, Marco Turi, Irene Togoli, Claudia Lunghi, Koulla Mikellidou, Luca Lo Verde, Alessandro Benedetto, Elisa Castaldi, Jan Kurzawsky, Akshatha. I have enjoyed so much the company of all of you during this journey.

I would like to thank also Massimiliano Di Luca, who provided me the opportunity to join his team, and for his patience, knowledge and invaluable support. A sincere thanks also to all the people of the Time Lab: Darren Rhodes, Ninja Horr, Susan Li, Hua-Chun Sun (Drew), Elia Gatti, Dario Pittera, for putting up with me during my visit at University of Birmingham. I enjoyed the time spent in Birmingham, and I have learnt so much from all of you, so thank you.

Last, but not least, a special thanks to my family and all my friends. You supported me so much during these years. Thank you!

Table of contents

Abstract

Despite decades of research concerning the sensory mechanisms for the processing of spatial, temporal and numerical information, several points still remain subject of debate. In this work, we will report a series of studies aimed at providing new evidence regarding the sensory mechanisms specific for the processing of space, time and number, and also to investigate the possibility that a common magnitude system might play a role in their processing. In the first part of the work, we examine the disruptive perceptual effects during eye movements ("saccades"), affecting the representation of space. Such distortion of space, thought to be related to the ocular-motor parameters and linked to visual stability processes, is not usually observed under normal viewing conditions – which give rise to the possibility that it could be due at least in part to the saccade automaticity and stereotyping reached in the typical experimental paradigms. Our results, however, showed that the pattern of saccadic effects was only marginally affected by practice over the course of the experiment and that performance of experts remained similar when tested in a condition leading to less stereotypical saccades. These results indicate that perisaccadic compression is a robust behavior, insensitive to the specific paradigm and to the level of practice with the saccade task.

In the second part, we will report two studies concerning the perception of time. In the first study we investigated the effect of motion adaptation on apparent time (i.e. the observation that adapting to fast motion causes a reduction in perceived duration of the subsequent stimuli), which has been previously tested using only simple translational motion. Our results showed that the adaptation-induced compression of time is specific for translational motion, while adaptation to complex motion, either circular or radial, did not affect perceived duration of subsequently viewed stimuli. These results show that such effect occurs only for uni-directional translational motion, ruling out the possibility that the neural mechanisms of the adaptation occur at early levels of visual processing. In the second study, we investigated the predictions of a recent model concerning time perception (the State-Dependent Network model), to test whether it could be extended to different sensory modalities.

Our results showed that, while some of the constrains might be variable according to the specific sensory modality tested, the general predictions of the model hold under different circumstances.

In the third part, we will present a study concerning the perception of numerosity, and the idea of number as a primary perceptual feature. Recent works showed that like other perceptual attributes, numerosity is susceptible to adaptation, but this idea has been challenged claiming that adaptation may operate via related mechanisms, such as texture-density. To disentangle this issue we measured the effect of adaptation on clouds of connected-dots (creating a robust underestimation of numerosity) and unconnected dots. We showed that adaptation to the same number of dots as the test causes robust adaptation of the connected, but not of the unconnected dot-pattern, suggesting that adaptation occurs at neural levels encoding perceived numerosity, rather than at lower levels responding to the number of elements in the scene.

Finally, in the fourth part, we investigated the possibility of a generalized magnitude system. To find further evidence for such system, we tested the effect of motion adaptation on perceived numerosity, as it as been previously tested on perceived time. Our results showed a partially similar pattern of results, suggesting a common general system, but showing also that mechanisms for time and numerosity could partially different. In the second study, we investigated the interplay between space and time – namely, the possibility to exploit spatial information to improve temporal judgments. Our results showed that such interplay is actually possible: providing additional information about where an event is bound to happen, improve the temporal resolution in judging when the event will happen – which supports the idea that such magnitudes might be encoded within a common metric. We also tested the possibility of exploiting timing information (audio-visual asynchrony) to perform spatial judgments concerning size and distance in depth, but without finding any influence of such cues on visual size and distance judgments. Although we could not completely rule out this possibility, our results suggest that the influence of timing cues would not be as strong as other visual cues, and so it might be limited to a small range of circumstances.

We conclude that: (1) saccadic distortions and are not a by-product of specific methodologies, but are strictly related to saccade ocular-motor param-

eters; (2) time perception is supported by a distributed mechanism, deeply rooted into the sensory streams; (3) numerosity is a primary perceptual attribute, and numerosity adaptation acts via number-specific mechanisms; (4) space, time and number might be processed by a generalized magnitude system with a common metric, but their processing could exploit partially different mechanisms.

Chapter 1

General introduction

This work concerns the study of how we perceive space, time and numbers, how these features interact, and what mechanisms are used to process them. But our first question is: why space, time and number? The answer is very simple: because these features represent some of the very basic descriptors of the visual world, and by studying them, we study how the brain processes the features that will form the building blocks of our conscious experience. Since we base most of our behavior on sensory information, and that most of our experience is shaped and defined by our sensory systems, studying how sensation and perception work represents a step into the highly complex and articulated functioning of our brain and our conscious experience.

While it is now well accepted that our conscious lives and our experiences are made possible and determined by the processes of our brain, it is surprising that despite its central role some of the brain functions and processing mechanisms are still poorly understood. Besides the cognitive abilities that form our higher-level capacities, like intelligence, reasoning, memory, and even art, the more basic, relatively low-level functions represent probably the ground on which such unique human abilities are built. Knowing that, it appears clearly what is the importance of studying perception: to understand the basis of our rich conscious experience and of our incredible cognitive capacities.

Throughout this thesis we will present a series of investigations into some functional mechanisms subserving the processing of spatial, temporal and numerical information. Particularly, our aims in approaching the study of sensory processing is to search for evidence suggesting how such mechanism works.

However, achieving enough evidence to completely understand even only a limited portion of the mechanisms dedicated to different characteristics of the sensory world is a far utopist purpose. What we did – and how science usually works – is to search for some of the missing pieces in a bigger puzzle. We did not search for ultimate answers, but for the details that might eventually contribute to find the answers.

In this introduction we will provide some general information about the rationales and paradigms used to study the perception of space, time and number, while we will describe with more details the literature concerning these different features in the next chapters.

1.1 Some words about methodology

How can we study perception? One of the more successful strategies used in psychophysics, is to somehow deceive the sensory systems, and check how perceptual responses, and perception itself, change. To do so, we usually exploit very brief stimuli, flashed for just fractions of a second, or we try to change the representational properties of the system with procedure like adaptation. In the next sections we will briefly describe some of the procedures exploited to study the perception of space, time and number, while a more detailed and specific discussion about previous experimental results will be provided in the introductions to the next chapters.

1.1.1 Eye movements and the representation of space

The field of space perception is a very broad field, encompassing the study of several features with several methodologies. For the aims of these works – as also better explained in the specific chapter (Chapter 2) – we will focus on the study of how the visual system achieves a stable representation of the world, in the face of very rapid and continuous movements of the eyes. To study this ability, most of the studies have concentrated not on situations where the representation of the external world is reliably stable, but on cases in which it is not, trying to find combinations of stimuli leading to distortions and visual "illusions". Indeed, the rationale here – but it is a common feature of psychophysical studies in general – is to study not when the system works, but the circumstances that make the system “fail”. Analyzing the errors of the sensory system is probably one of the most common strategies to find out how the functional mechanisms work, and possibly what are the neural substrates of a given process. While in our everyday lives we all know that eye movements apparently cause no disruption of the visual experience, the crucial step concerned the discovery that very brief stimuli, presented just before or during an eye movement, are conversely subject to very large distortions, especially concerning their perceived position. Indeed, across a large number of locations and timings, brief stimuli are consistently localized in a wrong position, usually onto the target of the eye movement (e.g. Ross et al., 1997; Ross et al., 2001; see Chapter 2 for more details about this effects). What do such distortions tell us about the underlying neural system? Many researchers think that the pattern of distortions affecting stimuli presented during eye movements should give us information about how the visual system might construct a stable representation of the world (see Burr & Morrone, 2012 for a review). Indeed, such compressed spatial metrics resemble the dramatic alterations of neuronal receptive fields during observed in the period immediately before the saccadic onset. Moreover, the interactions observed between different stimuli presented during saccades seem to reflect integration processes performed by the visual system in the attempt to maintain vision stable. So, given this pattern of results, studying what happens during saccades and precisely mapping distortions and interactions between different stimuli appears as a crucial task for the understanding of the underlying neural processes. In Chapter 2 we will provide more details about the perceptual effects of saccades, and we will present an experimental investigation concerning the perisaccadic mislocalization effects.

1.1.2 Time and number and visual adaptation

As in the case of visual stability and the representation of space, also in the field of time perception many studies focused on perceptual distortions, trying to discover how the system works analyzing its “errors”. One of the most powerful techniques used to investigate perceptual mechanisms – and that recently has proven to be useful also for time and number perception – is the "adaptation" technique. Such technique requires simply to look at a sustained stimulus for a relatively long period of time, and then to perform a perceptual judgment (such, for instance a discrimination judgment – i.e. judging a given feature of the "adapted" stimulus compared to another stimulus not affected by adaptation). During this period, often referred as the adaptation phase, the machinery of our perceptual system somehow recalibrates according to the statistics and features of the adaptor stimulus, which causes a change in the responses to a subsequent stimulus. Indeed, what is usually observed with such technique is a distorted perception of the stimuli presented after the adaptation phase. Many theoretical frameworks tried to address what happens at the neural level during adaptation. Generally, adaptation is thought to be a form of recalibration adjusting neural responses – and consequently altering the representation of the external stimuli – according to the recent history of stimulation, as a sort of plasticity governed by a sophisticated set of mechanisms (Solomon & Kohn, 2014).

Again, as in the eye movement example, the pattern of distortions caused by adaptation is thought to be a signature of the underlying mechanism, and it is generally thought that being affected by these processes is an index of the existence of a specific mechanism for the processing of the sensory feature investigated. In other words, if a specific mechanism for the processing of a given feature exists, so it should be adaptable, and this is a distinctive characteristic of the primary perceptual attributes. Indeed, the fact that numerosity is a highly adaptable feature has been taken by several authors as a clear indication of the existence of a dedicated brain mechanism for the processing of numbers, which would represent a basic and fundamental perceptual attribute (this is however a debated topic, so we will discuss it later in the specific section). In Chapter 3, Chapter 4 and Chapter 5, we will provide a more detailed discussion of how adaptation has been exploited to explore the mechanisms for time and number processing, and how such technique could be useful to shed light on the possible common mechanisms processing such features.

1.1.3 Space, time and number discrimination

However, the adaptation technique is only one example of the psychophysical paradigms exploited to investigate the machinery of perception. Other strategies to find out how the system works include studying perceptual judgments (with, for instance, the same discrimination task used to assess the effects of adaptation) modulating the features of the stimuli: how do changing a feature affects numerosity or temporal judgments? How does perceived duration change with increasing luminance, size, or the numerosity of a stimulus? How does perceived numerosity change if I join pairs of elements? Time and numerosity are two fields very rich of evidence obtained by this simple strategy, which provided clear indications of the relationship between them and other properties of the stimuli. So, such paradigms provide a good test bench to find how different combination of features affect temporal and numerical judgments, as well as for the investigation of mutual interactions between different magnitudes. So with some of the studies that we will present in this work, we also investigated how the representation of space, time and number changes modulating some other features and experimental parameters (Section 3.3, 4.2) or providing additional information in a different domain (Section 5.3 and 5.4).

1.1.4 Temporal order

A crucial aspect of time perception is the temporal order of two or more events. Investigations of perceived temporal order have been particularly focused on multisensory perception: on how we perceive the synchrony of two sensory signals in different modalities, in the face of a series of issues like different times of arrivals due to different information speed, difference in signal transduction and neural latencies, and so on (Matteson, 1970; Allison et al., 1983). Indeed, in our sensory world, many natural events are perceived as a combination of two or more signals: a ball hitting a wall generates both a visual signal (i.e. bounce) and an auditory signal (i.e. a sound). If the ball hit us, we also have the additional information provided by our somatosensory signals. While very often perception is investigated taking into account an individual modality, perception in the real environment is essentially multisensory (Vroomen & Keetels, 2010). Moreover, the perceived temporal order of two events has also proven to be not only determined by the physical properties of the stimuli, but it has pointed out that is a highly malleable perceptual feature, and that is subject to recalibration and adaptation effects (i.e. Fujisaki et al., 2004; Vroomen et al., 2004; Di Luca et al., 2009). The simplest paradigms exploited to study multisensory temporal order are the Synchrony Judgments (SJ) or Temporal Order Judgments (TOJ) paradigms, in which the subject has to simply report whether two stimuli (i.e. a sound and a visual flash, in the typical audio-visual paradigm) were presented simultaneously or not, or which one of the two was presented first. In the last part of Chapter 5 (Section 5.3), we will provide more detailed information about the literature on perceived temporal order, and we will report a study concerning how the temporal resolution of such judgments could be improved by predictions allowed by additional spatial information.

1.2 Conclusion

As we have emphasized at the beginning of this introduction, reaching a complete understanding of the dedicated and common sensory mechanisms for the processing of spatial, temporal and numerical information, is a task that goes beyond the purpose of this work, and that will probably require many more future investigations. In the next Chapters, however, we will present a series of evidence that may contribute to this higher purpose, showing, first, that the perceptual effects affecting the metrics of space are robust and independent from other confusing factor, and hence reflect only the underlying processes thought to support visual stability. Moreover, we will also show that time and number are features deeply rooted into the sensory processing streams, and provide evidence supporting the idea of a common magnitude system for space, time and number.

Chapter 2

Spatial vision: saccades and visual stability

2.1 Introduction

Research on spatial vision encompasses the study of processes and mechanisms that allow us to interpret the features of the visual world, exploiting many disparate techniques and paradigms – from single cell recording to neural network simulations, functional neuroimaging and human psychophysics. Here we will focus on one specific problem: visual stability and vision during eye movements. The question of how the brain achieves visual stability and perceptual continuity across each movement of our eyes is an old question that has fascinated many scholars across centuries. Indeed, even if most of the time we are unaware of it, our eyes shift from one position to another very rapidly (3-4 times per second), in order to explore the environment and search for useful information around us, or in response to salient stimuli, like a rapid change or the appearance of an object in the visual field. The purpose of such continuous scanning strategy is to achieve a more or less complete representation of the visual world and to inspect important features and objects, given that we can analyze with high details only the portion of the image that falls in the very center of the retina, i.e. the fovea. Regions of the image that fall outside the fovea are processed with increasing fewer details the more we move on the retina far away from it. Such high-resolution properties arise from the fact that the fovea is tightly packed with a high density of photoreceptor, while their density progressively falls toward the periphery of the retina. So, despite our common experience of a coherent and detailed visual world, the only region of the retina that can extract accurate information from the visual scene is the very small sized fovea, and to inspect the fine-grain details of different portions of an image, we need to move our eyes in order to reach the region of interest with our central fovea.

This could seem counterintuitive: why do not we have the same density of photoreceptor in any given point of the retina, in order to analyze the finer details of the image also in the periphery of the visual field? This specialization arise from a simple economic reason: if the human eye would have evolved a higher resolution in every given point of the retina, it would need so many photoreceptor that the optic nerve should be many times larger than it actually is (many centimeters) – a condition that is not actually compatible with our body structure. So, to analyze the information in the visual world while keeping a reasonable number of cells and axons, the best strategy is probably to develop only a tiny region with a high density of photoreceptor, and move it very rapidly and frequently across the visual field in order to get as more useful information as possible.

However, each eye movement – also termed "saccade" – poses a series of issues to the visual system. First of all, the eye is a highly unstable sensor, and the very high speed and rates of saccades cause very frequent and rapid shifts of the image on the retina. High-speed motion is a big problem for our visual perception, given that it could cause a blur of the image, due to the fact that the decay of activity of visual cells is not instantaneous (a feature called "visual persistence"). Moreover, from the unique point of view of a visual neuron – at least in the early stages of the visual pathways – a rapid shift of the image could be interpreted as an actual motion of the objects in the outside world: the topographic representation of the visual image shifts with every saccade, and neurons have to "figure out" whether some of the objects (or the entire visual world) are actually moving, or the motion is only due to the shift of the image on the retina. Nevertheless, we never perceive a blurred image when we saccade from an object to another, and we never perceive the visual world as moving – a condition that could be extremely confusing and disadvantageous for the survival and for the interaction with objects and other people. Conversely, irrespective of the number and magnitude of our eye movements, we always perceive a stable and relatively defined outside world. So, the brain must have developed a complex machinery to avoid blurred images and to distinguish between real object motion or motion of the retinae.

Another compelling problem for our visual system concerns the continuity of the image across different fixations. Despite the discrete nature of our visual perception, our conscious experience is continuous and seamless, without any interruption or discontinuity. How can the brain match pre- and post-saccadic images in order to keep a stable representation of the identities of the objects in the environment? One of the possible strategies to achieve such seamless continuity and correspondence is to generate a stable spatial map of the external world, in which objects are represented in spatiotopic (or at least craniotopic, head-centered) coordinates. This is a highly debated problem, and the points of view concerning this issue are usually polarized between the idea of a fast reconstruction of retinotopic maps from one fixation to another and the idea of a stable spatiotopic map, representing object positions in their real-world coordinates, that drives the matching of pre- and post-saccadic images. In the first case (the retinotopic hypothesis), the visual system would maintain only a map of what is currently on the retina at any given moment in time, and that this map would be simply updated from one fixation to another (for a review, see Wurtz, 2008). Conversely, according to the idea of a spatiotopic map, a more or less complete map representing the spatial (or at least cranio-centered) coordinates of the external objects would exist, and this map would be updated by the different retinal images across eye movements. Regarding the idea of a rapid updating of retinotopic maps across eye-movements, no higher-level spatial maps would exist across the different level of the visual pathways, and the external world would be represented only in retinotopic coordinates. Even if this topic is still subject of debate, evidence from both imaging (d'Avossa et al., 2007; Gardner et al., 2008; Crespi et al., 2011) and psychophysical studies (Melcher & Morrone, 2003; Ong et al., 2009; Melcher, 2005; Wittenberg et al., 2008) seems to point out the existence of spatiotopic maps. Overall, the picture emerged from such experimental observations, is that beside the transiently updated retinotopic maps, a slower (Zimmermann et al., 2013), more long-term map might be developed, which would keep the updated spatial information in a spatiotopic, real-world, reference frame (Burr & Morrone, 2012). Interestingly, such spatiotopic map would not be a static snapshot of the world, but a highly plastic map keeping information about many features of the external stimuli, such position, speed and duration, as well as their numerosity. This idea is indeed consistent with other accounts (Walsh, 2003; Burr et al., 2010) concerning the possibility that such attributes – space, time and number – might be intrinsically related, sharing functional mechanisms and neural substrates (see Chapter 5). Moreover, they also seem to share a spatiotopic map representing information in real-world coordinates.

2.1.1 Signals contributing to visual stability

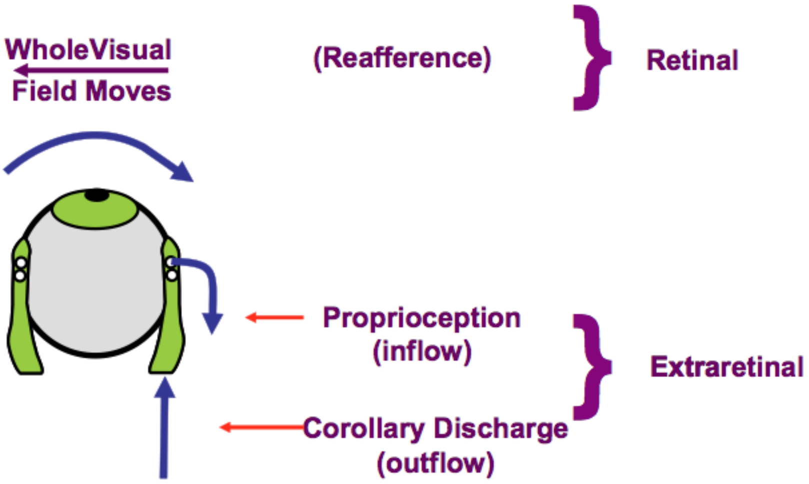

Figure 2.1: **Signals that might contribute to visual stability.** Three classes of signals are available for contributing to visual stability. One purely retinal signal is the “reafference” signal, provided by the whole-world motion created by a saccade. Conversely, the other two kinds are extra-retinal signal. The first is the proprioceptive signal signaling the position of the eye in the orbit. The second is the corollary discharge, which represents a copy of the motor command of the impending eye movement sent to the visual areas. Adapted from Wurtz, 2008.

Before discussing the evidence suggesting a representation of the visual world based on the updating of purely retinotopic map, or of a higher-order spatiotopic map, it could be useful to take into account which retinal and extraretinal signals might – or might not – contribute to visual stability, carrying information about the eye position and the parameters of each incoming saccade. How could the system discriminate between an object moving in the real world and motion created by the displacement of our eyes? Three different kinds of signals can contribute to disambiguate these two possibilities: the visual reafference signal, the proprioceptive eye position signal and the corollary discharge of the motor command sent to the oculomotor muscles (Figure 2.1).

First, the visual reafference signal (von Holst & Mittelstaedt, 1950) represents the whole-world motion signal created by the saccades. This optic-flow signal might be exploited to recognize self-motion instead of motion of the external environment. However, given the high speed of saccadic eye movements, such motion would be severely blurred, which makes this signal not so useful for fast movements, while could be more informative in the case of slower eye movements (i.e. smooth-pursuit eye movements) or whole-body motion through the environment.

Another signal that might be exploited by the brain to facilitate the perception of a stable world is the proprioceptive information signaling the eye position in the orbit. Several studies have examined this possibility. For instance, a critical test bench for this hypothesis would be to see what happens if the eye is artificially displaced during saccade preparation, and this is what was actually tested on the monkey across several studies. If proprioceptive information about eye position is taken into account, so an electrical stimulation to the brain that causes a shift of the eyes during the preparation of a saccade should lead to a compensation of the saccade dynamics for the new eye position created by the artificial displacement. However, it has been shown that a displacement created by stimulating directly the motor neurons does not lead to a compensation, while stimulating synapses at the level of the superior colliculus lead to a compensation that takes into account the new eye position created by the stimulation itself, suggesting that proprioceptive signals are not sufficient to carry positional information about the eyes, while a signal generated at higher-level oculomotor areas is required (Sparks et al., 1987; Schiller & Sandell, 1983; Sparks & Mays, 1983). Moreover, it has also been shown that compensation for eye displacement caused by stimulation at the level of superior colliculus occurs even if the proprioceptive afferents from the eyes are damaged, further suggesting that proprioceptive information does not play a role in the fast processes for visual stability that accompany each saccade. On the other hand, such signals may be more useful in the slower long-term calibration of the oculomotor system (i.e. to reduce the conflict between intended versus actually performed eye movements). Finally, the third class of signals that might be useful to maintain visual stability is represented by the "corollary discharge" (or "efference copy") signal (Sperry, 1950; von Holst & Mittelstaedt, 1950) – that is, a copy of the ocular-motor command also sent to visual areas to inform them about an impending eye movement. Several lines of evidence suggest that an extraretinal corollary discharge might be the most important factor allowing the maintenance of visual stability during saccades. First, Sperry (1950) observed that rotating the eyes of a fish by 180 degrees made it moving in circle. This finding was interpreted as a direct effect of a distorted corollary discharge signal resulting from rotating the eyes: normally, when the fish moves forward, a corollary discharge signals that the shift of the visual image on the retina is related to the self-motion; with eyes rotated, the corollary discharge not only failed to compensate for retinal motion, but even accentuate it, leading to circling. Also van Holst & Middlestadt (1950) found similar evidence for the importance of the corollary discharge (which they termed "efference copy", with the same meaning), studying the optokinetic responses of flies to rotating pattern of stripes. The idea of corollary discharge has however a long history, and many philosophers and scholars foresee the implication and the importance of an extraretinal signal for visual stability. For instance, Descartes speculated about such machinery, reasoning about the motion of the visual world occurring when he passively moved his eye by pushing on it, which created a markedly different experience compared to an active movement of the eye (Descartes, Treatise on Man). Von Helmholtz also recognized the implication of the motion of the visual world created by moving passively the eye, and referring to it as an "effort of will" conceptualized the role of a corollary discharge signal in avoiding such motion when the eye is moved actively. Moreover, recently it has also been shown that afterimages that are anchored to the retina (retinotopic) remain stable when the eye is moved passively, while they are displaced with the saccade when the eye is actively moved (Bridgeman, 2007). Another line of evidence suggesting the role of the corollary discharge signal comes from studies in which the subjects attempted to make eye movements while either oculomotor muscles or the entire body were paralyzed. In the Stevens et al.'s study (1976), the whole body of the experimenters (John Stevens) was completely paralyzed, and he experienced the impression of motion (i.e. impression of a general displacement of the visual field, not an actual motion of objects from one position to another) with each attempt to make a saccade (in the same direction of the intended saccade) – which is a clear index of the machinery trying to compensate for the eye-movement.

2.1.2 Perceptual distortions and physiological processes at the time of saccades

It has been shown that saccades are accompanied by several visual distortions, regarding the perception of stimuli presented very briefly around the time of saccadic onset (e.g. Lappe et al., 2000; Ross et al., 2001; Morrone & Burr, 2005; Binda et al., 2009). Although objects in the external environment usually last longer than a single perceptual "sample" (in this case, a fixation), presenting very short stimuli has been regarded has a promising psychophysical method to uncover the mechanisms for visual stability. What happens to such brief stimuli?

One of the first perceptual effects that have been considered is the so-called "visual saccadic suppression" or "saccadic masking". The idea that the visual system might block visual inputs during each eye movement was first developed more than a century ago, from the observation that for a human observer it is impossible to see its own eye movements in a mirror (Erdman & Dodge, 1898) – you can only see your eyes at the first position, and then in another position, while another observer can clearly see the movement of the eyes. Such process was first named "central anesthesia" – a momentary lack of sensitivity of the visual system to external inputs (Holt, 1903). The idea of a central anesthesia was later referred as saccadic suppression, and was related to the fact that we never perceive motion or the motion blurring that should follow from the high-speed shift of the eyes during a saccade (Burr et al., 1994). However, the idea of a complete anesthesia of the visual system was later challenged showing that the suppression is not complete, with several studies pointing out a 2-3 fold elevation in threshold during eye movements (Volkmann, 1962; Zuber & Stark, 1966; Riggs et al., 1976), that under normal viewing condition would be negligible. Moreover, such saccadic suppression of visual input in the form of threshold elevation would also be not sufficient to account for visual stability. Dodge (1900, 1905) and Woodworth however took a radically different point of view on saccadic suppression, hypothesizing that vision during saccades would not differ from vision during fixation, and the lack of sensibility to image motion (such seeing our own eye moving looking in a mirror) would be due to the high speed of saccades – too high to resolve the visual image. Anyway, even such explanation has proven to be erroneous. Indeed, first it has been showed that object motion could be resolved at velocities greater than 1000 deg/s (Johnstone & Riggs, 1979), and that fast object motion does not decrease visual sensitivity (Burr & Ross, 1982), but only shifts the contrast sensitivity curve down on the spatial frequency axis. So, a motion at 200 deg/s – a typical value for saccades (Carpenter, 1977) would easily blur the high end of the spatial frequencies spectrum, but concurrently enhancing the sensitivity for lower spatial frequency, that under conditions of stationary viewing would not be visible. So, given this enhancement of sensitivity for very low spatial frequencies, the question is why such low frequencies are not perceived during fast eye movements. To shed light on this issue, and on the issue of motion of the visual image during saccades, Burr et al. (1982) investigated how the selectivity and sensitivity to different spatial frequencies change during saccades, pointing out that contrast sensitivity for low spatial frequencies in substantially reduced during saccades, and that this suppression results from the selective damping of motion-selective mechanisms.

How does this perceptual damping of the motion-selective system is implemented by the brain? Regarding the physiological processes underlying such saccadic suppression of visual information, it has been recently shown that before each saccade the magnocellular visual pathway is selectively suppressed (Burr et al., 1994): contrast sensitivity for luminance-modulated low-spatial frequency sinusoidal gratings is strongly reduced when stimuli are presented during saccades, whilst sensitivity to high-spatial frequency or isoluminant gratings is spared. This pattern of results clearly show a selective reduction in the activity of magnocellular neurons, which are color-blind and most sensitive to low spatial frequencies and high temporal frequencies, while parvocellular neurons, which are sensitive to color-defined and high-spatial frequency stimuli, seem not affected by saccade execution. A selective suppression of the pathways related to low spatial frequencies and motion information can thus account for the elimination of image blurring and world-motion during saccades.

Related to this suppression of motion and low-frequency information, another perceptual effect observed at the time of eye movements is the saccadic suppression of image displacement. Such effect was first discovered by Ditchburn (1955), and then rediscovered independently by Wallach & Lewis (1966) and Brune & Lücking (1969). The saccadic suppression of displacement, which concerns the observation that a sudden displacement of a stimulus at the time of a saccade goes unnoticed, shows that information about the position of a visual object is severely degraded during eye movements, and large displacements of the visual world do not reach awareness. Interestingly, such lack of perceptual awareness seems not paralleled by corresponding behavioral errors in spatial orienting. Indeed, it has been shown that motor orientation responses, such pointing toward a visual target, are not affected by the suppression of displacement: human observers can point correctly toward a visual object, irrespective of the fact that they had or had not detected its displacement (Bridgeman et al., 1975; Prablanc & Martin, 1992), suggesting that the information about the displacement is preserved at same stage, but it is not available to perception.

Another of the major saccadic effects is the peri-saccadic compression: if a subject is asked to indicate the position of a stimulus presented just before or during the execution of a saccade, the response is usually wrong. Indeed, such peri-saccadic stimuli are dramatically mislocalized away from their actual position, and usually perceived near or superimposed to the saccadic target. This distorted perception of stimuli positions, which leads to error in localization performance, resembles a shift of perceived locations in the direction of the eye movement, toward the saccadic target. But what happens if you present a stimulus outside the trajectory of the saccade, beyond the target of the eye movement? Even in this case, the apparent position of a brief peri-saccadic

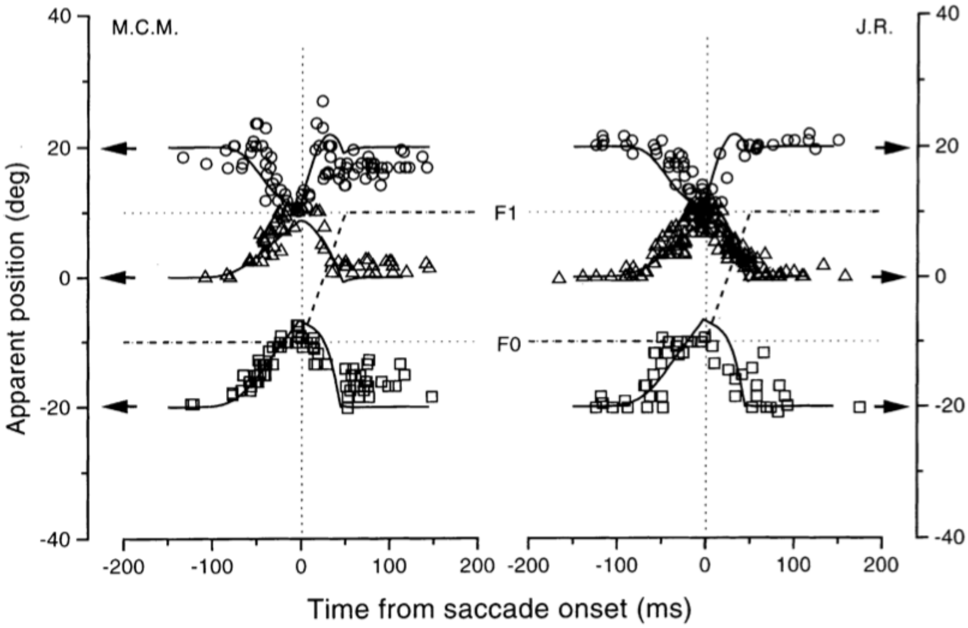

Figure 2.2: **Apparent position of peri-saccadic stimuli.** Apparent position of a vertical bar briefly flashed at position -20 deg, 0 deg (center of the screen) and +20 deg, as function of the presentation time relative to saccadic onset (note that the initial fixation point is located at -10 deg, while the saccadic target at +10 deg, relative to the center of the screen). Each point represents a single observation. The dashed line represents the saccade trajectory from the initial fixation point to the saccadic target. Solid curves represents a simulation of the localization behavior (See Ross et al., 1997 for further details about the model). Adapted from Ross et al., 1997.

stimulus would be distorted, and the stimulus would appear to be displaced towards the saccadic target or coincident with it, displaced in the direction against the saccade trajectory (Figure 2.2).

Regarding the neural substrates and processes related to the perisaccadic compression, it has been proposed that the "predictive remapping" of visual receptive fields – a process that is thought to be related to visual stability – could account for such mislocalization effects (Burr & Morrone, 2010). Remapping of visual receptive fields (RF) was first noted by Duhamel and colleagues (1992), recording the activity of visual neurons before and during saccades. They observed that just before the onset of an eye movement, the RFs of many neurons in the lateral intraparietal cortex (LIP) shifted in the direction of the incoming saccade, starting to respond to stimuli presented in the posi-

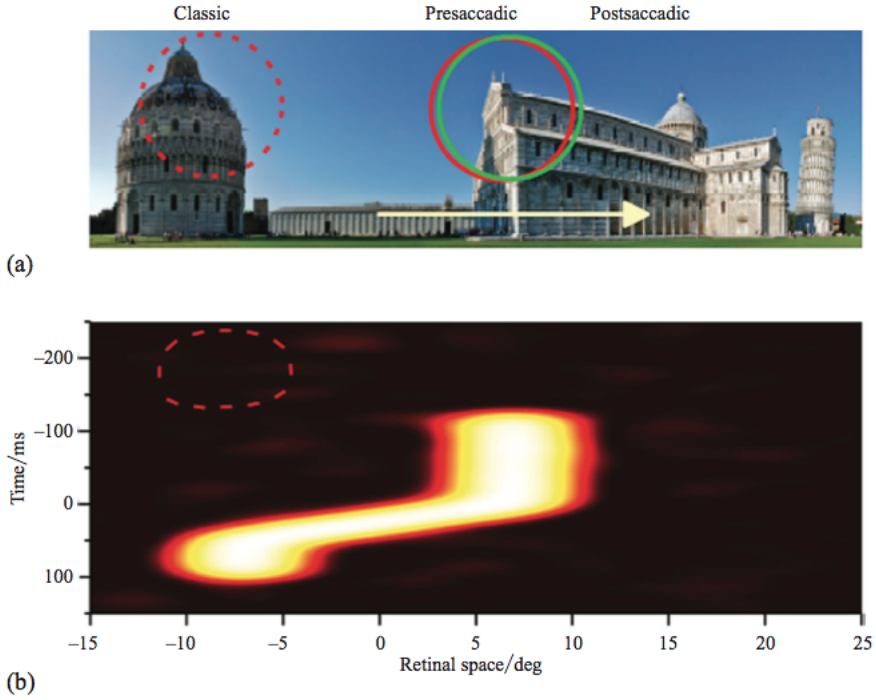

Figure 2.3: **Predictive remapping.** (a) Illustration of the “remapping” process. At the initial fixation stage, the “classical” receptive field falls on the left side of the image (the baptistery; dashed red circle). Then, just before the onset of the eye movement (white arrow shows the change in gaze position), the cell remaps its receptive field displacing it to a distance equal to gaze shift that will be created by the impending saccade – so, the remapped receptive field falls on the facade of the cathedral (solid red circle). As the eye moves, the receptive field returns to its classical position (on the retina), thus, it remains fixed (in spatial coordinates) on the facade until the end of the saccade. (b) Schematic representation of the spatiotemporal response selectivity of the remapping neuron as it returns back to its resting position. When represented in retinal coordinates, this “spatiotemporal” receptive field is oriented along the same trajectory of the gaze shift, and hence annuls it. The dashed ellipses show the tuning of the receptive field before the remapping. Adapted from Burr & Morrone, 2012.

tion that they will occupy after the completion of the movement (the so-called “future” RF). While neuronal receptive fields were considered a fixed properties, resulting from hardwired connections, such neurons – termed “remapping neurons” – nevertheless showed the ability to shift their RF in order to anticipate the movement, responding to what they will “see” after the conclusion of the saccade (Figure 2.3).

Could this anticipatory shift of visual RF be linked to visual stability? At first glance, the RF shift toward the future position not only would not be useful to solve the problem of visual stability, but it could even exacerbate it. Indeed, first the RF shift on the retina, then the retina itself shift with the movement, actually doubling the displacement of the RF from its original position (Burr & Morrone, 2012). However, the crucial factor of this process would not be the initial displacement, but the transient nature of such process, and the returning of the RFs to their "classical" position during the execution of the saccade (Burr & Morrone, 2011). Indeed, after the initial shift, when the saccades actually starts the RFs gradually return to their starting positions: doing so, the RFs remain fixed on the same position while the eyes are moving, in the time comprised between the end of the previous and the start of the new fixation, thus maintaining perceptual continuity by means of a sort of "transient craniotopicity". While physiological studies observed the initial shift, the returning of the RFs toward their resting state position has not been explicitly documented yet. However, it must occur at some stage, since after the movement the cells continue to respond according to their basic RF. During such period, the spatiotemporal dynamics of the movement creates a RF slanted onto the retinal space, in the same direction of the saccade – thus counteracting the displacement caused by it (see Figure 2.3b).

Even if such return to the resting position during the actual execution of the saccade, and the related period of transient craniotopicity created by the relaxing RF, has not been yet observed experimentally, also several indirect observations seem to support this idea.

First, the pattern of localization of brief stimuli presented around the time of an eye movement (Honda, 1989; Ross et al., 1997; Morrone et al., 1997) which, as mentioned above, resembles a compression of the visual space around the location of the saccadic target – and particularly the dynamics of mislocalization of pairs of stimuli – could give us some insight about the transient craniotopicity hypothesized to play a role in the maintaining of visual stability. Recently, Cicchini and colleagues (2013) studied the relation between the phenomena of perceptual stability of steady objects and the mislocalization of brief stimuli testing the pattern of mislocalization of pairs of bars briefly presented near the onset of a saccade. To do so, they presented the test stimulus – a white bar – with variable timing around the saccadic onset, preceded or followed by a similar black reference bar, and asked the participants to localize both the stimuli.

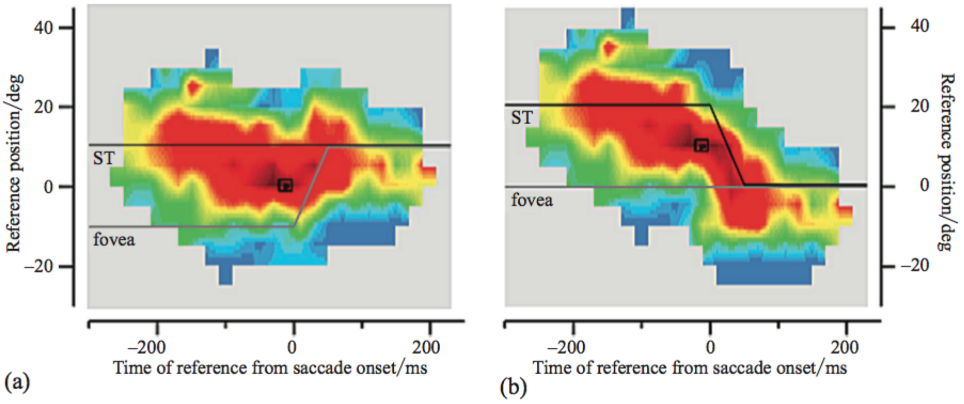

Figure 2.4: **Spatiotemporal interaction field.** Spatiotemporal map of interactions between a perisaccadic probe presented between -20 and 0 ms at screen position -1 deg (black square). The axis show the time of the reference bar (abscissa) and the horizontal position of the reference bar (ordinate) in both (a) screen and (b) retinal coordinates. The gray lines represent the position of the fovea, while the black lines represent the position of the saccadic target. The interaction indices at these coordinates were interpolated and smoothed to generate the map (note that hot colors indicate strong interactions). Adapted from Burr & Morrone, 2012.

What they found, is that while in the cases where the stimuli were presented well before or after the saccade localization, performances were fairly accurate; on the other hand, when they were presented around the onset of the saccade the perceived position of the test bar was strongly attracted toward the reference bar – a result that shows that the classic pattern of mislocalization is severely altered when a reference bar is presented before or after the test bar. This effect of attraction, which however does not confound the identities of the two stimuli, is also specific for the orientation and location of the bars, and does not occur with orthogonal stimuli or with large vertical displacements. Interestingly, such interaction between the test and the more stable reference bar occur over a huge spatiotemporal range, spanning about 20 degrees of visual angle and 200 ms (Figure 2.4): across the positions and timing showed in Figure 2.4, the two bars are seen together in the same position. When such spatiotemporal map of interactions, showing the strength of the attraction between the two stimuli, is plotted in screen (real-world) coordinates (Figure 2.4a), the field is oriented quite symmetrically along the time axis, but when the interactions are represented in retinal coordinates (Figure 2.4b), the field assume a clear slanting in space-time, following the trajectory of the saccade. So, this slanted shape of the interaction field shows that perceptually similar stimuli are likely to be integrated on the basis of their position in external, and not retinal, coordinates. What is the source of this spatio-temporal interaction field? On the one hand, these results support the idea of an extraretinal signal (see Section 2.1.1) involved in triggering the predictive RF shift and their later returning to the resting position, and thus creating the slant along the saccade trajectory. However, such mechanism is not sufficient, given the spread of the response field in space and time, and might only set a broad map over which stimuli could be integrated. While this mechanism selects the right features that should be integrated, another visual mechanism is needed, that would perform such integration under the constraints set by the remapping process.

Intriguingly, a similar receptive field slanted in space-time was also observed studying saccades with a very different technique – that is, classification images (Panichi et al., 2012). Such technique consists of a very simple task (i.e. detecting a brief target and reporting whether it was presented in the upper or lower visual field) target embedded in noise, and then analyzing the noise patterns leading or not to a correct detection of the target. By adding al the correct detection and the "false alarms" (i.e. the trial in which the subject reports erroneously to have detected the target) and subtracting the trials that conversely led to correct rejections or misses (i.e. when the target was present but the subject did not report it), lead to the reconstruction of a global mean noise image (a "classification image") that is thought to represent the shape of the response field of the neural mechanism underlying the process. When applied to brief perisaccadic stimuli, this technique showed a classification image that extends well beyond the stimulus both in space and in time, closely resembling a spatiotemporal receptive field, and very similar to what was later found by Cicchini et al. (2013) even with a very different task.

2.1.3 Conclusion and outline of the next sections

Taken together, these results suggest the involvement of spatio-temporal receptive fields slanted along the direction of the saccade, that integrate the feature of the external objects and maintain perceptual stability by means of a transient craniotopicity – a process that involves an extraretinal signal, triggering the start of the remapping process, and visual processes integrating the stimuli. Such conclusion is supported by many studies analyzing the pattern of visual distortions at the time of saccades, but, however, most of these effects are studied under very controlled circumstances, with very brief, unnatural, stimuli, and with paradigms requiring a large amount of saccades. On the other hand, none of the aforementioned distortions could be noticed in more common, real-life, circumstances, leaving the possibility that the level of saccade automaticity and stereotyping might have a role in the patterns of results obtained in eye movements studies. In the next chapter, we will report a study (Fornaciai & Binda, 2015) aimed to test whether such spatial distortions at the time of saccades are actually related to the eye movements and determined by the ocular-motor parameters, or whether they could be only a byproduct of the unnatural context in which they are studied, were subjects perform a great number of highly stereotyped eye movements and have to cope to very brief stimuli.

2.2 The effect of saccade automaticity on perisaccadic space compression

2.2.1 Introduction

Saccades pose a major perceptual problem in that each new eye movement changes the mapping of external objects on the retina. While the visual system is usually able to overcome this challenge and ensure stable and seamless perception, there are conditions where vision is disrupted during saccades: for example, when visual stimuli are brief, flashed in and out of view within some 100 ms before or after the onset of the saccade.

These briefly presented stimuli are subject to systematic perceptual distortions, but it is important to note that all these distortions are never observed in everyday life, for two main reasons. First, illusions only occur for briefly flashed stimuli, which are very rare in natural scenes. Second, and most importantly, illusions vary widely with the time of flash occurrence relative to the onset of a saccade. As a consequence, perisaccadic distortions are typically observed in laboratory conditions that encourage production of highly stereotyped eye movements: saccades directed to one or few target locations, repeated thousands of times so as to focus data collection in the short perisaccadic temporal window where illusory perceptions occur. These conditions might differ in important ways from natural eye movement behavior, especially considering that different types of saccades are associated with different patterns of brain activity (Johnston & Everling, 2008; McDowell, Dyckman, Austin, & Clementz, 2008).

In particular, saccade control recruits three areas, the Lateral Intra-Parietal cortex, the Frontal Eye Fields and the Superior Colliculus, which are also involved in the representation of perisaccadic visual space – as indicated by the fact that visual receptive fields change transiently at about the time of a saccade in LIP (Duhamel, Colby, & Goldberg, 1992), FEF (Umeno & Goldberg, 1997; Sommer & Wurtz, 2006; Zirnsak, Steinmetz, Noudoost, Xu, & Moore, 2014) and SC (Walker, Fitzgibbon, & Goldberg, 1995). Modulation of activity in all three areas has been observed in experiments comparing the execution of saccades with different levels of volitional control or automaticity. Fronto-parietal activity associated with saccade execution differs depending on whether saccades are specified based on symbolic instructions (e.g. anti-saccades) or spontaneously triggered by the presentation of peripheral targets (targeting saccades) – as observed in several neuroimaging studies of human subjects (McDowell et al., 2008) and specifically in the FEF (Everling & Munoz, 2000) and LIP (Gottlieb & Goldberg, 1999) of non-human primates. Both FEF (Bruce & Goldberg, 1985) and LIP (Gottlieb, Kusunoki, & Goldberg, 1998) activity depends on the salience of the saccade target stimulus, with enhanced responses to salient visual transients (e.g. abrupt onset of a stimulus) targeted by spontaneous saccades. Although they are a simple and spontaneous behavior, saccades targeting visual transients can be automatized with practice. Repetition of saccades in one direction over hundreds of trials leads to shorter and less variable saccade latencies (Basso & Wurtz, 1998) (i.e. saccade reaction times), and this is accompanied by stronger neural activity preceding saccade execution in Superior Colliculus (Basso & Wurtz, 1997, 1998). These changes occur in the intermediate layers of the Superior Colliculus, which project to the Frontal Eye Fields and have been shown to carry a ‘corollary discharge’ of the eye movement command – that is, a copy of the eye movement command that may help maintaining a stable representation of space in spite of the displacement of retinal images caused by the saccade (Sommer & Wurtz, 2002, 2006, 2008). Thus, there is evidence that the level of automaticity with which saccades are executed has an impact on neural activity in just those areas that play a key role in the representation of visual space. This raises the question whether visual space distortions observed during saccades are specifically associated with the execution of highly stereotyped saccades – rather than spontaneous, natural phenomena.

Here we set out to directly address this question, using two experimental approaches. In the first, we tested the effect of practice on a repetitive saccade task. We compared a group of expert participant with a history of several thousands trials on the specific saccade task we used, with a group of novices, who had never participated in eye-movement related research, and practiced the saccade task over the course of the experiment. Second, we studied how the localization performance of experts changed when the saccade task was modified to increase the level of volitional control. We changed the stimulus display so that the saccade target location was marked by a steadily visible stimulus, and subjects initiated saccades upon hearing a sound. The removal of the peripheral onset and the added complexity of the task, which no longer exploited our spontaneous tendency to saccade at a sudden peripheral onset, should lead to modification of saccade parameters – notably, latency (Walker, Walker, Husain, & Kennard, 2000; Rolfs & Vitu, 2007).

Our prediction is that if perisaccadic compression is specifically associated with the repetitive execution of stereotypical saccades, it should be reduced in conditions that promote the variability of saccade parameters, and in novice observers – with an idiosyncratic pattern of perisaccadic mislocalization that gradually normalizes with practice.

2.2.2 Materials and methods

Subjects

Twenty-four subjects participated in the experiment, after giving their written informed consent. Sixteen of the subjects were completely inexperienced in psychophysical experiments involving eye movements ('novices') and were tested with the basic 'abrupt onset' paradigm; data collection started immediately after the instructions were given (no practice trial). The other eight were 'expert' observers (with a history of >1000 trials in experiments involving eye movements and including the two authors); these were tested with both the 'abrupt onset' and a less usual 'steady-on' paradigm (described below). All subjects reported normal or corrected to normal vision. Experimental procedures were approved by the local ethics committee (Comitato Etico Pediatrico Regionale - Azienda Ospedaliero-Universitaria Meyer - Firenze (FI)) and are in line with the declaration of Helsinki.

Apparatus

The experiment was performed in a quiet and dark room. Subjects sat in front of a monitor screen (40 x 30 cm) at a distance of 57 cm, with their head stabilized by a chin rest. Stimuli were generated using the PsychoPhysics Toolbox routines (Brainard, 1997; Pelli, 1997) for Matlab (MatLab r2010a, The Mathworks, inc.) and presented on a CRT monitor (Barco Calibrator Line) with a resolution of 800 x 600 pixel and a refresh rate of 120 Hz, driven by a Mac Pro 4.1. Eye movements were monitored in-synch with visual presentations using the EyeLink 1000 system (SR Research, Canada) and the Eyelink toolbox for Matlab (Cornelissen, Peters, & Palmer, 2002). Eye position and pupil diameter of the left eye were measured with a frequency of 1000 Hz by means of an infrared sensor mounted below the screen, which allowed for unrestrained binocular viewing of the display. At the beginning of the experimental session, a standard 13-point calibration routine was performed.

Stimuli

Stimuli were presented against a homogeneous red background (Commission International d'Eclariage coordinates: x = 0.624, y = 0.343; luminance = 23 cd/m^2); the localization probe was a green vertical bar (CIE coordinates: x = 0.292, y = 0.597; luminance = 60 cd/m^2), 1 deg wide, straddling the full screen height (30 deg). It was presented for a single monitor frame, at variable horizontal locations and timings.

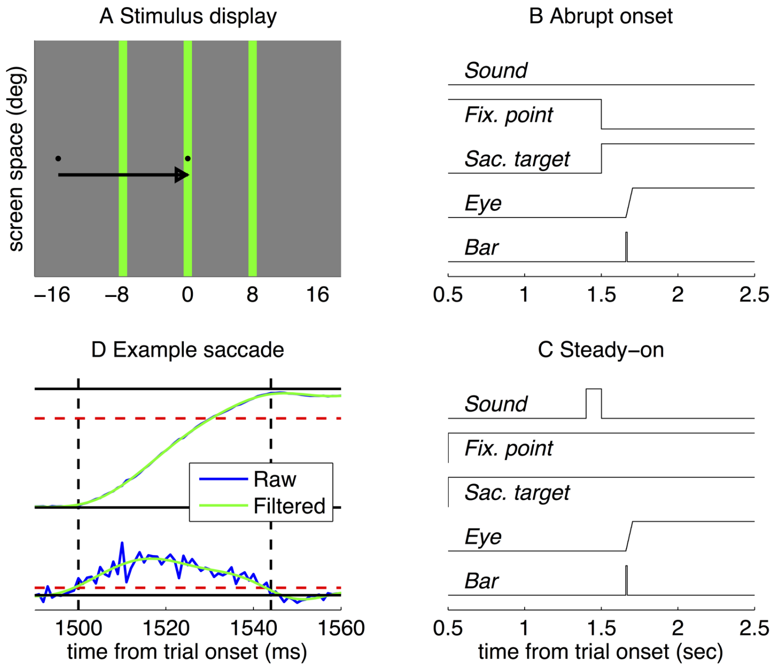

Figure 2.5: **Methods.** A. Approximately to-scale representation of the stimulus display; the arrow shows the required saccade (not part of the display) and the three vertical green bars indicate the tested positions (only one was tested in each trial). B-C. Timecourse of presentations in the abrupt onset target and the steady-on target paradigm respectively. D: example eye position (top) and velocity traces (bottom), before and after the Fourier filtering. Dashed red lines show the velocity threshold for detection of saccade onset and offset (marked by the vertical dashed black lines) and the exclusion criterion for saccade amplitude (the saccade had to be >12 deg, 3/4 of the required saccade amplitude).

Task and procedure

In all cases, subjects made rightward saccades from a small black square located 16 deg left of screen center (the fixation point) to an identical black square presented at the screen center (the saccadic target) – see Figure 2.5A. The timecourse of presentation of the two stimuli, however, differed in the two tested paradigms – see Figure 2.5B-C. In the 'abrupt onset' paradigm, the saccade target appeared after a variable delay (1500 ± 100 ms) and simultaneously with the extinction of the fixation point, giving the impression of a point that jumped between the two locations and that the subjects were instructed to promptly follow with their eyes. In the 'steady-on' paradigm, both the fixation point and the saccade trials remained always visible; subjects were signaled to start a saccade by a sound (100 ms white noise burst, approximately 80 dB). In all cases, the probe bar was presented at one of three horizontal positions, randomly intermixed across trials: at -8, 0 or +8 deg relative to the saccadic target (negative meaning leftward; Figure 2.5A). About 1 sec after the probe bar presentation, the mouse cursor appeared at a random position (drawn from a circular Gaussian distribution with mean at screen center and standard deviation of 4 deg) and subjects adjusted it to match the perceived position of the bar. In the rare cases where subjects failed to detect the bar ($2.4 \pm 0.7\%$ of trials), they were instructed to click in the bottom left corner of the screen – so that the trial could be discarded from the analyses. Response collection triggered the start of the next trial. Trials were administered in blocks of 24 separated by short breaks; each subject was tested in at least ten blocks per experiment. In the first five blocks, the probe bar was presented immediately upon detection of the saccade onset – calculated online as the first of two consecutive time points where horizontal eye velocity exceeded 100 deg/sec. In the rest of the blocks, the time of bar was defined a priori based on the subject's saccade latency (median across all previous trials in the experiment) and an average intended delay of ± 50ms. Triggering in the first half of the trials was aimed to avoiding that the early trials of novice subjects be wasted over non-perisaccadic bar presentations, maximizing the probability to reveal practice effects. However, for simplicity, the same procedure was also adopted with expert subjects. All blocks also included a minority of trials where the signal for starting the saccade was withheld and localization was measured during fixation.

Data analyses

An offline analysis examined the horizontal eye position traces. Visual inspection indicated some contamination from high-frequency noise. We therefore proceeded to filter the traces in the frequency domain (applying Fast Fourier Transform or FFT to each trace, multiplying the frequency spectrum by a Gaussian centered at 50 Hz with standard deviation of 10 Hz respectively, then reconverting the traces to the time domain via the inverse FFT). Figure 2.5D shows the raw and filtered traces for eye position and velocity for a representative trial. Filtering had virtually no impact on the estimation of any saccade parameter except for peak velocity – the absolute values of which were of course higher for the unfiltered traces and less orderly related to saccade amplitude. We then re-estimated saccade onset (first of two samples exceeding 100 deg/sec velocity) and consequently re-determined the time of probe-bar presentation relative to the saccade; we also estimated the other saccade parameters: offset (first of two samples falling below the 100 deg/sec threshold), duration, amplitude and peak velocity. Trials were discarded from further analyses if no saccade could be detected ($< 1\%$ in all cases), the saccade was anticipatory (negative latency, $8.6 \pm 3\%$ and $4.0 \pm 2\%$ in the 'abrupt onset target' paradigm, for novices and experts respectively; $5.3 \pm 2\%$ for Experiments in the 'steady-on target' paradigm) or smaller than 12 deg (3/4 of the required amplitude, $7.2 \pm 2\%$, $7.5 \pm 4\%$, $5.4 \pm 2\%$), leading to the exclusion of $18.4 \pm 4\%$, $12.3 \pm 6\%$ and $11.0 \pm 4\%$ of trials. Localization performance in the valid trials was analyzed by sorting trials based on the time of bar presentation relative to saccade onset; for each tested position of the bar, average localization was computed in partially overlapping bins of 30ms width (overlap: 20ms); the curves averaged across observers are shown in Figure 2.6. We summarized localization performance by computing, in the same bins, a compression index and a shift index. These were estimated in Lappe et al. (2000): the compression index is the standard deviation of reported bar positions, normalized by the standard deviation of the actual bar positions; the shift index is the mean of the reported bar positions relative to the mean of the actual bar positions. Based on these definitions, a compression index of 0 indicates maximal compression and 1 indicates no compression; a positive shift index indicates an overall bias in the direction of the saccade. We verified that using an alternative definition of compression and shift indices, e.g. slope and intercept of the reported vs. actual positions as in Zimmermann et al. (2014), did not alter the results. Statistical comparisons across groups/conditions were performed on a subset of trials, where the bar was presented within 15 ms from the saccade onset (i.e. between -15 ms and +15 ms from saccade onset) leading to the maximum expected mislocalization. For the analysis of practice effects in novices, trials in this bin were sorted by presentation order and mislocalization indices and saccade parameters were computed in the first through tenth percentiles. Besides performing ordinary statistical tests (t-tests) we evaluated their statistical power using Bayesian methods. Specifically, we computed the Bayes Factor following Dienes (2014). This requires entering a description of the data (t statistics and its standard error) and a description of the theory (the distribution of the statistics under the "alternative hypothesis"). We corrected the standard error of the t-statistics by 1 + 20/(df x df) as recommended for degrees of freedom df < 30. We only computed the Bayes Factor for comparisons of compression index values, were the distribution of the statistics is easily defined: a uniform distribution bound within -1 and 1 (the difference between compression indices, which vary between 0 and 1). Conventionally, a BF larger than 3 implies support for the alternative hypothesis, BFs between 1/3 and 3 indicate weak or no evidence for either hypothesis and a BF smaller than 1/3 implies strong evidence in support of the null hypothesis.

2.2.3 Results

Our first experiment compared perisaccadic localization in two groups of subjects: expert subjects (with long prior experience in experiments employing the same eye movement task, here referred to as the abrupt-onset paradigm) and novices (with no prior exposure to psychophysical tasks involving an eye movement). We analyzed their eye movement performance – in terms of the main saccade parameters: peak velocity, amplitude and latency – and how these changed over the course of the experiment, i.e. with practice.

As expected, practice had a strong effect on saccade latencies. Their distribution differs markedly between groups (Figure 2.6A), being much longer in the novices (two-sample t-test: t(22): 6.21, p<0.001) and more variable from trial to trial (two-sample t-test on the standard deviation of saccade latency values: t(22): 6.08, p<0.001). Over trials, novices' saccade latencies decreased (Figure 2.6B; paired t-test comparing latencies in the first vs. the fourth quartile of trials: t(15): -6.05, p<0.001) and so did their trial-by-trial variability (paired t-test: t(15): -4.77, p<0.001, not shown). At the beginning of the experiment, novice saccade latencies were almost twice as long as the experts; within the approximately 200 collected trials, they normalized to values similar to that of the experts (though still slightly longer, two-sample t-test comparing saccadic latencies in experts vs. the novices' last quartile of trials: t(22): 2.41,

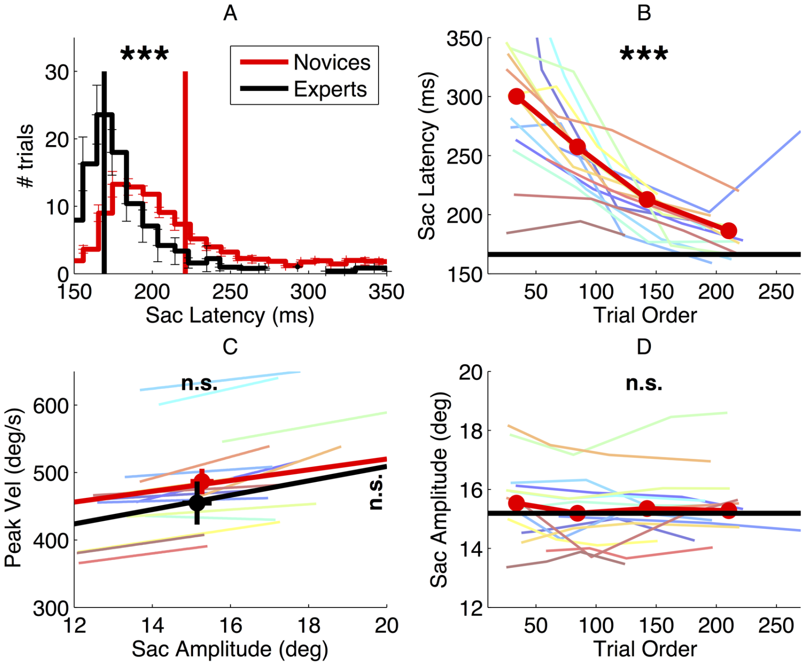

Figure 2.6: **Effect of practice on saccade behavior.** Saccade latency and main sequence in novice vs. expert observers and evolution of novices' saccade behavior over the course of the experiment. A: saccade latency histogram, computed for the individual subjects and then averaged, with error bars showing s.e. in each bin. Vertical lines mark the grand averages of saccade latency. Asterisks at the top of the panel report the result of a two-sample t-tests comparing the averages in the two groups (*** = $p<0.001$). C: saccade main sequence, plotting peak velocity against amplitude. Colored lines give the best fit line for the individual novice subjects; thick red and black lines give the average fit for the two subjects groups, and symbols with errorbars show grand averages and their s.e. The results of two-sample t-tests comparing the average peak velocity and amplitude in the two groups are marked on the right and at the top of the graph respectively (n.s. = not significant). B-D: Saccade latency and amplitude for novice subjects, computed after ranking trials according to their presentation order and averaging values in the four quartiles; colored lines show the individual novice subjects, the thick red line gives the averages in the novices group; asterisks report the results of paired t-tests comparing the first vs. the last quartile. For reference, the average in the experts' group in also shown (black line).

$p<0.05$).

Practice did not affect the other saccade parameters (Figure 2.6C-D). Saccade amplitude and peak velocity are comparable between novices and experts (two-sample t-test on amplitude values: t(22): 0.27, p=0.791; two-sample t-test on peak velocity values: t(22): 0.96, p=0.348) and so is the variability of these indices across trials (all $p>0.2$). Moreover, neither parameter changed systematically over the course of the experiment (paired t-test comparing saccade amplitude in novice early vs. late trials: t(15): -1.86, p=0.083; peak velocity: t(15): -0.37, p=0.718). Having established that practice had a robust effect on key descriptors of saccade behavior, namely saccade latency and its trial-to-trial variability, we went on to analyze its effect on localization performance.

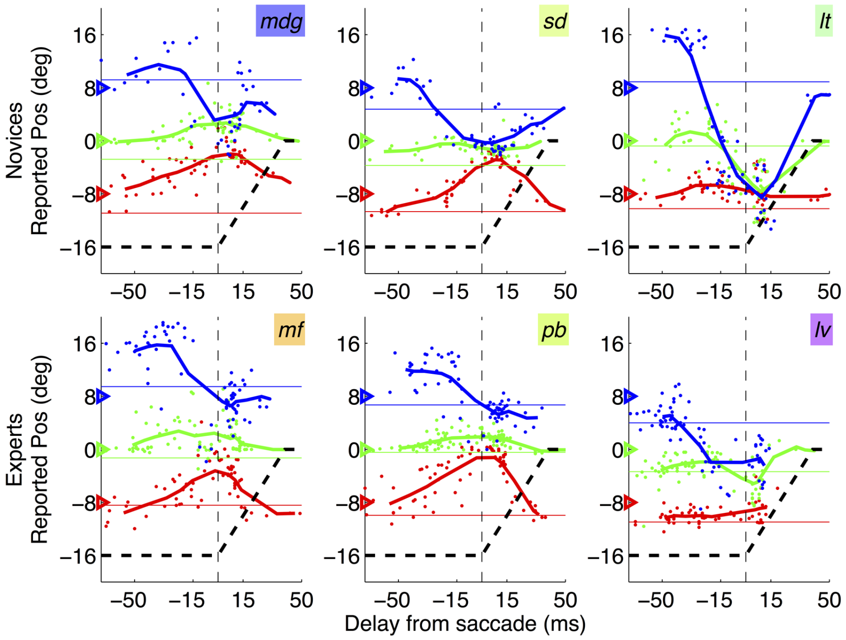

Figure 2.7: **Mislocalization curves in novices vs. experts.** Pattern of mislocalization in three representative novice subjects (top row) and three expert subjects (bottom row), all tested with the abrupt onset paradigm. Thick lines give running averages of localization judgments (averages in bins of 30 ms stepping by 10 ms) and dots show individual trials data. Subjects' initials are given by the text insets, with colors corresponding to those used in figures 2 and 4 (novices) and 5 (experts). The triangles on the left y-axis indicate the veridical position of the bars; the black dashed line shows the trajectory of the saccade, from the fixation point (-16 deg) to the saccade target (0 deg, or screen center).

Figure 2.7 shows results from three representative novice subjects and three experts (top and bottom rows respectively), plotting the average reported position of each tested position as function of the delay of bar presentation relative to the saccade onset. In all cases, the pattern of localization errors is consistent with the established phenomenon of perisaccadic compression – stimuli presented at about the saccade onset tend to be seen as compressed within a small spatial region. There is substantial inter-subject variability as to the 'focus of compression': bars may be seen as compressed toward the saccade target or left/right of it. However this variability, which was noted previously (Morrone, Ross, & Burr, 1997), appears to be present in both novice subjects and experts – and it did not correlate with any of the analyzed saccade parameters, including saccade landing (see correlation analyses below). In order to test for systematic differences of localization behavior between the two groups, we defined two indices of perisaccadic mislocalization: compression and shift index – as in (Lappe et al., 2000).

Figure 2.8A shows that novices and experts have closely matched values of compression index (two-sample t-test in the 30 ms straddling the saccade onset: t(22): -0.93, p=0.364). Bayesian statistics confirm that the lack of statistical significance is not due to lack of power: the Bayes Factor (BF) given the observed data (t: -0.93, se: 0.09) is 0.16 – a BF $< 1/3$ is conventionally interpreted as strong evidence for the null hypothesis, implying that compression is the same in novices and experts. Figure 4C shows shift index values, which are very variable across subjects (reflecting the idiosyncratic 'focus of compression' seen in Figure 2.7). There is a tendency for more negative values in the novices group, but the difference does not reach statistical significance (two-sample t-test in the 30 ms straddling the saccade onset: t(22): -0.96, p=0.347).

As a more direct test for practice effects, Figure 2.8B&D show the variation of novice compression and shift indices over the course of the experiment. Comparing early vs. late trials (the first vs. the fourth quartile), we found no trend for the compression index to change with practice (paired t-test: t(15): 1.08, p=0.299). Again we use Bayesian statistics for the comparison of compression index values and find that the Bayes Factor (BF) given the observed data (t: 1.08, se: 0.07) is 0.15, i.e. BF $< 1/3$ or strong evidence in support of the null hypothesis of no variation of compression index values. The shift index computed in the last quartile of the experimental trials is reliably more negative than at the beginning of the experiment (paired t-test: t(15): -4.44, p<0.001). However, this can hardly be interpreted as an effect of practice, given that the shift index in the novices grew progressively apart

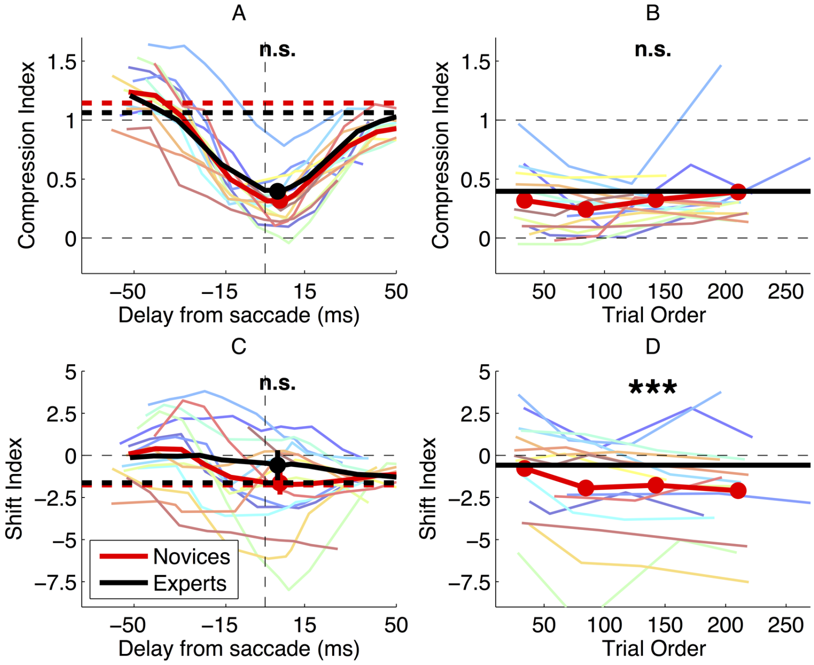

Figure 2.8: **Effect of practice on mislocalization indices** Mislocalization and saccade behavior in novice and expert subjects. Two-sample t-tests compared the averages in the two groups; the results (n.s. = not significant, * = p<0.05, ** = p<0.01, *** = p<0.001) are marked at the top of the panel (top and right in panel B for values on the x and y-axis respectively). A & C: compression and shift index plotted against the time of bar presentation. Thin colored lines show timecourses for the individual subjects (running averages as for Figure 2.7). Thick lines give the averages across subjects over the timecourse (continuous curves) and in fixation (dashed horizontal lines); filled symbols and their error bars show averages and s.e. in the 30 ms bin spanning the saccade onset, which the t-tests compared (results given at the top of the graph; n.s. = not significant). B & D: evolution of compression and shift index values over the course of the experiment. Trials with the probe presented in the 30 ms straddling saccade onset were ranked according to their presentation order, and averages were computed in the four quartiles. Thin lines give the results for the individual subjects and red thick lines the average; for reference, grand-averages for the experts groups are also shown (black horizontal thick lines). Paired t-tests compared the first and last quartiles and results are marked at the top of the panels (n.s. = not significant, *** = p<0.001)

from the experts group. In our second experiment, we focused on our expert observers and manipulated the saccade task. Rather than letting saccades be driven by the sudden appearance of the peripheral target, we had both the fixation and saccade target always visible (‘steady-on’) and an auditory cue instructing the initiation of a saccade.

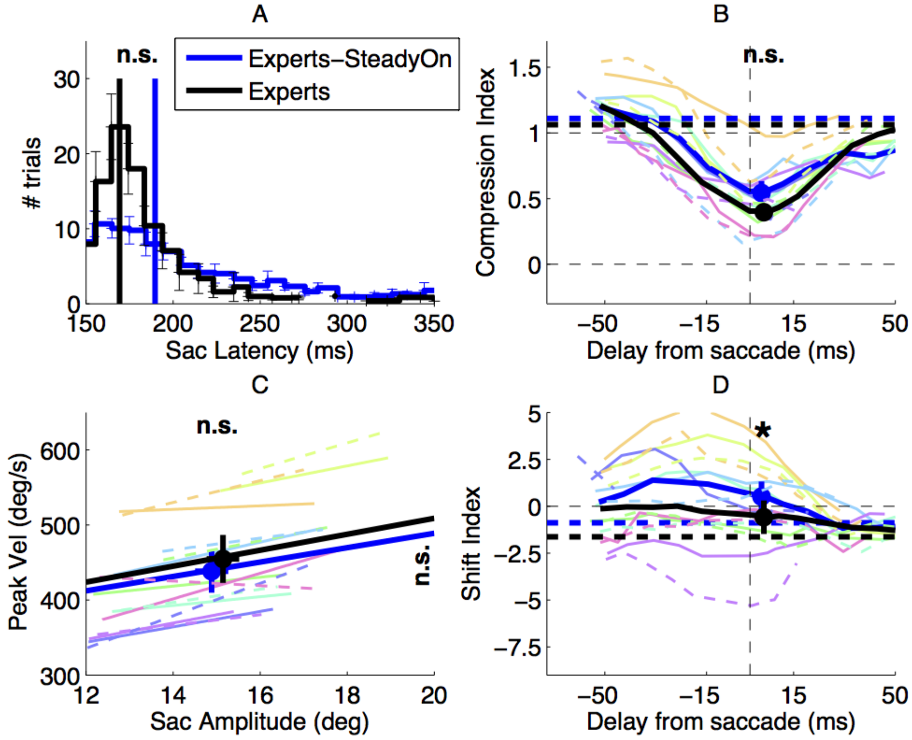

Figure 2.9: **Effect of saccade target presentation mode** Saccade behavior and perisaccadic mislocalization in expert subjects, tested with the abrupt onset vs. the steady-on saccade target paradigm. A & C: saccade latency and saccade main sequence (same conventions as in Figure 2.6 A & C); B & D: compression and shift indices of perisaccadic mislocalization (same conventions as in Figure 2.8 A & C). Paired t-tests compared the averages in the two conditions and the results (n.s. = not significant, * = p<0.05) are marked at the top of the panels A, B and D and at top and right of panel C (for values on the x and y-axis respectively). Thin colored lines in panels B-D show individual subject data: dashed for the abrupt onset, continuous for the steady-on saccade target paradigm.

Consistent with a more volitional nature of the saccade behavior in this condition, we find systematic differences in the saccade latency distribution (Figure 2.9B). There is an increase of trial-by-trial variability of saccade latency (paired t-test on the standard deviation of saccade latency values: t(7): 5.22, p<0.01). Contrary to Experiment 1, this is not accompanied by a significant change of average saccade latency (paired t-test: t(7): 1.56, p=0.163). All other saccade parameters are well matched across conditions. There is no difference of average saccadic amplitude (x-axis of Figure 2.9D, paired t-test on average saccade amplitude: t(7): -1.45, p=0.191) or its variability (paired t-test on the standard deviation of saccade amplitude values: t(7): 1.12, p=0.301) and no difference of average peak velocity (y-axis of Figure 2.9D; paired t-test: t(7): -2.35, p=0.051). The change of saccade latency distribution was not accompanied by changes of perisaccadic compression – the compression index values does not differ significantly between conditions (Figure 2.9A, paired t-test in the 30 ms straddling the saccade onset: t(7): 2.22, p=0.062). However, we observe a tendency for shift index values to be more positive in the steady-on than in the abrupt onset target paradigm (paired t-test: t(7): 3.23, $p<0.05$) – the opposite of the trend observed comparing novices to experts. In our last set of analyses, we pooled data from experts and novices collected with the abrupt-onset paradigm and analyzed the correlations between mislocalization indices and saccade parameters across subjects. Compression and shift indices do not correlate with any of the parameters we analyzed. In particular, there is no significant correlation between peak velocity and compression index values (Pearson's linear correlation coefficient R(22): 0.071, p: 0.741) and no correlation between the shift index (which describes the location of the 'focus of compression' discussed in relation to Figure 2.6) and saccade landing position (R(22): 0.050, p: 0.818).

2.2.4 Discussion

Our two experiments compared the localization of briefly presented perisaccadic stimuli during different types of saccades, associated with different distributions of saccade latencies: spontaneous 'targeting' saccades to a sudden peripheral onset, made with different levels of practice, or saccades instructed by an auditory stimulus. Our main finding is that the index describing perisaccadic compression (Lappe et al., 2000) remains similar in spite of large variations of saccade latency and its variability, suggesting that the level of automaticity with which the saccade task is performed does not influence the perisaccadic distortion of perceived space.

In the first experiment, we examined how perisaccadic compression indices vary as a function of practice on a simple targeting saccade task, in which saccades were repeatedly made to an abrupt peripheral onset occurring at the same location for the entire experiment. Novice observers (with no prior experience in saccade experiments) displayed longer and more variable saccade latencies compared with a pool of expert observers, and their behavior gradually converged towards that of the experts over the course of the experiment. A similar effect of practice on saccade latency has been reported for targeting saccades in non-human primates (Basso & Wurtz, 1998). We find that these large variations of saccade performance, both across subject groups and within novice subjects as function of practice, were not associated to any detectable difference of perisaccadic compression. In the second experiment, we compared performance of experts between the targeting saccade task described above and an atypical saccade task, where the saccade target was continuously visible and an auditory stimulus instructed subjects to initiate the saccade. Removing the sudden onset of a visual stimulus reduces it salience, affecting both behavioral responses (Yantis & Jonides, 1984) and neural responses in several areas – particularly areas tightly connected with eye movement control, such as LIP and FEF (Bruce & Goldberg, 1985; Gottlieb et al., 1998). In this task, we find that saccades have more variable latencies; this is consistent with stronger volitional control (associated with more variable behavior, Carpenter, 1999). However, contrary to what is typically found for voluntary saccades, the average saccade latency does not increase. This may be due to the relative simplicity of the task, which does not impose the time requirements of processing complex instructions and/or the meaning of symbolic cues that are often used for guiding voluntary saccades (Walker et al., 2000). Whichever the origin of the increased variability of saccade behavior, we found that this was not accompanied by a change of perisaccadic compression – which was indistinguishable from that observed for targeting saccades.

Our findings are in line with previous work that manipulated the degree of automaticity of saccade behavior by varying the predictability of the saccade target location, and observed minor or no changes of perisaccadic compression (Maij, Brenner, & Smeets, 2011). Our results also agree with work comparing perisaccadic mislocalization during saccades with different levels of volitional control: ‘pro-saccades’ targeting a peripheral onset and ‘anti-saccades’ made in the direction opposite to the peripheral onset (Awater & Lappe, 2004). Although many saccade parameters distinguished the two types of eye movements, the patterns of perisaccadic mislocalization were markedly similar. Ostendorf et al. (2007) specifically investigated the relationship between the main saccade parameters and perisaccadic compression; while we both find that perisaccadic compression is unrelated to the distribution of saccade latency values, Ostendorf et al. (2007) found a significant correlation between perisaccadic compression and peak saccadic velocity that was not present in our dataset – nor in Maij et al.'s (2011).

Our second experiment may also be considered in the light of a recent study (Zimmermann, Morrone, et al., 2014) that manipulated the features of the saccade target and observed the near elimination of perisaccadic compression when all visual transients associated with the target appearance were eliminated – by removing the saccade target stimulus altogether and instructing subjects to saccade to an unmarked location. One of the hypotheses put forward to explain this finding is that compression is affected by the shift of attention triggered by the sudden appearance of the saccade target. However, Zimmermann et al.'s (2014) manipulation affected an additional factor that is known to affect localization: spatial references (Lappe et al., 2000), clearly reduced when the saccade target presentation is withheld. In this sense, our second experiment may be seen as a logical counterpart of Zimmermann et al.'s (2014): we removed visual transients associated with the saccade target, while maintaining its role as a strong and stable visual reference. We did not find a dramatic change of compression as in Zimmermann et al. (2014). This supports an alternative hypothesis proposed by these authors, that compression is influenced by the presence of the saccade target rather than its salient sudden appearance and the shift of attention that the latter would trigger. This is consistent with much evidence that visual factors, and specifically spatial references, play a major role in space perception at the time of eye movements (e.g. Deubel, 2004) and specifically in shaping perisaccadic mislocalization effects – as shown in a variety of paradigms that analyzed real saccades (Cicchini, Binda, Burr, & Morrone, 2013; Zimmermann, Morrone, et al., 2014), simulated saccades (Ostendorf, Fischer, Gaymard, & Ploner, 2006; Zimmermann, Born, Fink, & Cavanagh, 2014) and interrupted saccades (Atsma, Maij, Corneil, & Medendorp, 2014).

Like the compression index, we find that the shift index – the other index describing perisaccadic mislocalization – is fairly well matched across conditions and experiments. However, small significant differences did emerge, which could indicate that the two indices changed independently in the contexts we examined. Two observations call for caution in interpreting these. First, note that the majority of our data is concentrated in the interval immediately following saccade onset; while at the trough of the compression index, this interval is not optimal for testing differences in the shift index, which typically peaks before the saccade onset (Lappe et al., 2000). Also note that the shift index varied markedly across observers (both for novices and experts), yet this idiosyncrasy appeared to be a consistent trait of the individual subjects, which did not normalize with practice and did not correlate with any of the saccade parameters.

These results are reassuring of the face validity of the saccadic compression phenomenon. To our knowledge, the present study is the first to follow perisaccadic mislocalization from the very first exposure until subjects master the task. Finding virtually no change indicates that the phenomenon does not emerge from strategies that subjects develop in this laboratory setting. This is important for two reasons. First, the experimental conditions in which compression is measured are highly unnatural – we normally don't make saccades of constant direction and amplitude repeatedly over several minutes, and the neural substrates of saccade programming depend on the degree of automaticity of the saccade task: recruiting different areas (Johnston & Everling, 2008; McDowell et al., 2008) with a different temporal profile (Basso & Wurtz, 1997, 1998). Second, given hundreds of repetitions, subjects may develop strategies and adjust responses to cope with an unnatural context – such as the unnatural flashing of stimuli in and out of view within few milliseconds. An example of this comes from the saccade adaptation literature. Saccade adaptation is the change of saccade amplitude obtained by repeatedly displacing the target of the saccade while the latter is in-flight. Given hundreds of repetitions, this manipulation affects not only oculomotor behavior (shortening/lengthening of the saccades) but also leads to a global distortion of visual space (Zimmermann & Lappe, 2009, 2010, 2011; Schnier & Lappe, 2012) – as though the visual system had adjusted to the artificial mismatch between pre- and post-saccadic target location, incorporating it within a new and distorted spatial metrics. Our observations indicate that, on the contrary, the compressed spatial metrics representing perisaccadic flashed stimuli does not emerge as an adjustment to specific stimulus and task conditions – being present in observers that are completely new to such conditions.

In conclusion, we show that the spatial compression observed for perisaccadic flashed stimuli is a robust phenomenon, insensitive to the specific paradigm used to drive saccades and to the level of practice with the saccade task.

Chapter 3

Time perception

3.1 Introduction

Although there have been major advances over the last few decades in our understanding of how the human brain processes many features of the sensory world, and especially how the mechanisms elaborating spatial information work, the perception of time still remain one of the less known products of neural computation.

The nature of time itself represents a long debated problem, with concepts from philosophy, psychology and physics intertwined. Everyone knows what time is, but giving a precise definition is something that occupied millennia of philosophical speculation. Time is a pervasive feature of our physical and sensory world and taking track of time is a crucial factor for several human abilities and for the society itself. It is so pervasive that its role in our lives is highlighted across scales spanning over 15 orders of magnitudes (Buonomano, 2007). Noteworthy, from the extremely short timescale of nanoseconds recorded by atomic clocks to the more familiar ranges of seconds, minutes, hours and so on, humans exploits the same technology to record time, with the precise atomic clocks taking the time for a huge range of tasks, from the synchronization of signal from different satellites to even adjustments to the length of the year in our calendar.

But what about our inner experience of time? Despite the fact that a specific sensory system devoted to the sense of time does not exist, animals in general possesses an intrinsic ability to keep track of the passage of time.

This ability is crucial for survival in the environment, where a precise timing for actions is mandatory in a large range of situations, and for functions like perception and – in the case of the more evolved animals – cognition (Ivry & Spencer, 2004; Buhusi & Meck, 2005). Similarly to our technological devices, animals can also exploit temporal information across a huge range of scales, from microseconds to milliseconds for the finer processes toward the daily cycles of food intake, sleep and circadian regulation – obviously without the extreme precision of an atomic clock, but with enough sensibility to face rapid changes in the environment and anticipate future events.

3.1.1 Temporal information across different timescales

The smallest scale of temporal information that biological systems can exploit is the microsecond range. This timescale is particularly important to determine the difference in time for which a given auditory signal reaches the two ears – a process crucial for fine localization of sound sources. Such processes could seem very difficult in a system (i.e. the brain) where signals (i.e. neuronal action potentials) last hundreds of milliseconds, but it could be implemented by biological systems by means of "coincidence detection", and exploiting the time necessary for propagating a signal along the axons as a physical delay line. Indeed, some neurons are particularly capable of detecting simultaneous signals or small differences between them. Such a sensibility allows for the detection of very small differences in the time of arrival of auditory signals at the two ears, with a sensibility that in extreme cases (like the barn owl, for which auditory localization is fundamental for hunting small rodents) allow the discrimination of interval differences in the order of 10 µs.

Beside timing in the very fast timescale of microseconds, useful for auditory processes like sound localization, time in the range from milliseconds to seconds and even to minutes appears to be most important for functional behavior in the environment, and underpinned by the most sophisticated processes. A great number of studies have pointed out that animals, particularly humans, have evolved complex and precise mechanisms for timing in this range, especially considering that this scale of temporal information comprises the complex sounds that many species exploit for communication, and that humans can exploit a pure temporal code for communication (i.e. the Morse code). Human languages, for instance, have many temporal properties that are necessary for comprehension of speech, such the ability to correctly segment a stream of words, fundamental for the recognition of prosody and overall meaning. This timescale is also particularly important for behavior, since reactions to relevant stimuli, action planning and event anticipation in this duration range comprise a multitude of processes fundamental for survival (from avoiding a predator or choosing the right timing to catch a prey, to start the car when the traffic light turns green). Our inner experience of time, as well as most of the sensory and sensory-motor processes underlying our conscious experience of the world and the interaction with it, falls into this timescale. So, for the aim of this work, an in-depth discussion of the theories and models proposed to account for time perception will be focused on time perception in the millisecond-to-second range (see next section).

Beyond the millisecond and second time ranges, and even beyond minutes, animal track the passage of time by means of the circadian cycles. This timescale is particularly useful to anticipate changes in the environment, such cyclical changes in light, temperature or the behavior of other animals (i.e. predators), and for regulating food intake and sleep cycles. Given the intrinsically slow nature of timing in this daily scale, the mechanisms involved here are very different from timing in the shorter ranges, and so not likely governed by the fast neuronal dynamics or by specific features of neurons. More likely mechanisms underlying the circadian regulation rely on protein production and long-lasting changes in their concentration inside cells. Indeed, circadian rhythms are not a unique feature of animals, but also vegetal and single-cell organisms can regulate their behavior according to daily cycles.

Reviewing the literature on time perception requires particular attention to the time range, since different timescales are subserved by markedly different mechanism – for instance in the case of microsecond, millisecond and slower cycles. Classically, the field of time perception has been more concerned with temporal durations comprised between 100 ms and a few seconds, given that, as mentioned above, most of the sensory and sensory-motor processes are based with such range of durations. Inside this range (especially from 0.1 to 1 s) many interesting properties have been identified. First, even if it is a debated issue, there is the notion of “indifference interval”: i.e. an interval, approximately around 700 ms, for which there are no tendencies toward under- or over-estimation, which are usually observed in other ranges (Eisler et al., 2008). Second, the highest sensitivity for rhythmical stimuli is around 300-800 ms (Drake & Botte, 1993; Friberg & Sundberg, 1995). Finally, the typical preferred tapping rate in a tapping task in usually from 350 ms to 700 ms (different for children, adults and older people). Some authors also suggested a dissociation between interval shorter and longer than 1 s. For instance, Rammsayer et al., (2008) showed different pharmacological effects on different intervals, and many studies pointed out that while shorter intervals are mostly based on sensory processing (they benefit from some automatic processing), the processing of intervals longer than one second involves the support of cognitive resources (Rammsayer & Lima, 1991; Lewis & Miall, 2003b; Hellström & Rammsayer, 2004). However, such turning point between mechanisms at 1 s is somehow arbitrary, and could be variable (Grondin, 2010). Nevertheless, some recent studies also pointed out that different mechanisms – that otherwise would be considered mutually exclusive – could underlie the processing of temporal information in the sub-second and supra-second ranges (Mauk & Buonomano, 2004), and so different processes could actually coexist, working on different durations. This issue will be of particular interest for the debate between different theories about temporal mechanisms in the millisecond-to-second range, and hence will be discussed more precisely in the next section.

3.1.2 Theories and models of time perception

Across many decades of theoretical and experimental work, several models of duration perception have been proposed. However, no consensus exists on how and where in the brain temporal information is processed, and so the different models of time perception are highly debated. Different frameworks trying to address how the brain keep track of time exploited different metaphors and tried to localize the hypothetical temporal mechanism in several brain areas that seemed suitable to underpin temporal information processing, but broadly speaking the models fall into two classes. On the one hand, we have the “dedicated models”, which assume a specific mechanism, exclusively dedicated for measuring time passage, such as the classic clock-counter model and its several variants, and more recent models exploiting oscillatory processes such the beat-frequency model; on the other hand, we have the "intrinsic models", such as the State-Dependent Network (SDN) model, where the information about event duration is derived from the intrinsic activity of neural populations responding to non-temporal features of the stimulus, thus intrinsically acting as interval timers – i.e. other sensory or cognitive processes that are not specifically dedicated to time processing act additionally as interval timers. In the second class of models, time processing would be ubiquitous throughout the brain, rather than being a specialized mechanism. Another difference between these two class of models, is that while the dedicated models usually hypothesize a mechanism localized somewhere in the brain (although with some exception, where a clock-like mechanism is represented by a network of brain areas distributed across the cortex; Harrington & Haaland, 1999; Lewis & Miall, 2003; Ivry & Schlerf, 2008), intrinsic models usually concern a distributed mechanism, not localized in a particular area or brain region.

Besides the most classical Clock-counter and the most recent SDN model, a large number of variants and competing frameworks have been proposed across many years, each one yielding both some successful predictions that have been observed experimentally and evidence that seems not supporting it, increasing the complexity of the topic.

"Dedicated" models of time perception

The class of dedicated models of time perception is probably the most conspicuous, and represents the "classical" view on how the brain might process temporal information. These models very often exploit the metaphor of a clock or a stopwatch, starting with a given signal and stopping when the signal ends, generating a sensory representation of the duration of each given event that can be compared with a reference trace stored in memory, which give rise to our psychological experience of duration. One of the most prominent and studied model of human time perception in this class – but probably the most prominent for time perception in general – is the "Clock-Counter" (or "Pacemaker-Accumulator") model (Creelman, 1962; Treisman, 1963). This model, which proposes a modular system with different stages, is the best example of the

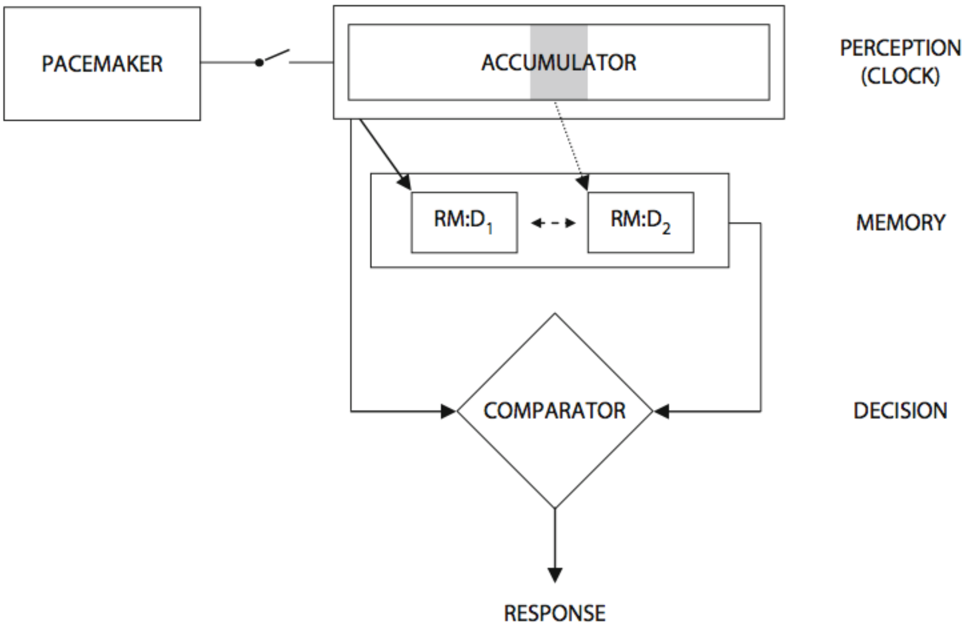

Figure 3.1: **Pacemaker-Accumulator model.** Illustration of the information-processing version of the Scalar Expectancy Theory. At the perception (or clock) stage, there is a pacemaker-accumulator device, with a switch between the two components. The arrows indicate that once a given level of accumulation is reached, it is with another representation in reference memory that the interval for a given trial will be compared. The accumulator level that has to be reached may vary from trial to trial (gray zones in the accumulator). The bidirectional arrow at the memory stage indicates that different reference memories may exert mutual influences. Adapted from Grondin, 2010.

stopwatch metaphor. The first stage in this information-processing model is the sensory or clock stage, structured as a pacemaker-accumulator mechanism with a switch device between them: when the switch is closed, the pulses emitted by the pacemaker are collected by the accumulator; once the stimulus ends, the switch opens, and the number of pulses in the accumulator are compared to a previously reinforced memory trace of a reference accumulator reading at the comparator level. Between the clock and comparator levels, there is a memory stage, which, on the one hand, can collect a given duration provided by the accumulator if such duration is somehow reinforced (i.e. reward for keeping a specific time), and on the other hand provides the reference trace for the comparator stage (Figure 3.1).

One of the major goals of this framework, in its contemporary versions, is to take into account and implement the scalar property (Gibbon, 1977), which describes the way in which the mean and standard deviation of the response distribution (usually in a duration reproduction task) covary. Such scalar property closely resembles the Weber's law (Weber, 1851 – see Buhusi & Meck, 2005) – a milestone law that is obeyed by most sensory dimensions. Indeed, while the Clock-Counter metaphor forms the ground onto many different models have been developed, the most cited version of the Pacemaker-Accumulator accounts is probably the Scalar Expectancy Theory (SET; Gibbon, 1977; 1991; 1992; Wearden, 2003), which is founded on the critical assumption of the scalar property. According to such property, participants' responses in a temporal task follow a normal distribution around the criterion (or reference) duration, with the width of the response distribution proportional to the criterion duration. In other words, the variability of the judgments increases linearly with the mean of the durations to be represented. In the classic Pacemaker-Accumulator models, the scalar property is represented by the assumption that the timing error is proportional to the criterion duration: the longer the duration, the bigger the expected error in the clock processes. Interestingly, this property holds not only for behavioral responses, but also for measures of neural activation such ensemble recording or hemodynamic responses measured with fMRI (Meck & Malapani, 2004; Hinton & Meck, 2004).

However, different models attribute the source of timing errors to different stages. The most often reported source of such errors is the pacemaker level, which pulse mode could be either represented by a deterministic distribution, like in Treisman's (1963) model, where the frequency of clock pulses was modulated by the arousal, or by a stochastic (Poisson) process, like in Creelman's (1962) model. Less commonly, the timing errors are attributed to a counter (or accumulator) failure (Killeen & Fetterman, 1988). Another stage that has been linked to timing error is the switch stage, also related to marking errors. Such stage is thought to be controlled by attention (Meck, 1984), and hence attentional resources allocated to the timing task determine the precision of the collected pulses: when the attention is fully allocated to time, the accumulation process is at its maximum and the number of timing errors drop significantly, while perceived duration is lengthened. The switch level is also associated with marking errors – i.e. latencies in the closing and opening of the switch in response to the onset and offset of a stimulus, which create a discrepancy between the physical onset/offset and the internal representation of them. While such latencies might explain some peculiar features of duration discrimination performance (Grondin, 1993; 2003; 2010), some authors attribute the source of the entire variability in timing processes to the discrepancy between physical and internal onset/offset (Allan et al., 1971). Finally, another source of variability in timing task that must be taken into account is represented by the memory and the decisional stage, especially in tasks involving the comparison of learned reference intervals with probe durations (Gibbon et al., 1984).

The success of the Pacemaker-Accumulator model arises from the fact that large amounts of its prediction have been experimentally tested (Matell & Meck, 2000; Buhusi & Meck, 2005). One of the most studied predictions of this model is the aforementioned scalar property, but many other predictions, such as the dependence of the experience of time on some non-temporal factors like arousal or the levels of some neurotransmitters have been successfully tested experimentally. Perhaps one of the first experimental investigations about the neural substrates of the pacemaker used a pharmacological approach, providing considerable support for a dissociation between the different stages of the pacemaker-accumulator mechanism. Particularly, it has been shown that in both humans (Meck, 1983; Maricq & Church, 1983; Meck, 1986; Matell et al., 2004) and animals (Rammsayer, 1993) dopaminergic alterations affect selectively the clock stage of the model, in terms of a deceleration of subjective clock speed in proportion to their affinity with dopamine D2 receptors, while cholinergic drugs selectively affect the memory stage, with cholinergic activity in the frontal cortex proportional to the absolute error of a temporal memory translation constant (Meck, 2002) – a parameter of the scalar property responsible for the scalar transformation of inputs generated by a sensory clock and stored in memory.

However, the connection between dopamine and the speed of the internal clock has been challenged by several observations. For instance, patients with Parkinson's disease show a peculiar behavior in temporal tasks: when asked to remember two durations and reproduce them, the responses tend to migrate toward each other (the "migration" effect; Malapani, 1998) when tested off the dopaminergic medication – a pattern of results that does not support the idea of a direct and specific dopaminergic modulation of the speed of the internal clock. Moreover, other evidence that does not support this connection regard the inconsistency between the small behavioral effect of dopaminergic drugs (Matell et al., 2004) and the level of dopamine released in the striatum observed in vivo (Holson et al., 1996).

While the Pacemaker-Accumulator model suffers from the fact that its neurobiological underpinnings are not so well defined, many other dedicated models have been proposed, basically resolving some of the original problems but unfortunately introducing other issues (Matell & Meck, 2000).

Another model that exploits a pacemaker-accumulator device is the Behavioral Theory of Timing (Killeen & Fetterman, 1988), according to which a given signal starts a pacemaker that drives the animal into different "behavioral" states according to the number of pulses elapsed. Interestingly, also this framework is consistent with the scalar property, and the successive states of activity in the striatum have been proposed as a possible neural underpinning of this model, but the reinforcement-based modulation of the pacemaker rate hypothesized to be responsible for the scalar property does not hold up further experimental findings (Bizo & White, 1994a; Bizo & White, 1994b; Bizo & White, 1997; Matell & Meck, 2000).

Beside the Clock/Stopwatch metaphor best represented by the Pacemaker-Accumulator models, other dedicated models exploited different approaches to address the mechanisms governing our experience of time. Two remarkable different types of "clocks" are represented by the idea of a *process-decay* mechanism, and by an oscillator mechanism working by means of *coincidence-detection*.

Regarding process-decay mechanisms, the first example is the Multiple Time Scales model (Standdon & Higa, 1999). According to this model, the role of clock is subserved by the decay of memory strength, with specific patterns of decay representing different durations. Another example of process-decay is the Spectral Timing Model (Grossberg & Schmajuk, 1989), where the assumption is that different rates of neuronal activation lead to differential habituation rates. In turn, the combination of different habituation rates generate different pattern of activity in response to different durations. Both these models achieve the scalar property as the Pacemaker-Accumulator models, but their possible neurobiological substrates are not clear, and both fail to explain some perceptual phenomenon such the gap or the reset effect (Matell & Meck, 2000).

The second alternative to the Pacemaker-Accumulator device is the oscillator/ coincidence-detection mechanism. One model implementing an oscillator mechanism that takes advantage of the physical regularities in the temporal structures of the environment is the Dynamic Attending Theory (DAT; Jones & Boltz, 1989). Such model is based on two main processes: a non-linear oscillator and an attentional pulse rhythm process. Combining these two processes give rise to an attentional rhythm, with the period and phase adapted to the temporal structure of the external stimuli. Moreover, such oscillator process is highly adaptable, trying to get closer to the rhythmicity of the environmental stimuli, synchronizing the attentional pulses to the onset of the stimuli (Barnes & Jones, 2000; McAuley & Jones, 2003). This model, which emphasizes the role of expectation and attentional modulation, can account for a large number of phenomena occurring when regularities are available in the environment. Another interpretation of the oscillator mechanism for time perception is the Multiple Oscillator Model (Church & Broadbent, 1990; 1991). Such model hypothesizes the existence of a large number of oscillators with different periods, with time coded as the combination of half-phase readouts across all the oscillators: specific combinations of periodic neural events would represent specific durations. One limit of this model is represented by the fact that longer durations must be coded by neural events with a correspondingly longer period (up to 204 s), while experimental data pointed out oscillator processes only up to 60 s in the basal ganglia (Ruskin et al., 1999).

A theory involving oscillating processes worth particular attention is the Beat Frequency model of timing (Miall, 1989) and its recently updated variant Striatal Beat Frequency (SBF) model (see Figure 3.2; Matell & Meck, 2000; Matell & Meck, 2004; Matell & Meck, 2005). In this latter instantiation, the striatum would have a central role as coincidence detector. Coincidence detection represent an integration rule according to which neurons are induced to fire only when receiving a large number of simultaneous inputs in a small temporal window (5 – 20 ms). A reinforcement of particular synapses can take place by means of Long Term Potentiation or Long Term Depression (LTP and

LTD, respectively) (Wickens et al., 1996; Charpier & Deniau, 1997) induced by the concurrent stimulation by large arrangement of cortical input and simultaneous dopaminergic burst from the substantia nigra pars compacta (SNPC). More precisely, the striatum activity from cortical input coupled with depolarization of the striatal units usually causes LTD. However, when such depolarization is further coupled with dopamine release from the SNPC, synapses are potentiated by means of LTP. LTP and LTD can thus represent a biological substrate for learning, changing the synaptic weights according to the pattern of activity and creating specific patterns of activity necessary to successfully activate the striatal spiny neurons (Arbuthnott et al., 2000; Shultz et al., 1997). Exploiting such potentiation/depression mechanism, striatal neurons would be particularly suitable as coincidence detectors, firing only when the previously reinforced cortical inputs are active at the same time. So, striatal spiny neurons would be the optimal substrate to support the Beat Frequency model. According to such model, at the onset of a stimulus a burst of dopaminergic activity starts a synchronized oscillating activity with a variety of periods across the cortex. Furthermore, the coincident firing of such cortical oscillator would be detected by striatal spiny neurons previously trained by means of the LTP/LTD process. The output of these units would be then integrated by the basal ganglia output nuclei (globus pallidus, subthalamic nucleus, entopeduncular nucleus, substantia nigra pars reticulata) and sent to the thalamus for the expression of the behavioral outcome. Moreover, the flexibility of this model is further increased by the assumption that the thalamic activity itself can dynamically modulate the cortical and striatal activity through multiple closed and open loops.

Has such a pattern of striatal activity been observed experimentally? The assumptions of SBF are supported by investigations concerning the coding of predictive signals by dopaminergic neurons (Fiorillo et al., 2003). Indeed, under conditions in which there is a certain degree of uncertainty about whether or not a reward would be delivered, dopaminergic neurons show a peculiar pattern of activity that closely resembles the processes hypothesized by the SBF: a burst of activity at the trial onset, another burst after a duration close to the expected time of the reward, with sustained activity throughout the interval between the two burst. This pattern of activity is actually very similar to

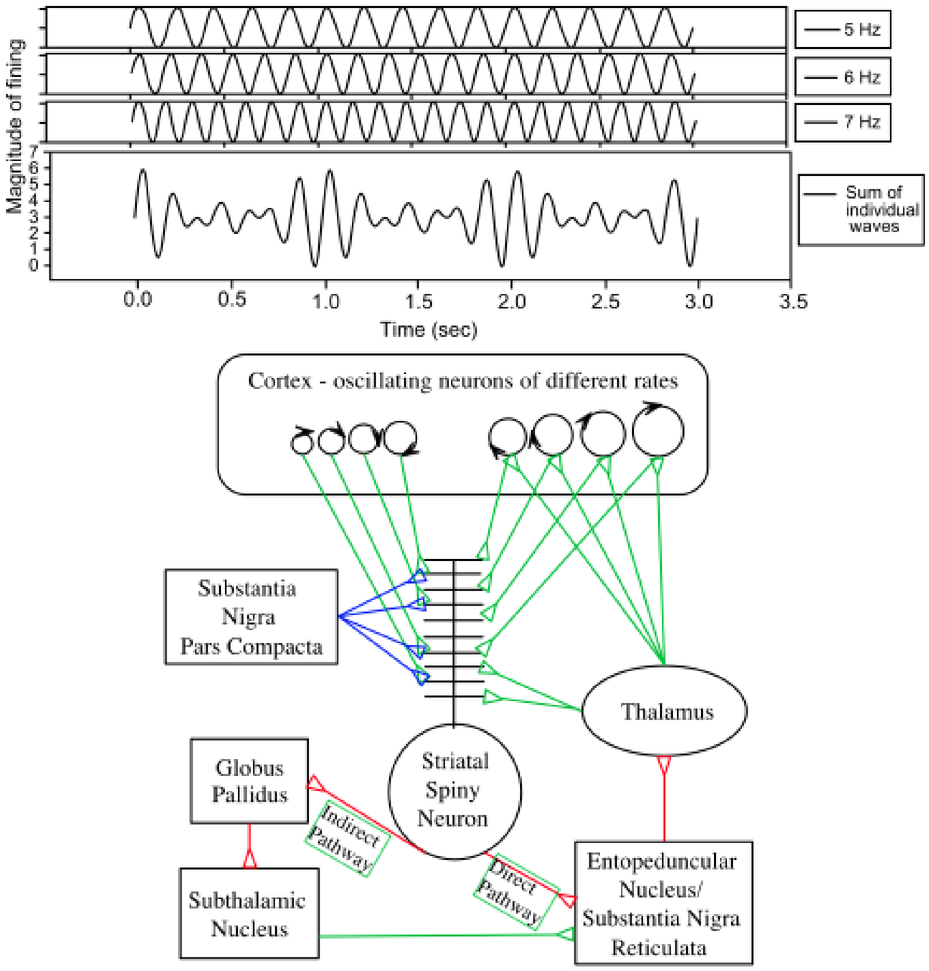

Figure 3.2: **Graphic representation of the machinery of the Striatal Beat Frequency model.** (top) Three different oscillating patterns generated by cortical ensembles, and the unweighted summation of them over time – which represents the input to the striatum. (bottom) Information flow through the cortico-striatal-thalamo-cortical loop. Arrows indicate the direction of the information flow. ed arrows are GABAergic, green arrows are glutamatergic, blue arrows are dopaminergic. The dopaminergic input to the striatal neuron provides a reinforcement signal to selectively weight cortical and thalamic inputs. Different striatal neurons participate in direct and indirect pathways. Adapted from Matell & Meck, 2004.

the mechanisms of coincidence detection and reinforcement described above: the first burst would synchronize the oscillator, the sustained activity would represent attentional processes, and the burst at the expected reward time could represent the updating of cortico-striatal connections and transmission

(Shultz, 2000).

Intriguingly, it has been recently pointed out that such coincidence-detection mechanism might be actually another instantiation of the more canonic spike counting strategy (i.e. counting the amount of activity elicited by a given stimulus), like the other side of the coin. Indeed, some mathematical investigations concerning the nature of these two neural strategies indicated that both coincidence-detection and spike counting could be used to solve the problem of comparing different spike patterns, according to the precision needed and the acceptable error in the specific task (van Rossum, 2001). Namely, in tasks where the precision needs to be relatively high, and hence the maximum acceptable error is small (a situation that well describe a task concerning a continuous variable, like time), a more suitable solution would be to exploit coincidence detection. On the other hand, when the precision need not be so high (like in the case where neuronal spike patterns represent a discrete variable, such numerosity), a sufficient solution may be to simply count the two spike patterns representing the different stimuli to be discriminated. Such a continuum between the extremes of coincidence-detection and spike counting, suggested by these results, supports the idea of a common neural system processing at least temporal and numerical magnitudes (Meck & Church, 1983; Brannon & Roitman, 2003; Walsh, 2003; Feigenson et al., 2004) (see Chapter 5 for further information about evidence of a common mechanism for time and numbers).

“Intrinsic” models of time perception

Across many years of research, the idea of a centralized, dedicated clock supporting our experience of time has proven to be the most successful framework, also supported by the fact that time is something that transcend all the sensory modalities. However, some recent models tried to account for time perception without hypothesizing a centralized and amodal clock, relying instead on more biologically defined circuit and neuronal properties (Matell & Meck, 2004; Meck et al., 2008; Ivry & Schlerf, 2008). Such class of theories, named “intrinsic” models offers a radically different point of view on temporal mechanisms, assuming that there are no specialized mechanisms underlying the perception of time. Instead, temporal information would be represented intrinsically in the dynamics of neuronal network dedicated to other functions.

However, some of the theories in this class proposed temporal mechanisms that might be limited to brain areas with particular properties, namely the ability to maintain a sustained activity without direct sensory stimulation (Brody et al., 2003; Reutimann et al., 2004). For instance, according to Brody and colleagues (2003), the time-dependent activity of prefrontal neurons in the delay period between two stimuli might encode time itself, thus localizing the substrate of time processing in the prefrontal cortex, without attributing a specialized role for time to such brain region. Another instantiation of such an intrinsic mechanism for time perception is the possibility that both the time elapsed from a stimulus and the expected time of a given action that has to be done (in delayed response tasks) might be encoded by the ramped activity of neuronal ensemble, which might provide a working memory representation of the stimulus and the time until the response (Lebedev et al., 2008).

On the other hand, time processing might be ubiquitous across the brain, arising as an intrinsic property of the dynamics of modality-specific networks (Buonomano, 2000). One possibility is that the information describing the duration of a stimulus or an event in any given sensory modality might be decoded by the magnitude of neural activity elicited by such stimulus (Pariyadath & Eagleman, 2007). Such idea is supported by several findings about repetition suppression, which showed that presenting the same stimulus many times leads in parallel to a decrease in neural activity associated with such stimulus and to a contraction in subjective duration (Pariyadath & Eagleman, 2007). For instance, this pattern of suppression affecting repeated stimuli has also been found when a set of digits is presented in their standard order (1, 2, 3, 4, 5...), which leads to the perception of a longer duration for the digit “1”, and a decreasing subjective duration for the subsequent digits. However, when the digits are presented scrambled (3, 1, 5, 2, 4...), each digit is perceived as equally longer in time – which can be easily explained by the fact that with the scrambled order each digit cannot be anticipated, and thus receive a similar amount of neural processing. So, according to Pariyadath and Eagleman’s (2007) idea, the process leading to the decrease observed in neural activity (measured with fMRI) found by Henson & Rugg (2003), is the same that leads to the “compression” of duration for the repeated or predictable stimulus.

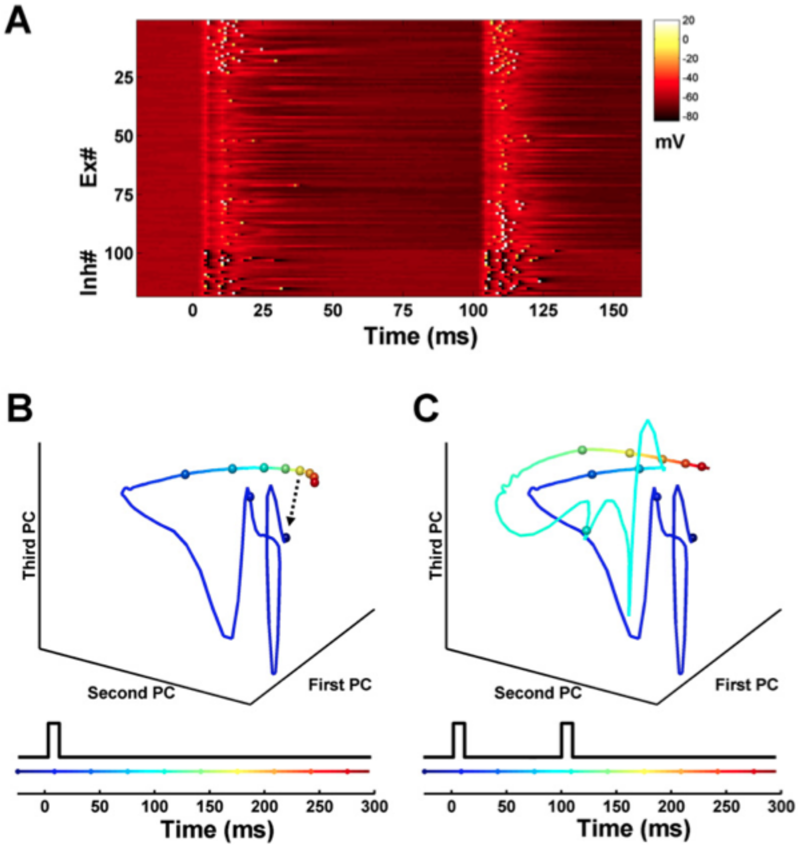

Figure 3.3: **State-Dependent Network Simulation.** (A) The panel represents a voltage plot of a subset of neurons in the simulated network. Each line represents the voltage of a single neuron in response to two identical events separated by an inter-stimulus interval of 100 ms. The first 100 lines represent 100 excitatory ("Ex") units (out of a total of 400 in the original simulation), and the remaining lines represent 25 inhibitory ("Inh") units (out of 100 in the simulation). Each input produced a depolarization across all neurons in the network, followed by subsequent inhibition. While most units exhibited subthreshold activity, some units spiked (represented as white pixels) to both inputs, or exclusively to the first or second. The excitatory units are sorted according to their probability of firing to the first (top) or second (bottom) pulse. This selectivity to the first or second event was due to the difference between network states at different point in time (t = 0 and t = 100 ms). (B) Processing trajectory related to the three principal components of the network in response to a single pulse. As showed, there is an abrupt and rapidly evolving response beginning at t = 0, followed by a slower trajectory. The fast response is due to the simultaneous depolarization of a large number of units, while the slower change reflects the short-term synaptic dynamics and slow IPSPs. The speed of the evolution of the trajectory represented in state-space can be visualized by the rate of change of the color code and by the distance between the 25 ms marker spheres. Because synaptic properties cannot be rapidly "reset," the network cannot return to its initial state (arrow) before the arrival of a second event. (C) Trajectory in response to a 100 ms interval. Note that the same stimulus produces a different fast response to the second event, departing from the current trajectory. Adapted from Karmarkar & Buonomano, 2007.

Perhaps one of the most representative instantiations of the intrinsic models is the State-Dependent Network (SDN) model (Buonomano, 2007; Karmarkar & Buonomano, 2007). The general idea behind this framework is that temporal information would be encoded by the dynamic variation of time-dependent neuronal and network properties, such cellular and overall activation state – i.e. duration processing might arise from the dynamically changing activity of neural networks processing other attributes of the stimuli, and each given neural network might be intrinsically capable of time processing. During the processing of a sensory stimulus – i.e. a sound – cellular properties and activation states across a population of neurons continuously change, and the change itself might contain the information indicating how much time is elapsed since the start of the activity. Considering a neuronal network containing N neurons, and coding the possible activation states of each neuron as either a 0 (the neuron is not active) or 1 (the neuron is active), the overall activation state of the network at any given point in time (t) can be represented as a N-dimensional vector (where N represents the number of neurons in the network). Such a vector defines the coordinate of a point in a N-dimensional space (with as many dimensions as the number of neurons in the network), representing a particular state of the network at the time point t. When the network is working, processing a stimulus, the activity of each neuron would be rapidly modulated over time, causing the N-dimensional vector of the states to change. By drawing each point described by the changing vector at any time point t+1, these changes can be represented as a trajectory in the N-dimensional space (Fig. 3.3B and 3.3C show a depiction of such trajectory in a three-dimensional space, adapted from Karmarkar & Buonomano's (2007) simulation). According to Buonomano's framework, monitoring this "neural trajectory" should allow for the extraction of temporal information, particularly from the difference between the starting and the final points. Such neural trajectory driven by time-varying changes can be modulated not only by the stimulus properties, but also by internally generated activity carried by recurrent connections. These internal source of modulation offer an interesting property: even if a stimulus is not modulated in time, the neural trajectory would continue to change (Broome et al., 2006; Rabinovich et al., 2008), contributing to the processing of the stimulus intensity or its duration in time. Moreover, while different stimuli elicit different spatiotemporal pattern of activity (Buonomano et al., 1997; Broome et al., 2006; Churchland et al., 2007), for a given stimulus, the dynamic of its neural trajectory should evolve over time in a reproducible manner – another property that could allow not only for the extraction of information about the stimulus itself, but also about the temporal dimension (for instance, the time elapsed since the onset of the stimulus).

However, considering only the internal source of activity and the external input leading to changes in the neural trajectory is too simplistic. Neurons have a large array of properties that vary over time during the activity, and that could modulate the activity itself – a dimension identified by Buonomano & Maass (2009) as the network's internal "hidden" state (as such properties are not accessible to downstream neurons), that must be taken into account. One of the strongest source of modulation that can shape the way the neurons respond to external inputs and influence each other in a timescale of few hundreds milliseconds, is short-term synaptic plasticity, that takes place as depression or potentiation of synaptic connections (Zucker, 1989; Markram et al., 1998; Reyes & Sakmann, 1999; Dobrunz & Stevens, 1999; Zucker & Reger, 2002). Given that such short-term depression/facilitation is often as stronger as the long-term synaptic plasticity (Marder & Buonomano, 2003), which provides a mechanism for the formation of long-lasting memory traces, the short-term modulation might provide a trace of the recent history of stimuli processed by the network. So, over a short timescale, after the activity elicited by a stimulus, the network would be in a slightly different initial state if the same stimulus is repeated after a few hundreds milliseconds, and this possibility would also provide a potential mechanism to extract information about the inter-stimulus interval (the time between the stimuli), given the short-term nature (i.e. decaying over time) of this form of synaptic plasticity. Beside short-term plasticity, a large number of other time- and activity-dependent variables can contribute to the modulation of a network's state, shaping the output of neurons. For instance, some other properties that might contribute to the hidden stare are: NMDA (N-methyl-D-spartate) and ion channel kinetics (Lester et al., 1990; Johnston & Wu, 1995), Ca2+ concentration dynamics (Berridge et al., 2003; Burnashev & Rozov, 2005), inhibitory post-synaptic potentials (IPSPs; Newberry & Nicoll, 1984; Buonomano & Merzenich, 1998), and metabotropic glutamate currents (Batchelor et al., 1994).

How are such neural trajectories decoded in order to extract useful information about the properties of a stimulus? Considering a set of downstream, or "read-out" neurons, each of them would receive signals from large ensembles of input neurons, thus receiving a high-dimensional sample of the activity of the upstream network. Thinking of read-out neurons as linear discriminators, each of them computes a sum of the inputs weighted with their synaptic weights (i.e. the strength of the synaptic connections). Such a weighted sum performed by linear read-out neurons allows them to detect and process the activity associated with a specific external stimulation. Discriminating different external stimuli – which involves the separation of different neural trajectories according to the external stimuli that caused them – with a given degree of sensibility would be possible thanks to the high dimensionality of the input to read-out neurons. Particularly, the nonlinear (where the nonlinearity arise from the intrinsic complexity of the interactions between the internal state and the external stimulation) projection of stimuli into a higher-dimensional state space makes linearly separable even classes of stimuli that otherwise, with fewer dimensions, would not (Cover, 1965; Legenstein & Maass, 2007).

Evidence for the mechanism proposed by the SDN model comes particularly from neural network simulation studies. For instance, Karmarkar & Buonomano (2007) developed a neural network simulation composed of many excitatory and inhibitory units that exhibit a wide range or synaptic and cellular time-dependent properties and short-term plasticity mechanisms. What they found is that such a network is capable of representing different durations encoding them as unique patterns of activity that gave rise to different evolutions of the network's state over time, even without an explicit mechanism providing a linear metric for the passage of time. Moreover, simulations of cerebellar networks showed a similar pattern of results, pointing out a series of processes that closely resemble the hypothetical machinery of the SDN model. The cerebellum is indeed one of the brain regions that are thought to have a role in some aspects of the timing abilities, and particularly regarding motor timing – an idea strongly supported by experimental evidence (Ivry et al., 1988; Garcia & Mauk, 1998). For instance, it has been shown that modest lesions at the level of the cerebellar cortex can disrupt motor response timing (Perrett et al., 1993). According to this data, it has been proposed that the motor timing mechanism might rely on the spatiotemporal dynamic of cerebellar granule cell population, with these dynamics arising from the interaction between the internal states of the cerebellar networks with external inputs (Mauk & Donegan, 1997; Medina et al., 2000). This pattern of activity of the cerebellar granule cells would not only represent all potentially behaviorally-relevant stimuli, but also intrinsically contain a specific time stamp of the time that has elapsed since the onset of previous potential stimuli.

Besides neural networks simulations, several in vitro and in vivo experiments have shown that biological systems could intrinsically support the processing of temporal information (Goel & Buonomano, 2014). For instance, studies on electric fish (Carlson, 2009), crickets (Kostarakos & Hedwing, 2012) and frogs (Alder & Rose, 2000), identified classes of neurons responding selectively to signals (electric pulses in the first case, auditory patterns in the other cases) repeated at specific intervals – a behavior that suggests a clear selectivity for the temporal structure of the external stimuli. Moreover, neurons that respond selectively to the temporal and spatio-temporal structure of the external sensory signals have been identified also in the auditory system of birds (Margoliash, 1983; Doupe & Solis 1997) and mammals (Brosch & Schreiner, 2000; Sadagopan & Wang, 2009).

Other evidence suggesting a sensory-specific, intrinsic, timing mechanism is the observation that adaptation effects on perceived time are spatially-localized – i.e. specific for the region of the visual field where the adaptor stimulus was presented (Johnston et al., 2006; Burr et al., 2007). The possibility to selectively adapt a given portion of the visual field, leaving the other regions unaffected does not seem consistent with the idea of a centralized, amodal clock, while a timing mechanism deeply rooted in the sensory processing stream seems more suitable to explain this effect. The effects of motion and motion adaptation will be discussed more extensively in the next paragraph concerning temporal distortions, and in Section 3.2, where an experimental investigation about the specificity of the motion-induced duration compression for the motion profile used will be reported.

Overall, the emerging framework that offers a radically different viewpoint of the processing of temporal information provides an intriguing step towards a more biologically-plausible account for our experience of time. However, these state-dependent computations are not meant to completely replace a more centralized, amodal and cognitive mechanism for time processing, as proposed by many other frameworks. Conversely, different mechanisms could coexist and be exploited for different ranges of durations – an idea that was anticipated even at the end of the 19th century, with some similarities with contemporary intrinsic frameworks, when Mustenberg suggested that very short intervals might be directly perceived by sensory mechanism (Ulrich et al., 2006). Indeed, the physiological machinery that drives neuronal networks could be useful to discriminate durations of few hundreds milliseconds, but such time-dependent properties might provide an inadequate resolution for longer durations (Buonomano & Karmarkar, 2002; Genovesio et al., 2006). Experimental evidence of the distinction between very short and longer durations is limited to a few studies, but seems to point out that the coexistence of more temporal mechanisms could be a suitable solution. For instance, it has been shown that a secondary task parallel to the timing task (Rammsayer et al., 1991; Ulrich et al., 2006) or pharmacological manipulations (Rammsayer, 2006) affect the judgment of longer interval (1 s), while leaving performances for shorter intervals (100 ms) almost unaffected. So the possibility that intrinsic and dedicated models might represent two faces of the same coins, applied to different timescales, remain an open possibility that would require further evidence to be elucidated.

3.1.3 Distortions of perceived time, cross-dimensional inluences and the intertwining between time and other magnitudes

Our psychological representation of time very often deviates from physical time. It is very common to experience a faster passage of time when we are having fun (time "flies"), while in a boring situation time seems to pass very slowly. Another very common illusion, is the "stopped clock" illusion: when we suddenly move our gaze toward a clock, the second hand seems momentarily stopped in place, before continuing to move with its regular pace after a time that subjectively appear to be more than a second. This effect could be placed among the many perceptual effects of saccades, and, indeed, many researchers started in recent years to look at saccades as a source of temporal illusions, which might be useful to elucidate some aspects of, on the one hand, the processes for visual stability (see Chapter 2), and, on the other hand, the mechanisms supporting time perception. Such “stopped clock” effect (or “saccadic chronostasis”) was first explained by Yarrow and colleagues (2001), hypothesizing that the gap leaved by the eye movement (due to the suppression of visual signals; see Chapter 2), might be filled retrospectively with the immediately post-saccadic image, leading to the increased duration of a still second hand. However, Morrone et al. (2005) further showed that the apparent duration of stimuli presented during saccades is strongly reduced (or compressed), up to a factor of 2: the temporal interval between two visual (but not auditory) stimuli briefly presented across a saccade is underestimated. This observation is particularly interesting, because during saccades different magnitudes are distorted in a similar way (Burr et al., 2010). For instance, comparing the saccadic compression of time with the localization distortions leading to a compression of space (discussed extensively in Chapter 2), the two effects appear to be very similar: the time course of both the effects is very close, occurring for stimuli presented from shortly before to shortly after the eye movement, and with the maximum effect for stimuli presented near the saccadic onset (Morrone et al., 2005). Moreover, it has also been shown that saccades compress also symbolic and non-symbolic numerical magnitudes, again with a very similar time-course (Burr et al., 2010; Binda et al., 2011; 2012; see Chapter 5 for a more detailed discussion of similar effects on the representation of space, time and numbers). Even if the mechanisms leading to saccadic effects are not yet completely clear, the effects observed on space, time and number seem suggest a partially shared mechanism supporting the perception of different magnitudes.

Indeed, many temporal distortions are linked to the modulation of other non-temporal features of the stimuli (Eagleman, 2008). Stimuli that are brighter, bigger, or containing higher numerosity, are perceived to last longer in time compared to stimuli with the same duration but with lower magnitudes along these dimensions (Xuan et al., 2007; 2009; Rammsayer & Verner, 2014; 2015). Particularly interesting, as in the case of the distortions at the time of saccades, is the intertwining between time and other magnitudes, such space and number: a modulation in one of these two magnitudes cause a change also in the temporal dimension of the stimuli, with apparent duration increasing as size or numerosity increases – as pointed out by the duration overestimation of bigger or more numerous stimuli compared to smaller or less numerous ones (see Chapter 5).

Another source of duration distortions is motion. It has been shown that moving stimuli are perceived to last longer in time compared to static stimuli, and faster stimuli have a longer apparent duration compared with slower ones (Brown, 1995; Kanai et al., 2006). Motion is intrinsically defined as a combination of spatial and temporal information (change in position over time), so this effect might not seem totally unexpected: as concurring spatial or numerical information could affect the representation of time, also concurring motion information could distort it. However, such modulation of perceived time caused by dynamic stimuli has been observed also using flickering stimuli (Kanai et al., 2006), suggesting that not only motion can change the representation of time, but generally a dynamic modulation of the stimulus over time (in this case, a modulation of stimulus' luminance).

Another kind of motion-induced distortion of perceived time is related to motion adaptation: a sustained period of adaptation with a fast moving stimulus causes a compression of the apparent duration of a subsequent, slower, stimulus presented in the same region of the visual field (Johnston et al., 2006), even when the effects of motion adaptation on motion itself (distortion in perceived speed) – which can alter perceived duration by themselves – are controlled for. Adaptation is known to be a recalibration process triggered by the sustained presentation of the same stimulus for a relatively long period of time, which sometimes has dramatic effects on different stimuli viewed immediately after the adapted stimulus, and adaptability is considered as a characteristic of "primary" perceptual features (Boynton, 2004; Kohn, 2007; Solomon & Kohn, 2014). Indeed, many visual features showed adaptation effects, like color, shape, orientation, contrast, etc. Motion itself shows different adaptation effects, like the Positional Motion Aftereffect (Snowden, 1998; Nishida & Johnston, 1999), well represented by the classic "waterfall illusion", or the velocity aftereffect, affecting the perceived speed of moving stimuli (Thompson, 1981). Usually, such effects are in terms of a repulsion from the features of the adaptor stimulus: fast moving or flickering stimuli leads to a reduction in perceived speed, while slower stimuli cause an increase; positional shifts after motion adaptation are in the opposite direction of the adapted motion. A peculiar feature of the motion-induced compression of apparent time is that the effects are not asymmetric: while high-speed adaptation (usually 20 Hz) causes the aforementioned reduction in perceived duration, slower speed adaptation (usually 5 Hz) does not affect apparent time. What is the precise mechanism behind such temporal effect is still not clear. However, a more comprehensive discussion of the debate concerning this effect will be provided in the next chapter.

Other temporal distortions concern the difference in perceived duration between repeatedly presented stimuli and novel, unexpected stimuli; or, the effect of predictability on subjective duration (which was previously discussed in the section about the different models of time perception) (Pariyadath & Eagleman, 2007). When an oddball stimulus is suddenly presented interrupting a stream of a stimulus repeated many times, its apparent duration is overestimated by about 50% (Tse et al., 2004; Ulrich et al., 2006). Moreover, this effect of subjective time dilation is not related to a generalized slow-down of the entire visual (or auditory) scene – which should lead, beside the time dilation, also to a change in the perceived temporal frequency of other sensory information – but it is actually a specific modulation of purely temporal information. Also, an increased saliency of the stimulus does not influence duration judgments, but, conversely, the time dilation is entirely dependent on the contrast between a repeated stimulus and a novel one. Interestingly, Pariyadath & Eagleman (2007) pointed out, as also described above (Paragraph 3.1.4,) that the effect concerns also the predictability of abstract numerical sequences: in a regular sequence of number, the first stimulus to be presented (i.e. 1) is perceived to have a longer duration compared to the following stimuli (i.e. 2, 3, 4, 5...). However, if the sequence is scrambled (3, 5, 1, 4, 2...), the effect disappear, showing that such modulation is likely to be dependent on higher-level mechanisms processing not only the very basic features of the stimuli, but also more abstract concept like the meaning of numerical symbols. This latter result seems also suggesting a link between the processing of time and numbers by means of a partially common mechanism: when the number are presented in their correct order, the predictability of the sequence (i.e. increasing number) leads to dilation of the first stimulus followed by a decrease in subjective duration for the other numbers; on the other hand, when the sequence is scrambled, the usual order of the magnitudes is violated, and so it is no more possible to create an expectation, and so no effect on subjective duration is observed.

3.1.4 Conclusion and outline of the next sections

While the precise mechanisms of time perception still remain not completely understood, research on temporal distortions represent a promising field to clarify at least some aspects of these mechanisms, and to test the predictions of the different theories. Even if a unitary framework to account for all the experimental observations is still lacking, this stream of research represents a fundamental milestone to constrain current and future models concerning the perception of time.

In the following sections, we will discuss a series of experiments concerning human time perception abilities and the underlying functional mechanism. In the first experiment, we have investigated the selectivity of the motion-induced duration compression for the specific motion profile used as adaptor and test stimulus – a feature of this effect that would be very important to constrain the current hypotheses concerning the machinery behind such effect, that represents a good example of the relation between time and the combination of different features (Fornaciai et al., 2014; Fornaciai, Arrighi, & Burr, Under review). Furthermore, we will present an experiment concerning the predictions of the State-Dependent Network model (Buonomano & Maass, 2009). In this work, we investigated the influence of the InterStimulus Interval (ISI) on duration discrimination performances in different sensory modalities, according to the idea that neural networks supporting the processing of temporal information need a proper reset time between two consecutive durations in order to discriminate them (Fornaciai, Markouli, & Di Luca, In preparation).

3.2 The effect of complex motion adaptation on perceived time

3.2.1 Introduction

Despite the large amount of work that the field of time perception has attracted in the past decades, the mechanisms exploited by how brain to process temporal information are still not completely understood. One strategy to shed light on the neural and sensory processes for time perception is to study how perceived duration might be modulated by other processes or features of the stimuli.

There is a good deal of evidence that event duration can be influenced by many factors, including motion (see Section 3.1). For example, objects in motion, or contrast-reversed over time, seem to have longer duration than stationary objects (Kanai et al., 2006; Brown, 1995; Roelofs & Zeeman; 1951). Perceived duration is also affected by adaptation: after viewing fast translating stimuli, the perceived duration of a subsequent, slower, stimulus presented in the adapted location is underestimated by about 20% (Johnston et al., 2006). Importantly, alteration of perceived time is independent of that produced by direct variations of stimulus speed, as it occurs even after compensating speed to offset the effects of adaptation (Burr et al., 2007) or adapting to sequences of fast-slow motion that do not affect perceived speed (Ayhan et al., 2009). Adaptation-induced aftereffects on perceived time do not generalize to the entire visual field but are spatially selective to the region that has been adapted (Johnston et al., 2006; Ayhan et al., 2009): Burr et al. (2007) went on to show that the spatial-selectivity is spatiotopic not retinotopic, suggesting that the reference frame of the adaptation-induced distortions of event time is in real-world, rather than retinal coordinates (but see also Bruno et al., 2010; Latimer et al., 2014; Burr et al., 2011).

It is still far from clear why adaptation to fast-moving stimuli should affect event-time, especially as it does so directly, not via perceived speed or other intermediate mechanisms. A good first-step towards understanding the mechanisms would be to define the conditions under which the compression occurs. Most previous studies have used simple translational or counter-phased motion (Johnston et al., 2006; Burr et al., 2007; Ayhan et al., 2009; Bruno et al., 2010; Burr et al., 2011; Curran & Benton, 2012). From the few available studies, it would seem that adaptation to complex motion does not strongly distort duration estimates. For example, Curran and Benton (2012) showed that adaptation to an upward drifting plaid defined by two superimposed sinusoidal gratings produced strong compression of perceived time, of about 10-15%. However, when the two adapting motion components comprised transparent clouds of dots (each individually distinguishable), duration estimates remained veridical. Arrighi and colleagues (2010) reported that adaptation to biological motion had little effect on perceived duration. Both studies suggest that adaptation to complex motion does not affect perceived duration.

To better understand the mechanisms underlying the interactions between visual motion and event time, we investigated in this study whether complex motions, such as expansion or rotation affect perceived duration. For simple and complex motions we measured the effects of adaptation to fast motion both on perceived speed and on stimuli perceived duration when the speed of the adapted and non-adapted stimuli were perceptually matched. We then investigated whether adaptation to one motion profile distorts temporal estimations of subsequent stimuli moving along different trajectories to test whether and to what extent distortions of perceived time generalized across different kinds of visual motion. Finally, we investigated whether adaptation to motion that entails multiple motion directions (characterizing rotational and radial motion but not linear translation) is a key factor in causing time distortions.

3.2.2 Materials and methods

Subjects

A total of 25 subjects participated in the several experiments of the study (mean age 25 years), with some participating in either a single or multiple experimental conditions. All participants had normal or corrected-to-normal visual acuity, gave informed written consent and were naïve to the purpose of the study (with the exception of Author M.F. who participated in all experiments). The study was approved by the local research ethics committee and adhered to the Declaration of Helsinki. In Experiment 1 we tested 12, 10, 6 and 5 participants for translation, radial, rigid and non-rigid circular motion condition respectively. In the experiment in which we manipulated the stimulus presented in the un-adapted locations we tested 6 subjects in all conditions whilst 5 subjects participated to the experiment in which the moving stimuli were cloud of dots and not gratings. In the condition in which we tested the reference frame of the adaptation aftereffects (either retinotopic or spatiotopic) we tested 6 participants for the translation, radial and circular motion condition, while 5 subjects were tested in the circular non-rigid condition. In the two experiments in which we investigated the effects of perceived speed on stimuli duration we tested 5 participants in both experimental conditions: "No adaptation" and "No speed-matching". We tested 3 participants in the experiment in which the speed of the adapting stimulus was adjusted to do not trigger changes in the perceived speed of the adapted stimulus, whilst 6 subjects participated to the experiment in which we tested different combinations of motion profiles for the adapting and test stimuli. Finally, in the final experiment in which we manipulated the visible are of the adapting stimulus we tested 8 participants with the only exception of the "full-patch" adapter condition in which we tested 5 participants. All the group sizes reported above include the only non-naïve subjects M.F. (author of the study) who participated in all the experimental conditions.

Stimuli and procedure

Visual stimuli were presented on a Barco CRT monitor (Barco Calibrator, screen resolution 1024x768, 32 bit colour depth, refresh rate 100 Hz and mean luminance 27 cd/m^2) subtending about 40°x 30°at subjects viewing distance of 57 cm. Stimuli were generated with the Psychophysics Toolbox V.3 (Kleiner et al., 2007) for MatLab (version 2010b) running on a PC computer. All experiments comprised a test phase consisting of the presentation of a sequence of two stimuli, test (variable duration) and reference (fixed duration) with subjects required to indicate which stimulus, test or reference, lasted longer (guessing whenever uncertain). On each trial the duration of the reference stimulus was constant at 500 ms, while that of the test was variable, determined by the adaptive QUEST routine (Watson & Pelli, 1987). While in the first trial of the session (usually 30 trials), the test duration was initially 500 ms (same as reference), in subsequent trials, the test duration was the PSE estimated by the QUEST routine, plus a random value drawn from a Gaussian distribution with standard deviation 30 ms. This procedure ensured that there was considerable scatter around the PSE, and that the number of "longer than" and "shorter than" trials were roughly equal. Usually 3-5 sessions were run for each condition, in randomized order. The final estimate of PSE was taken as the median of the best-fitting cumulative Gaussian function to all the data of a particular condition (percentage "longer than" against test physical duration). Examples of these functions are shown in Fig. 3.4 (lower panels) for duration estimates of two different kinds of motion profiles (translational and radial motion). On separate sessions, subjects performed duration discrimination after being adapted to fast motion. Each trial started with a moving stimulus subtending an area of 5°X 5°displayed for 7 seconds on given location of the visual field. On the following test phase, the test stimulus was presented in the adapted location and the reference in a neutral position at an equivalent distance from fixation. The total amount of adaptation-induced distortions on perceived duration was defined as the difference between the PSEs of the baseline and the adaptation condition (both conditions show in Figure 3.4 with black curves and symbols referring to the no-adaptation condition and data in red to the adaptation condition).

As well as measuring the effect on duration, we measured the effect of adaptation on perceived speed, using a similar procedure. Again one region was adapted to fast motion, and then test and reference were presented as before. Again the reference was constant (at 10 Hz) and the QUEST procedure determined the speed of the adapted stimulus to home in on the speed match. PSE was given as the mean of the psychometric functions, with a minimum of 60 data points (upper panels in Figure 3.4). In most duration discrimination experiments, the physical speed of the test (adapted) stimulus was increased to perceptually match that of the non-adapted reference.

We also used a condition where the stimulus in the adapted position was kept constant (TF 10 Hz, duration 500 ms, as the reference stimulus), while the stimulus in the non-adapted position was varied according to the quest routine. In this condition, the speed-matching was applied to the non-adapted stimulus, reducing its physical speed in order to match the perceived speed of the adapted stimulus.

We studied adaptation-induced distortions on perceived duration for four different kinds of motion profiles: a) linear translation; b) radial motion; c) rigid circular motion and d) non-rigid circular motion. Translating stimuli were vertical luminance modulated gratings (SF 1 cpd) drifting horizontally, changing drift direction every 2 seconds. Radially moving stimuli were concentric circular gratings with a spatial frequency ranging from 0.5 to 1.2 cpd (outer and inner border respectively), alternatively expanding and contracting, again changing direction every 2 second. Circular motion was tested with both rigid and non-rigid rotation. Rigid-motion stimuli were windmill-like rotating gratings (spatial frequency increasing from 0.5 to 1.2 cpd from the outer to the inner border), while non-rigid rotational stimuli were defined by a series of 7 concentric circle gratings, each of them with a spatial frequency of about 1 cpd. All stimuli were windowed within an annular mask (inner and outer diameters equal to 2 °and 10 °respectively), with borders blurred by a Gaussian smoothing (spatial constant equal to 0.15 deg) and were presented with a Michelson contrast of 90%.

In a separate condition, we studied translational, radial and circular motion profiles using random-dot patterns. These stimuli comprised 100 dots (50 black and 50 white), each of diameter of 0.15 deg, randomly positioned within a circular area of radius 5 deg. Dots moved along translational (left/right), radial (inward/outward) or circular trajectories, with the speed of each element kept constant at all eccentricities. Each dot had a limited lifetime of 5 frames, after which its position was reassigned randomly to another portion of the stimulus.

In one experimental condition we used an adaptation technique devised by Ayhan, Bruno, Nishida and Johnston (2009), adapting with gratings alternating in speed between 5 and 20 Hz, and varying the ratio of the two speeds to determine a stimulus that would not cause any change in apparent speed (see panel A in Fig 3.11). This ratio was then used as the adaptor, without the need to change the speed of stimuli in the adapted location.

In most conditions we measured adaptation-induced distortions on perceived duration with similar motion profiles of adaptor and test/reference stimuli. However, to test whether adaptation to a motion profile affects duration estimates for stimuli moving along different trajectories, we added two additional adaptation conditions in which subjects adapted to translational/radial motion, and subsequently estimated duration of a radial/translational stimulus. Finally, we also performed a series of experiments to investigate the role of multiple motion directions to investigate whether this is the key factor determining whether motion adaptation affect perceived time.

3.2.3 Results

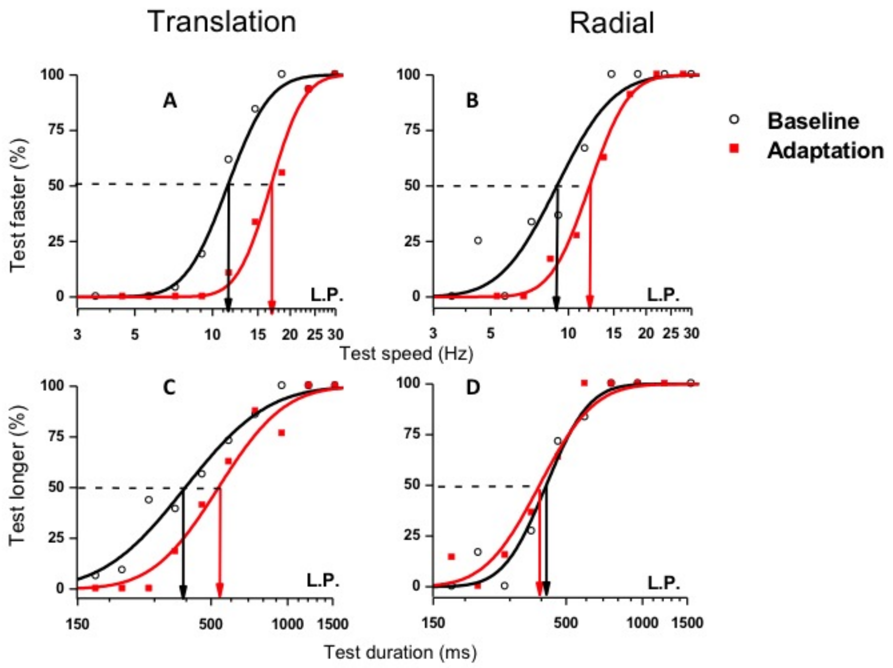

Figure 3.4: **Psychometric functions for the speed or duration discrimination tasks.** Psychometric functions of a typical subject showing the proportion of trials the test stimulus appeared to move faster or to last longer than the reference stimulus (fix duration and speed, 500 ms and 10 Hz respectively), as a function of speed (panels A&B) or duration (panels C&D) of the test. Data in black refer to the baseline condition in which the discrimination tasks were performed without adaptation. Data in red refer to the adaptation condition in which subjects, after adaptation to fast motion, discriminated either the speed or duration of the test stimulus presented in the adapted location, relative to a reference stimulus presented in a neutral position. The results indicate that adaptation to both motion profiles, translational and radial motion, distorted perceived speed as shown by the rightward shift of the red curves in panels A and B. However, apparent duration was compressed just after adaptation to translational motion (rightward shift of the red curve in panel C) but not to radial motion as shown by the almost overlapping curves in panel D.

We measured the effect of adaptation to motion on apparent speed and apparent duration (in separate sessions). Figure 3.4 shows typical psychometric functions – the proportion of trials where the test stimulus (in the adapted region) appeared to move faster or appeared of longer duration than the reference (in an unadapted region) – as a function of speed or duration of the test. Results after adaptation are shown by filled squares, baseline by open circles. For the speed measurements (panels A&B), adapting with both circular and translational motion had similar effects, shifting the red curves rightwards, towards higher speeds (to compensate for the reduction in perceived speed). The point of subjective equality (PSE) is given by the median of the psychometric functions. For both translational and circular motion, in the baseline condition the PSE was near the physical speed of the reference grating (10 Hz), but shifted to around 13-17 Hz after adaptation.

The effect on perceived duration was quite different (panels C&D). Adaptation to translational motion caused a clear change in perceived duration, requiring on average about 30% increase to compensate for the effect of motion-adaptation (which decreased apparent duration), consistent with previous research (Burr et al., 2007; Burr et al., 2011). However, adapting to circular motion had no effect on apparent duration: the PSEs for baseline and adapted condition were virtually identical. Note that for the duration judgments, the speed of the test was adjusted to appear identical to the probe, following similar procedures to those used in previous research (Burr et al., 2007; Burr et al., 2011).

Figure 3.5 shows the effect of motion adaption on apparent speed and duration, for all subjects for the four different conditions: translation, radial, rigid and non-rigid circular motion. The red arrows near the axes show the group averages. The magnitude of the effect varies considerably between subjects: but while translational motion clearly affects both perceived duration and perceived speed, the other forms of motion affect only speed, not duration, on average. We ran a Multivariate analysis of variance (MANOVA) with changes in perceived speed and duration as the two dependent variables and motion profiles as the independent variable at 4 different levels. The results indicate that there was no significant effect of motion profiles on changes of perceived speed of the adapted stimuli (F(3,29)=0.265, p=0.85), but a significant effect

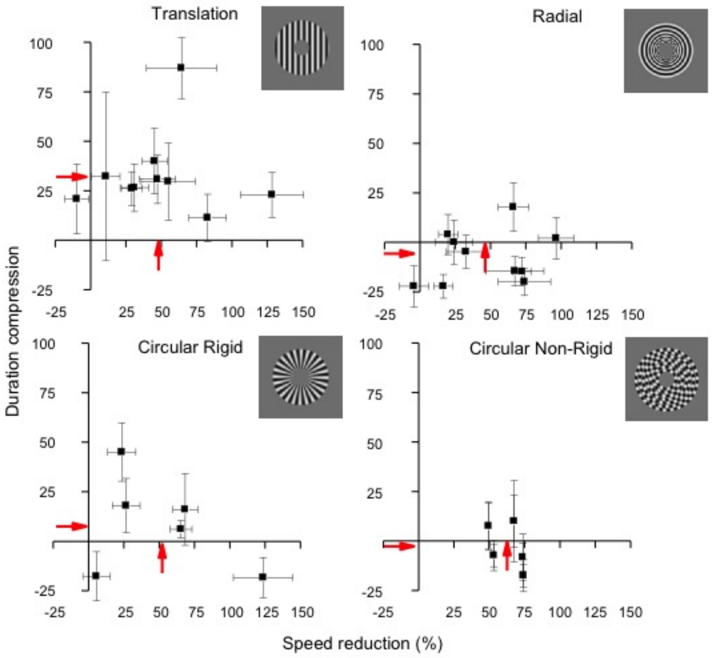

Figure 3.5: **Effects of adaptation to fast motion on perceived speed and duration.** Scatter plot showing the effect of motion adaptation to four different kinds of motion on both, perceived speed (abscissa) and duration (ordinate). While adaptation to fast motion compressed perceived duration regardless the motion profile taken into account (on average by about 50-60% as shown by the red arrows on the abscissa axes), perceived duration was affected by adaptation only to translational motion (average compression around 30%). After adaptation to non-translational motion, the duration estimates remained veridical.

of motion profiles on changes in perceived duration (F(3,29)=11.89, $p<0.001$). Post-hoc analysis showed that this difference was due to translational motion being statistically different from radial motion ($p<0.019$), circular rigid motion ($p<0.02$) and circular non-rigid motion ($p=0.02$) with no difference amongst the non-translational motion profiles (all p values > 0.2).

In all the experiments discussed above, we measured speed and time PSEs by modulating the stimulus in the adapted position. As this procedure tends to lead to very large compensations in speed, we repeated the experiment (with translational, radial and circular rigid motion stimuli), varying the stimulus in the non-adapted position, keeping constant the adapted one throughout the session to be certain that this procedure did not lead to measurement artefacts. The results are shown by the purple bars of Fig. 3.6. Although the effects

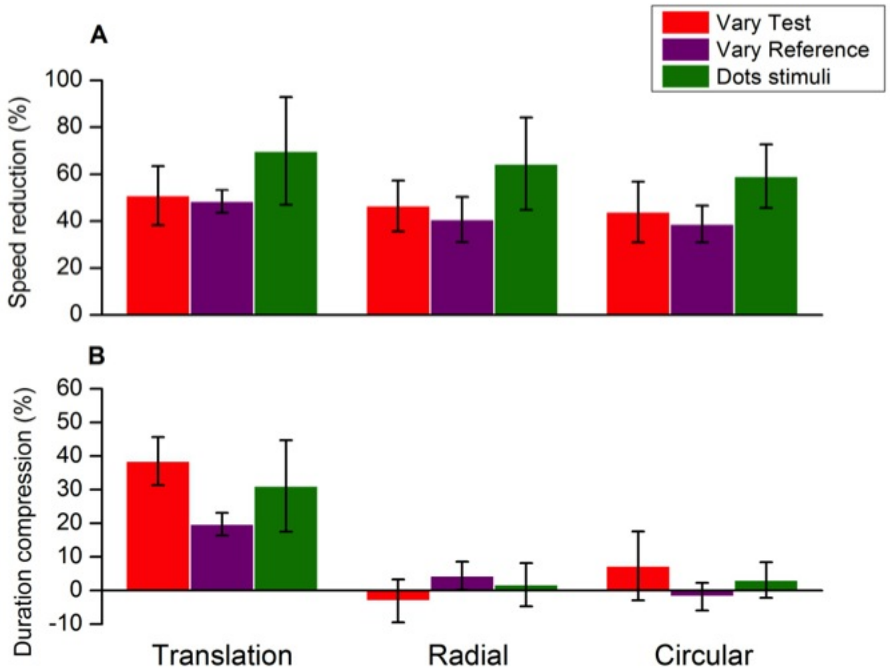

Figure 3.6: **Average effects of adaptation on speed and duration for three different experiment conditions.** Upper graphs show the effects of adaptation on perceived speed, for three types of motion, translation, radial and circular. The three experimental conditions are: “Vary test” – results taken from Fig. 3.5 (except for circular non-rigid condition), where stimuli in the adapted position were varied; “Vary reference” – where the stimulus in the non-adapted position was varied; “Dots stimuli” – experiment with random dots instead of gratings.

are slightly smaller, they replicate the findings of the previous experiment: we found a strong reduction in perceived speed after adaptation for all three conditions (baseline versus post-adaptation, $p < 0.001$, $p < 0.001$ and $p = 0.002$, respectively for translational, radial and circular motion, not different for the three motion profiles: $p = 0.420$). On the other hand, similarly to the previous experiment, we found a robust effect of duration compression (Fig. 3.6b) only for translational motion stimuli (baseline versus adaptation, $p < 0.001$, $p = 0.243$, $p = 0.756$), which was significantly different from the effect of the other two motion profiles (translation versus radial, $p = 0.026$; translation versus circular, $p = 0.009$).

Another issue of radial and circular rigid stimuli concerns the difference in spatial frequency from the inner to the outer portions of the patches: spatial frequency decreased in inverse proportion to the radius, which may introduce a variation in perceived speed in different portions of the stimuli, and hence create an acceleration or deceleration pattern from the centre to the periphery. To avoid this problem, we also used circular non-rigid motion stimuli, which showed a similar effect. But to be certain, we also repeated the experiment with clouds of dots moving along translational, radial or circular trajectories. Random-dot patterns have a broadband spatial frequency spectrum, which should not interfere with perceived speed or introduce acceleration or deceleration patterns by itself. As shown in Fig.3.6 (green bars), we found a very similar pattern of results compared to the previous conditions. While adaptation to these stimuli affected the perceived speed irrespective of the motion profile, the effect on apparent duration was again specific for the translational motion profile. (Baseline versus post-adaptation comparison significant for all the speed conditions, p-values < 0.002, with no significant difference between different conditions, p-values > 0.05. For duration baseline versus post-adaptation comparison significant only for translation, $p = 0.0257$; p-values > 0.05 for the other conditions; translational motion effect significantly different from the other conditions, $p = 0.044$ and $p = 0.028$, respectively for translation versus radial and versus circular). Overall, these two additional experiments reinforce the main one, showing that the observed pattern of results does not result from measurement issues.

Why should adaptation affect duration only for translating images? One possibility is that we are less precise at estimating the duration for translating images, and this imprecision leads to a larger adaption effect. For example, Van Der Burg, Alais and Cass (2013) showed a clear correlation between the magnitude of the adaptation effect and size of the simultaneity window. To test this possibility, we calculated the Coefficient of Variation (CoV) for the various tasks, defined as the standard deviation of the psychometric functions divided by the duration of the probe (500 ms). We plot these in Fig. 3.7A as a function of adaption magnitude (the change in PSE normalized by the pre-adaption PSE). It is obvious that there is no relationship between the two measures. Pearson's correlation coefficient $r = -0.02$, not significantly different from zero. In addition, there was no significant difference in CoV between the four conditions (one-way ANOVA on ranks, $H = 4.22$, $df = 3$, $p = 0.238$).

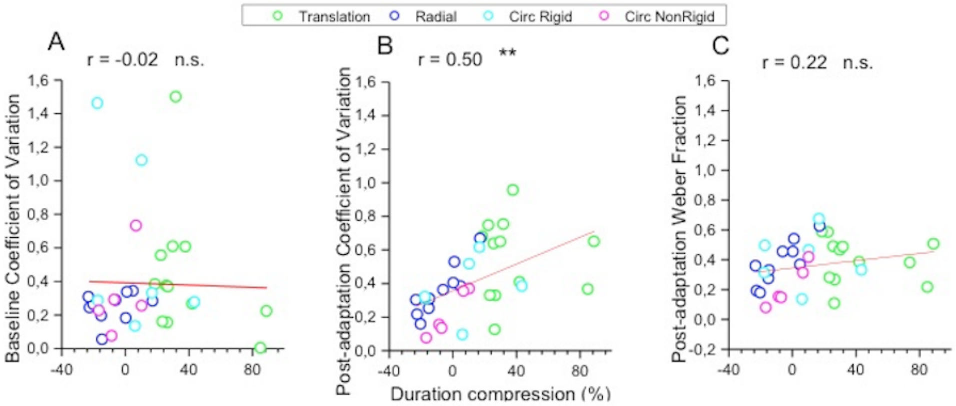

Figure 3.7: **The effect of motion adaptation on Coefficient of Variation (CoV) and Weber fraction.** The amount of adaptation-induced compression of perceived time (abscissa) plotted against the Coefficient of Variation (CoV, defined by thresholds for duration discrimination normalized by the reference duration) before and after adaptation (panels A and B respectively), for four different motion profiles (data in different colours). Whether pre-adaptation CoV did not correlate with duration compression, a robust correlation was found with post-adaptation CoV. This indicates that motion adaptation affected both PSE and JND in the duration discrimination tasks and this, in turn, maintained Weber-Fractions (defined by the ratio of JND to PSE), almost constant (data in panel C).

Thus it seems unlikely that the failure to adapt with non-translational motion results from poorer duration discrimination for that type of motion.

However, when post-adaptation CoVs are plotted against adaptation effect, there is a significant positive relationship: slope 0.39, r = 0.50, p = 0.003. This is interesting, as it suggests that adaptation affects not only PSE but also sensitivity to duration. This is brought out more clearly in Fig. 3.7C, which plots Weber fraction against adaptation effect: Weber fractions are like CoV, except that the thresholds are normalized by PSE instead of physical number. Here the slope of the regression is much lower than for the CoV (slope 0.12, r = 0.22, p = 0.111), as the normalizing PSE increases (for translational motion) after adaptation. A similar result was observed by Morrone, Ross and Burr (2005) during saccades: they found that saccades reduce both the PSE and the JNDs of temporal duration estimates, so the Weber fraction remains constant.

In previous studies, we have examined whether motion-induced time compression is spatially specific in retinotopic or spatiotopic coordinates: there

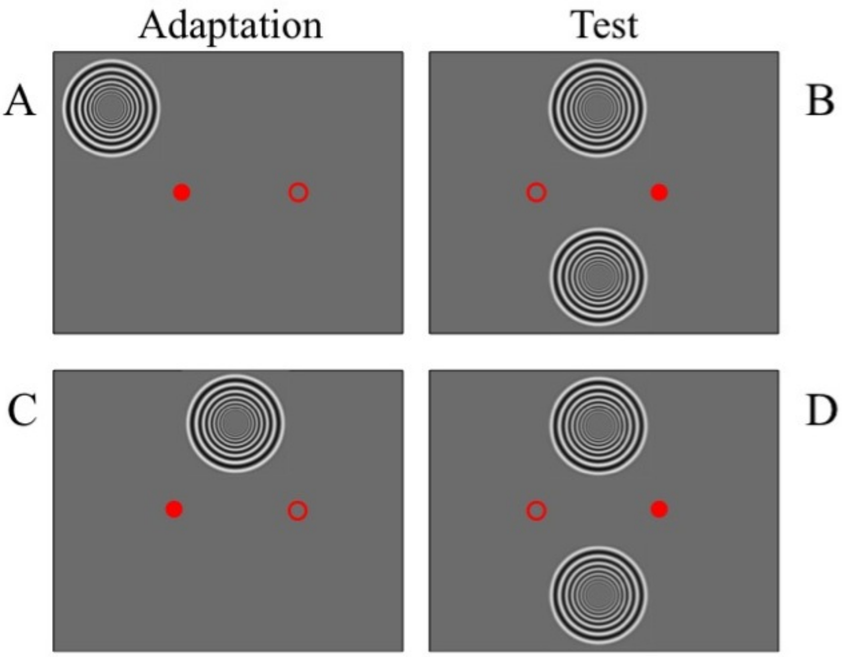

Figure 3.8: **Measuring adaptation aftereffects in retinotopic and spatiotopic coordinates.** A cartoon showing stimuli displacement for the three different experimental conditions. When the adaptation was in retinotopic (eye-centred) coordinates, subjects adapted to a fast moving stimulus displayed on the left of the fixation point, then they executed a rightward saccade and were presented with a test stimulus in the corresponding retinal adapted location (panels A&B). In the spatiotopic condition the adaptor was superimposed to the test stimulus after the saccade in external world coordinates but not in the retinal space (panels C&D). The full adaptation condition was identical to the spatiotopic condition but with fixation maintained throughout the trial. In all conditions, subjects had to judge either the duration or the speed of the adapted stimulus, relative to a reference displayed in a neutral location below the screen midline.

has been much debate of this issue, with some data suggesting that the effect is primarily spatiotopic (Burr et al., 2007; Burr et al., 2011), while others fail to confirm this (Bruno et al., 2010; Latimer et al., 2014). We repeated the measurements here using the method illustrated in Fig. 3.8. In the test phase of all conditions, subjects viewed a fixation point displayed 6.5°to the right of the screen centre (filled circle in the right-most panels), both test and probe gratings were to the left of the fixation point, the test above the probe below the screen midline. For the retinotopic condition (upper panels), the adaptor was placed to the upper-left of the pre-saccadic fixation point (displayed 6.5°

on the left of the screen centre as shown by the filled circle of the left-most panels), in the same retinal position as the test stimulus relative to post-saccadic fixation point. In the spatiotopic condition (lower panels) the adaptor was to the upper right of fixation, in the same screen position as the test after the saccade. In both cases, after adaptation was completed, the fixation point moved to its rightward position and subjects saccaded to it. The test grating appeared 500 ms after extinction of the adaptor, and the probe 500 ms after extinction of the test. The "full" adaption condition was like the spatiotopic condition, except subjects kept fixation on the left fixation point throughout the experiment.

Figure 3.9 shows the results for the four types of motion: translation, radial and rigid- and non-rigid circular motion. The upper panel shows the results for the effects on apparent speed. All forms of motion caused strong and significant effects in the full condition, around 40% reduction of perceived speed. Also in the retinotopic condition, the reduction in perceived speed occurred for all types of motion, to a similar extent as the full-adaption condition. However, in the spatiotopic condition, there was no significant change for any of the four types of motion. This result is consistent with the fact that the motion aftereffect is retinotopic, as previously reported (Addams, 1834; Wenderoth & Wiese, 2008; Knapen et al., 2009; Turi & Burr, 2012).

However, the results for duration compression are quite different. Only adaptation to translational motion yielded a significant compression of perceived duration regardless the reference frame the adaptor and test were superimposed in. As observed before (Burr et al., 2007; Burr et al., 2011), the effect was strong in the spatiotopic condition, nearly half the full adaptation condition. There was, however, also an effect in the retinotopic condition, which had not been reported in previous studies of our lab (Burr et al., 2007; Burr et al., 2011) but had been observed elsewhere (Bruno et al., 2010; Latimer et al., 2014). For the three types of complex motion, however, none of the conditions caused any compression of perceived duration. This shows that the difference between the effects of adaptation to translational and complex motion on duration perception is very robust.

Can complex motion influence perceived duration under any circumstance? To test this we measured the effect of varying stimulus speed on perceived

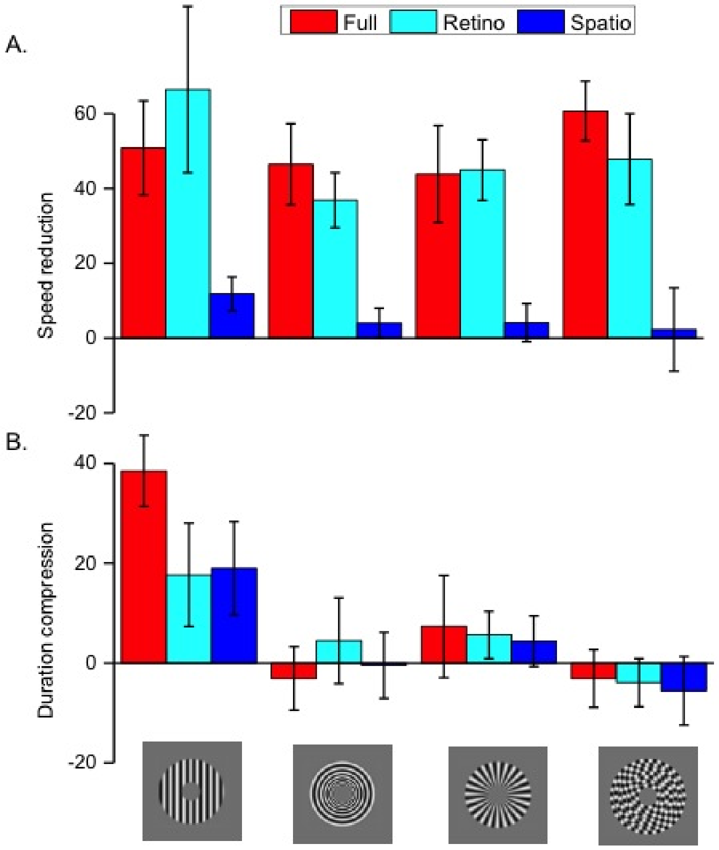

Figure 3.9: **The reference frame of motion adaptation aftereffects on perceived speed and duration.** Panel A: Reduction of perceived speed after adaptation to four different kinds of motion (translational, radial, rigid and non-rigid circular motion), in three different reference frames: retinotopic, spatiotopic or both (full adaptation). In the two conditions with a retinotopic component (retino and full, cyan and red bars respectively), the perceived speed of the adapted stimulus was robustly reduced. On the contrary, speed estimates for stimuli presented in the same spatiotopic location of the adaptor (blue bars in the figure), remained quite veridical with rather no change in the perceived speed of the adapted stimuli. Panel B: Effects of motion adaptation on perceived duration. For translational motion, we found a robust compression of perceived duration for the conditions in which the adaptor and test were superimposed in retinotopic or spatiotopic coordinates, with the sum of these effects to be similar, to that measured in the full adaptation condition (compression of about 40%). However, adaptation to radial and circular motion (both versions, rigid and non rigid) did not yield any compression of perceived duration after the speed of the test and reference stimuli was equated. This result indicates that motion adaptation compress perceived time just after adaptation to translational motion and this effect does not generalize to different kinds of motion profiles.

duration, using a technique similar to that of Kanai et al. (2006). The probe stimulus moved at 10 Hz, while the test moved at a slower speed, calculated

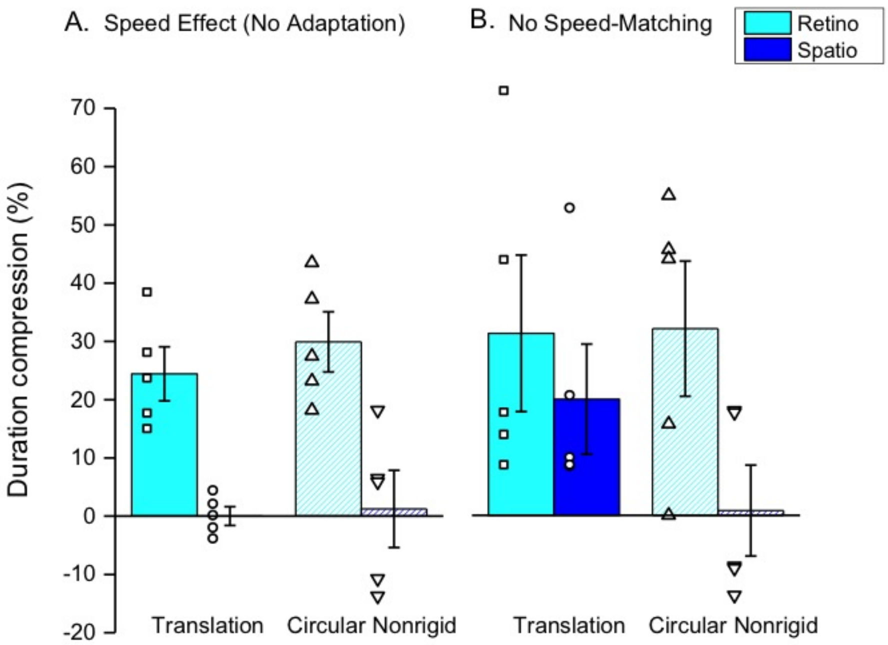

Figure 3.10: **The role of differences in stimuli perceived speed for adaptation-induced distortions on perceived time.** Panel A: The effect of the speed difference in estimating stimuli duration. Subjects performed a duration discrimination task between 2 stimuli with the speed of one of them reduced to "simulate" adaptation-induced distortions. The results show that duration estimates for stimuli moving at lower speed were compressed for both kinds of motion, translational and radial motion (full and hatched bars respectively). In one condition (data in cyan) we reduced the speed of the test stimulus according to the speed changes found for retinotopic adaptation whilst in the other (data in blue) speed reduction simulated adaptation in the spatiotopic reference frame (data derived from panel A of Figure 3.8). Panel B: Adaptation-induced compression of perceived time when the apparent speed of the adapted and reference stimulus was not equated. For translational motion, duration compression was found in both retinotopic and spatiotopic conditions with the first being strengthened relative to the values found in speed matched conditions (see panel B of Fig. 3.8) by the lack of speed compensation given that adaptation did not affect significantly perceived speed in the spatiotopic condition (blue bars in panel A of Figure 3.8). For radial motion, we found a consistent compression of perceived time in retinotopic condition. This indicates that adaptation to radial motion can affect perceived time, but not by adaptation.

to be the same as the speed reduction after adaptation to fast motion (values taken from Fig. 3.9). The bars of Fig. 3.10A show that varying image speed had a robust effect on perceived duration, causing a compression of around 25% for both translational and circular motion, as shown in the figure by solid and hatched bars respectively. A two-way RM ANOVA with Holm-Sidak multiple comparisons revealed a statistically significant difference between the duration estimates for the baseline condition (in which stimuli moved at the same physical speed) and the condition in which the test speed was reduced to simulate the effect of adaptation ($p < 0.001$ for both the motion profiles). The reduction in perceived duration occurred only in the retinotopic condition, because spatiotopic adaptation caused little change in perceived speed in either translational or circular motion (see panel A in Fig. 3.9), so the test and probe had the same physical speed during the match. These results indicate that circular motion can influence apparent duration; it is only *adaptation* to circular or radial motion that has no effect on duration when adapted and neutral stimulus perceived speed are equated.

The reason for using speeds corresponding to the reduction caused by adaptation was to simulate the effect of making duration judgments after adaptation to fast motion, without compensating for the effects on perceived speed. The bars of Fig. 3.10B show duration matches for an experiment where perceived speed was NOT compensated for: it was like the experiment of Fig. 3.9B, except that the test and probe gratings moved as the same physical, rather than apparent speed. Under these conditions, when motion adaptation occurred in retinotopic coordinates, the circular motion had a significant effect on perceived duration similar as that for translational motion (two-way RM ANOVA with Multiple comparison, Holm-Sidak method; baseline versus adaptation, $p < 0.05$ for both the motion profiles; no statistically significant differences between translational and circular motion, $p > 0.05$). In both cases, the reduction in perceived duration was similar in magnitude to the effect of varying speed (Fig. 3.10A). It was therefore probably caused by the same mechanism: the difference in perceived speed, rather than a direct effect of adaptation on duration. Importantly, there was no duration compression in the spatiotopic condition for circular motion (no difference between baseline and adaptation measures, $p > 0.05$; statistically significant difference between retinotopic and spatiotopic condition, $p = 0.036$), but only for translational motion (baseline versus adaptation, $p = 0.021$; no differences between retinotopic and spatiotopic effect: $p = 0.420$). It has been argued elsewhere (Burr et al., 2007) that the spatiotopic effect of duration component is the high-level direct effect of motion-adaptation on duration, not to be confused with the indirect effects via changes in perceived speed.

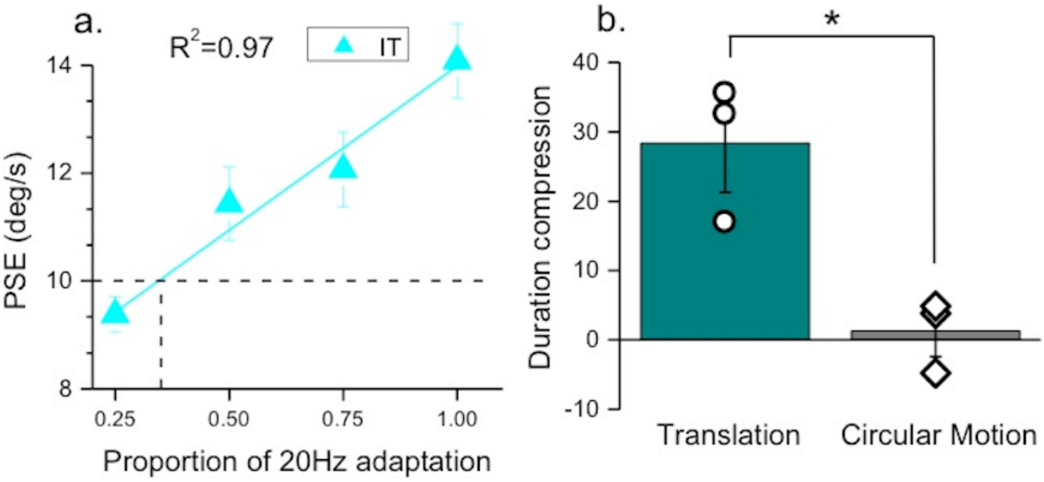

Figure 3.11: **Difference in physical speed between test and reference stimulus cannot account for adaptation effects on perceived time.** (Panel A) Changes in the perceived speed of the adapted stimulus for different proportions of fast (20 Hz) or slow (5 Hz) adapting speed (data for the typical subject I.T.). The speed ratio that did not change the perceived speed of the adapted stimulus (indicated in the figure by dashed lines) was chosen to measure the effect of adaptation on perceived duration without manipulating the physical speed of the test stimulus. (Panel B) Even with this technique, motion adaptation to translational stimuli distorted perceived duration of the test stimulus whilst adaptation to circular motion did not (data averaged across 3 subjects).

The results to date suggest that if the artefact of adaption-induced speed change is eliminated, there is no adaption-induced duration compression for complex radial or circular motion. However, if the speed change is not compensated, an effect emerges. Although we compensated for changes in perceived speed in two ways (changing either the standard or the test speed), to be certain that the results are not in some way an artefact of speed compensation, we measured the effect with another technique that causes no changes in apparent speed (Ayhan et al., 2009). We adapted with an interleaved mixture of fast and slow gratings (20 and 5 Hz) of variable ratio: the more high-frequency in the mix, the more it decreases the apparent speed of the test, and vice versa. For each subject we measured the effect on apparent speed as a function of mix ratio (for both translational and circular motion), producing curves like that of

Fig. 3.11A. We then chose the ratio that produced no speed-change for both, translational and radial motion for each subject, and measured the effect on duration for that ratio. The results are shown in Fig. 3.11B. Again, only for translational motion we found an adaptation-induced compression of perceived duration whilst estimates for circular motion remained rather veridical.

Adaptor or test?

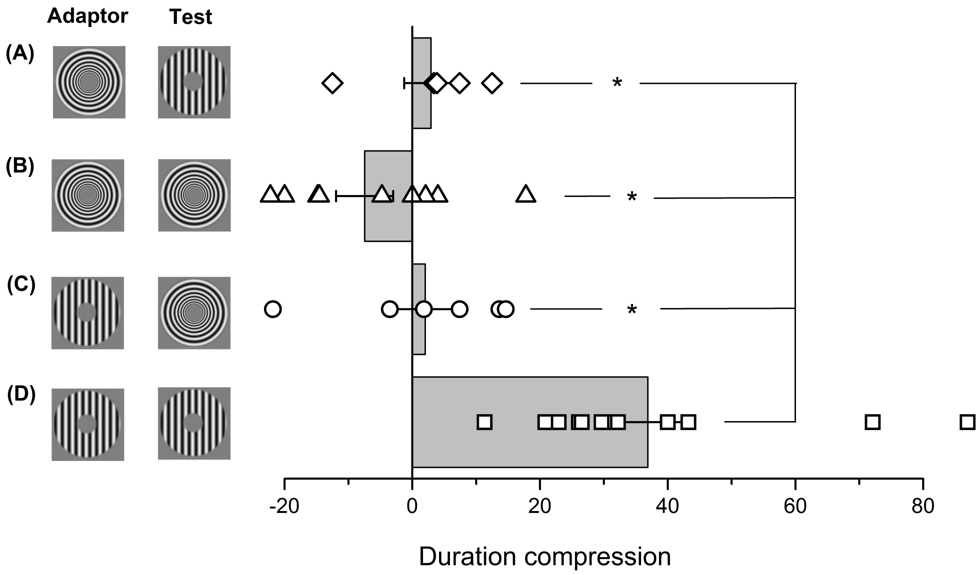

Figure 3.12: **Adapting or testing translational motion?** The amount of perceived duration compression induced by adaptation to several combinations of adapting and test stimuli. The results indicate that the only condition providing a consisted distortion of perceived time (average compression of about 40%) is the one just comprising translational motion (condition D in the figure). In all conditions involving radial motion, either as adopter or as test stimulus, duration estimated remained rather veridical (panels A, B & C).

All experiments reported so far consistently show that adaptation to translational motion, but not other types of motion, distorts perceived duration. Is it essential for the adaptor to be translating? Or the test? Or both? To address this issue we tested motion adaption induced duration compression with 4 different combinations of adaptor and test stimuli: radial-translational; radial-radial; translational-radial; translational-translational (icons in Figure

12). The bar graphs show that the only combination to cause duration compression was when *both* adaptor and test where translational motion (condition D; Signed Rank test, $p < 0.001$): if *either* the test or the adaptor were in radial motion, there was no effect as shown by a two-tailed paired t-tests ($t = 0.784$, $p = 0.47$, for condition A, $t = 1.922$, $p = 0.086$ for condition B, and $t = 0.178$, $p = 0.865$ for condition C).

Decomposing complex motion

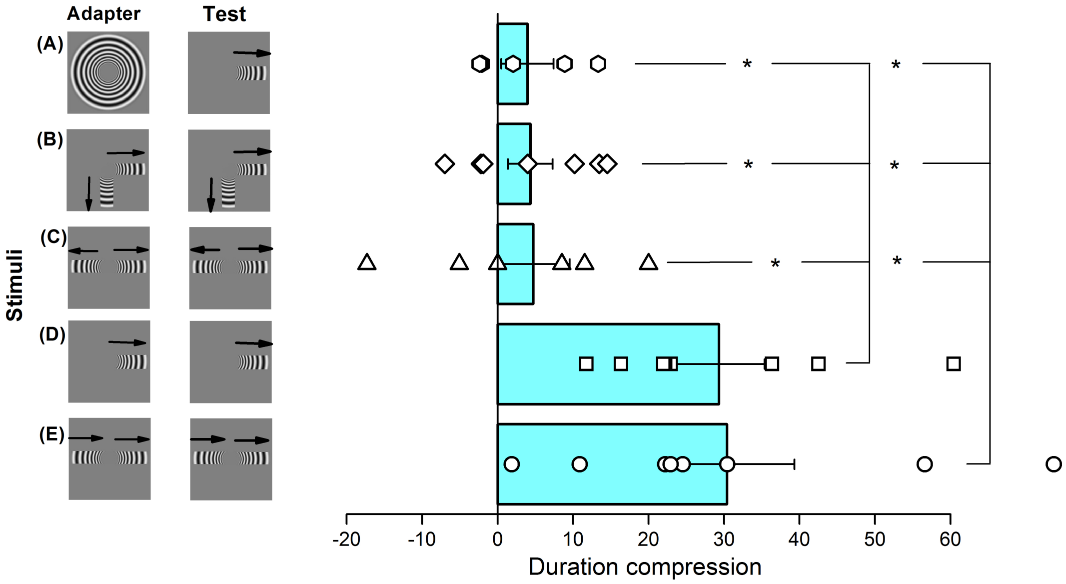

Figure 3.13: **Adaptation to multiple motion directions does not affect perceived duration.** Compression of perceived duration for several combinations of simplified versions of translational and radial motion. Adapting to a full radial patch did not affect duration of a subsequent radial patch in which a single motion direction was left visible (panel A). Similarly, when the radial patch was occluded to have two visible sectors either orthogonal (panel B) or collinear vignetting motion in opposite direction (panel C), motion adaptation did not compress perceived duration. However, when subjects adapted to a radial patch in which a single sector was visible (panel D), the duration of the subsequent stimulus was robustly compressed. The same occurred for the condition in which congruent linear motion was vignetted in two collinear sectors as that shown in panel E. Even in this case we found a strong compression of perceived time to indicate that motion adaptation affect time processing just in those conditions in which motion is defined by a single motion direction.

With translational motion, the local motion is, by definition, all in the same direction and speed. However radial and circular motion comprise motion vectors of all directions, including orthogonal and opposite direction. To test the importance of the different motion directions, we measured duration compression with simplified versions of translational and radial motion, comprising just one or two sectors (icons of Fig. 3.13). We first adapted with a full radial pattern and tested with a single sector, subtending 1.1x4.2°(Fig. 3.10A). This produced a compression of just about 4%, not significantly different from zero (2-tailed paired t-test: p= 0.26). Similarly, adapting to a simplified radial-motion stimulus defined by 2 motion directions – either orthogonal (Fig. 3.13B) or opposite (Fig. 3.13C) – did not yield significant changes in perceived duration of stimuli presented in the adapted location (p= 0.14 and p= 0.38 respectively).

However, adapting to a single motion sector of radial motion, where the local motion was almost in the same direction, did robustly compress the duration, by 29% (p = 0.0015: Fig. 3.13D). Also adapting and testing to two opposing sectors, both moving in the same direction causes strong duration compression, 30% (p = 0.006). This is clear: what determines whether adaptation causes duration compression is whether the motion is in the same direction over the entire field.

3.2.4 Discussion

The results of this study clearly show that adapting to fast translational motion reduces perceived duration, by about 30%; while adapting to more complex trajectories, such as radial or circular motion (rigid or non-rigid), or patches of grating with opposing or orthogonal motion, does not affect perceived duration. What seems to be important is that the motion be in a single direction.

Why does adaptation only to single-direction translational motion affect perceived duration? We can exclude that these results arise from poorer precision in duration judgments after adaptation to translation: there was no significant difference in the discrimination thresholds for the four types of motion used; nor there was a significant correlation between baseline threshold and magnitude of the effect for any of the experimental conditions. If the adaption depended on higher thresholds, one would have expected this correlation to be strong and significant, as has been observed in other conditions, such as adaptation of perceived audio-visual synchrony (Van der Burg et al., 2013). Clearly there is something qualitatively different about translational motion that adapts timing mechanisms.

It is important to note that while adaptation to circular motion does not compress duration, circular motion can affect perceived duration directly. The duration of a stimulus rotating at 10 Hz is perceived to be longer than stimuli rotating more slowly. Similarly, if the effect of adaptation to fast rotational motion on apparent duration is measured without compensating for the effect on perceived speed (so the adaptor and test stimuli appeared to be of different speeds), then this will also affect duration. However, when the effect on speed is compensated for directly, or if one uses an adaptation technique that does not change perceived speed (Ayhan et al., 2009), then there is no effect on perceived duration. The results clearly show that while all types of motion have a direct effect on duration, with slower-moving objects appearing of shorter duration (and vice versa), only translation causes an adaptation effect.

This result of selective adaptation-induced temporal compression finds broad agreement in the literature. As mentioned in the introduction, Curran and Benton (2012) showed that adaptation to two individually discernable sheets of transparent random-dot motion has little effect on perceived duration, whereas a plaid (also comprising two components) moving in a single direction does. On the other hand, Marinovic & Arnold (2012) have demonstrated adaption-induced duration compression with a rotating pattern of filled circles. It is not obvious why their results were different from ours. Possibly, the main difference was that the circle pattern was quite sparse and the individual elements quite large, so it was not perceived as a single rotating stimulus. It may be interesting to replicate the effects with Marinovic and Arnold's stimuli to try to understand better the differences. Hogendoorn and coll. (2010) also reported duration compression after adaptation to radial motion. However, the magnitudes of the effects were small, in the order of 10 ms, compared with 20-40% compression (around 200 ms) observed here and in previous studies (Johnston et al., 2006; Burr et al., 2007; Bruno et al., 2010; Burr et al., 2011). The methods used by Hogendoorn et al. were different from those here, it is difficult to make precise comparisons: but certainly the magnitude of the effects reported for adaptation to radial motion are far smaller than those usually reported for adaptation to translational motion and this is perfectly in agreement with the present results.

But why should translational motion, and only translational motion, adapt a specific region of space so that subsequently presented stimuli have reduced duration? This is a difficult question, for which we do not have a complete answer. However, the data of this study help to constrain the explanations for adaption-induced duration compression. For example, Johnston and colleagues (Ayhan et al., 2009; Ayhan et al., 2011; Johnston, 2010) have suggested that duration compression results from preferential adaptation of magnocellular neurons in the lateral geniculate nucleus. Evidence in favour of this idea includes the fact that motion too fast to be perceived processed by cortical units (but presumably not too fast to elicit magno¬cellular activity in the geniculate) can cause duration compression (Johnston et al., 2008). However, the fact that circular motion is equally effective in stimulating neurons in the lateral geniculate but it does not cause duration compression, seems to rule out this possibility.

The key difference between stimuli that did and did not adapt duration was whether there was more than one direction of motion displayed in the adapting stimulus (even when the test stimulus was unidirectional). It is far from clear why this should be so. There is evidence that different cortical regions are stimulated by unidirectional and flow motion, possibly corresponding to area MT and MST (Duffy & Wurtz, 1991; Morrone et al., 1995; Smith et al., 2006; Morrone et al., 2000). TMS studies have linked activation of the MT/MST complex to time processing (Bueti et al., 2008; Salvioni et al., 2013): possibly only the area that responds to unidirectional motion is linked closely with the perception of duration. It may also be of interest that the MT complex receives input directly from sub-cortical structures, bypassing V1. The input comes from the pulvinar (Bridge et al., 2015), the lateral geniculate nucleus (Sincich et al., 2004) and the superior colliculus (Lyon et al., 2010). It is possible that this direct pathway is differently stimulated by uni- and multi-directional motion, leading to different adaptation effects.

It is still far from clear what functional role motion-induced compression of event duration may serve. We know that the adaptation is spatially specific (Johnston et al., 2006), and many studies suggest that the spatial specificity may be largely spatiotopic (although the current results, as well as much research from other laboratories (Johnston et al., 2006; Bruno et al., 2010; Latimer et al., 2014), suggest there may also be a retinotopic component). The spatiotopic component in the selectivity points to relatively high-level processing areas. However, it remains far from clear why adaptation to motion affects the timing properties of these areas; and far less clear why the adaptation should need to be unidirectional.

We are fully aware that the above discussion is speculative. However, while the interpretation may be speculative, the facts are very clear: adaptation to fast translational motion causes robust changes to perceived duration, while adaptation to circular or radial motion, or to two sectors of motion in opposing or orthogonal directions does not. We believe that understanding this phenomenon will be important in understanding fully the mechanisms underlying event duration, and their interconnection with the dynamic perception of space.

3.3 Effect of inter-stimulus interval on auditory and visual duration discrimination

3.3.1 Introduction

Recent models of time perception departed from the hypothesis of a centralized and amodal clock, and rely instead on biologically plausible circuitry and neuronal properties (Matell & Meck, 2004; Meck et al., 2008; Ivry & Schlerf, 2008). Particularly, the “State-Dependent Network” (SDN) model (Buonomano & Maass, 2009), proposes that estimates of perceived duration arise from the dynamically changing activity of neural networks that are processing other attributes of the stimuli. That is, according to this model any given neuronal ensemble should be intrinsically capable of time processing. Experimental evidence indeed suggests that the processing of very brief durations is deeply rooted in the sensory streams of processing (Bueti et al., 2008; Cicchini, 2012) giving credit to this hypothesis. Moreover, neural network simulations shows that the SDN model could account for numerous characteristic features of human temporal processing (Buonomano & Mauk, 1994; Buonomano & Merzenich, 1995; Buonomano & Laje, 2010; Goel & Buonomano, 2014).

A test bench for the SDN model is that the limiting factor in the precision of temporal estimate should be determined by neuronal networks processing. Two recent studies (Karmarkar & Buonomano, 2007; Buonomano et al., 2009) attempted to test the predictions of the SDN model using psychophysical measures of precision. First, what they tested is the effect of temporal context on the ability to correctly discriminate a given duration – according to idea that timing should be context-dependent, i.e. the representation should vary as function of the initial state of the network. To test this prediction, Karmarkar & Buonomano (2007) presented intervals marked by sounds that could be preceded by irrelevant distractors presented either with a fixed interval from the test duration or with a variable interval. What they found is a significant increase in the threshold for duration discrimination when the test interval was preceded by a distractor with a variable timing, while no impairment was found with the fixed distractor, suggesting that while the specific dynamics of activity elicited by the fixed distractor could be taken into account across many trials, the variable distractor gives rise to unpredictable activity that is not reproducible across trials. This in turn leads to additional variability in the state of the network and hence to a degradation of the precision of temporal estimates.

Subsequent studies have challenged the interpretation of these results. For instance, Spencer et al., (2009) showed that the impairment in discrimination performance disappear when either the distractor or the test interval are longer (300 ms instead of 100 ms, as first tested by Karmarkar & Buonomano, 2007). The difference in the effect of variable distractors suggests that the impairment is due to attentional effects, such an attentional capture caused by the distractor, which has an effect only when the stimuli were very brief (i.e. a failure to shift attention to the test interval after the distractor). Moreover, Burr et al. (2013) further showed that the effect is present even when the distractor follow the test stimulus or is presented in a different sensory modality. Both results are at odd with the strictly causal explanation offered by the State-Dependent Network model.

Beside the contextual effects, Karmarkar & Buonomano (2007) and Buonomano et al. (2009) further tested the predictions of the SDN model studying the effect of the Inter-Stimulus Interval (ISI) on the precision of duration discrimination judgments. In these studies, they presented empty intervals marked by sounds with either the same or different frequency, and either with a short or a long ISI between the two intervals to be compared. According to Buonomano and colleagues, intervals bounded by sounds with the same frequency should be processed by the same network; conversely, if the frequencies of the markers are different, the two intervals would be represented by the activity of different neurons (Reale & Imig, 1980; Morel et al., 1993). This allowed Buonomano and colleagues to test whether a sufficient interval is required to allow the network to return to the initial state after the first interval and properly process the second interval. The model predicts that intervals processed by the same network and separated by a short pause should lead to lower sensitivity to differences in duration. Indeed, this is actually what they found: higher thresholds for discriminating the two intervals, when the durations are defined by sounds with same frequency – an impairment that seems specific for temporal judgments, since in a frequency discrimination task they did not observe any influence of the ISI.

Since Spencer et al. (2009) pointed out a strict limit for the distractor effect predicted by the SDN model, in this study we first tried to figure out whether also other disrupting effects predicted by model – namely, the effect of a too short ISI on duration discrimination performances – are subject to such constraint (i.e. whether it is limited to very brief durations, around 100 ms). Furthermore, we also investigated the effect of ISI in the visual modality, in order to test whether different modalities are prone to the same contextual effects on temporal perception and whether they follow the same dynamic and constraints.

3.3.2 Experiment 1 – auditory duration discrimination

In the first experiment, we attempted to replicate the results reported by Karmarkar & Buonomano (2007) and Buonomano et al. (2009), but using a longer duration range. In the second part of the experiment, we further investigated whether durations partially encoded by two different networks are susceptible to the ISI influence, bounding durations with different sounds, in order to present the two intervals as either the same sequence of sound frequencies or an inverse sequence of sounds.

Materials and methods

Subjects. 12 subjects participated in the experiment (8 females, age ranging from 19 to 30 years), after giving an informed written consent. With the exception of two of the authors (M.F. and E.M.) who participated in the experiment, all the subjects were naive to purpose of the study, and were rewarded for their time with 6 GBP/hour. All the participants reported to have a good hearing. Experimental procedures were conducted according to the protocol approved by the ethics committee of the University of Birmingham, and are in line with the declaration of Helsinki.

Apparatus. The experiment was performed in a quiet dark room. Observers sat at a table, with head stabilized by means of a chinrest. Auditory stimuli were generated with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MatLab (MatLab 2013b, The Mathworks, Inc.) and delivered by a speaker (Fostex PM0.4n) positioned on the table at 60 cm in front of the participant. Stimuli were empty auditory intervals, delimited by two marker sounds one at the onset and one at the offset. Markers were 16 ms tones, either with 1 kHz or 4 kHz frequency (according to the condition, see below), comprising a 2 ms ramp both at the onset and offset. Responses were collected using a standard keyboard.

Procedure. In each trial, two auditory intervals were presented. Subjects were instructed to indicate which one of the two intervals seemed to have a longer duration, pressing the appropriate key on a keyboard. Participants completed at least eight blocks of 80 trials. The standard interval had a duration of 300 ms and was always presented first. The probe interval ranged between 150 ms and 550 ms in duration and was always presented second. The two intervals were separated by an ISI of either 250 ms (short ISI) or 750 ms (long ISI), chosen according to the procedure used by Buonomano and colleagues (2009). Both ISI conditions were randomly interleaved within each block. The two ISI conditions were combined with other two conditions in a 2x2 experimental design. In the first part of the experiment, while the standard interval was always defined by two 1 kHz tones, the second interval could be marked either by 1 kHz ("same frequency" condition) or 4 kHz ("different frequency" condition) tones, with these two possibilities interleaved within each block. In the second part of the experiment, while the standard interval was defined by two markers having different frequency (1 kHz at the onset and 4 kHz at the offset), the markers of the second interval could either have the two frequencies in the same order (1 kHz onset/4 kHz offset, "same sequence" condition) or in the opposite order (4 kHz onset/1 kHz offset; "different sequence" condition).

Results

Figure 3.14 shows the results of Experiment 1. First, we analyzed participants' perceived duration (PSE) of the probe intervals (panels A and B). What we found is a systematic overestimation of the first interval especially with the short ISI, in most of the conditions (one-sample t-tests against the physical reference duration (300 ms): significant difference for both the conditions of the first part, t(11) = 2.759, p = 0.0186, t(11) = 3.945, p = 0.002 and for the same sequence condition, t(11) = 3.514, p = 0.004; no overestimation in the different sequence condition, t(11) = 1.131, p = 0.282;). A two-way ANOVA indicated an effect of the ISI on participants' PSEs (F(1,11) = 9.321, p = 0.011 and F(1,11) = 11.446, p = 0.006, respectively for the first and the second part of the experiment, while no significant differences were found among the different conditions: F(1,11) = 3.874, p = 0.075 and F(1,11) = 0.614, p = 0.450), and a multiple comparison procedure showed a statistically significant difference between PSEs in the short and long ISI conditions, in the same frequency, same sequence and different sequence conditions (t(11) = 3.033, p = 0.006; t(11) = 2.650, p = 0.015; t(11) = 2.477, p = 0.022, respectively), but not in the different frequency condition (t(11) = 1.843, p = 0.080), which, however, show a trend in the same direction.

Given that in most of the conditions tested we found a significant effect of ISI on perceived duration, to obtain a cleaner index of participants' precision in the task we chose to analyze JNDs normalized by duration estimates (PSEs) – that is, we computed the Weber Fractions (WFs). Such index, which takes into account variability and biases in perceived duration, could give us a more sensible measure of performance levels, discarding any artifact due variations

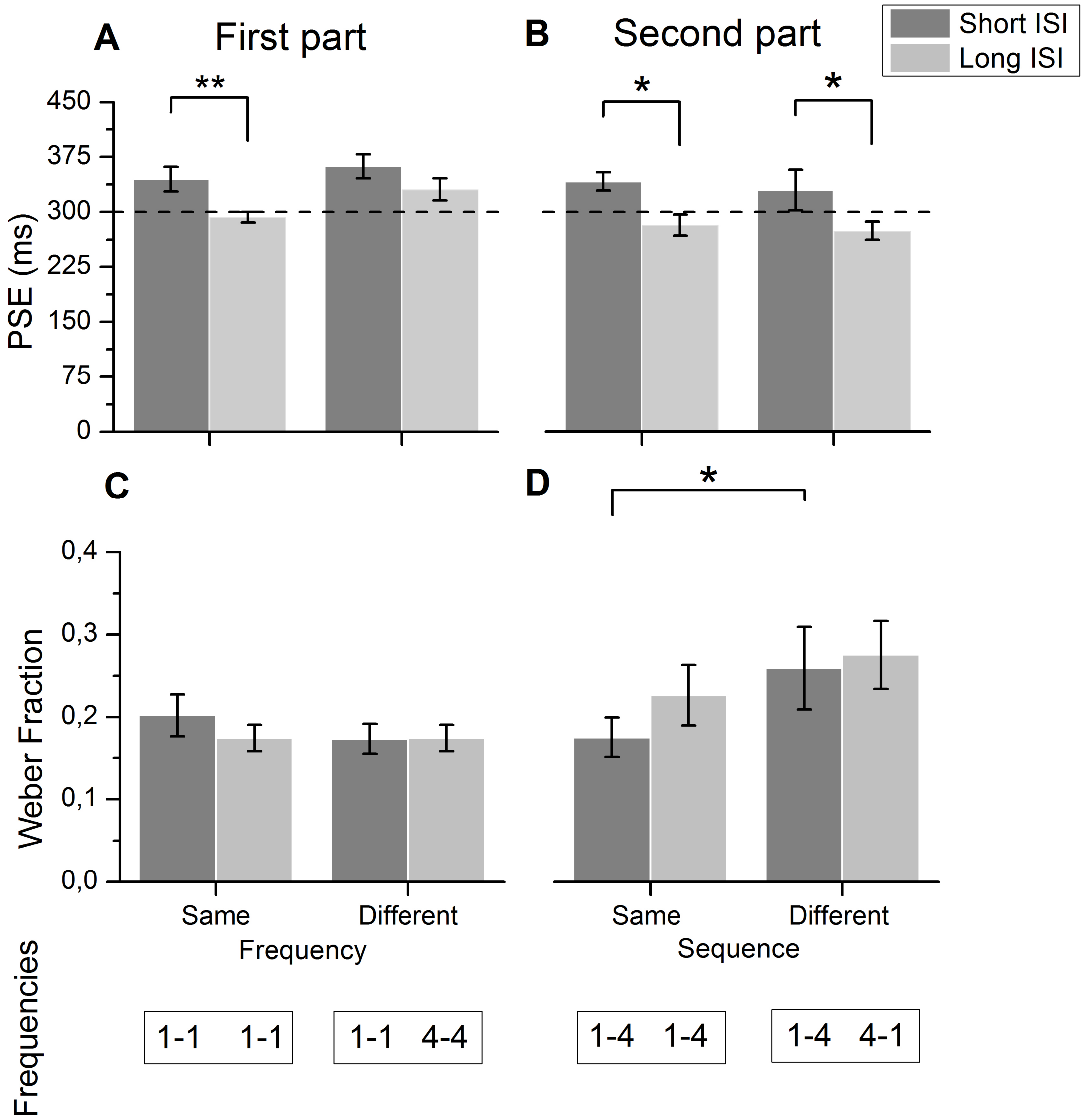

Figure 3.14: **Experiment 1 results.** (A and B) Duration of the probe interval that perceptually matches the 300ms standard (PSE). Results suggest a clear overestimation in the perceived duration of the first interval selectively for the short ISI condition, in three out of four conditions (same frequency, same sequence and different sequence). (C and D) Average participants' precision (WF) in the auditory duration discrimination task. (Bottom of the panel) Sequences of markers' frequencies in the different experimental conditions. Error bars represent the S.E.M. Asterisks represent significance of the multiple comparisons analysis: * $p < 0.05$, ** $p < 0.01$.

in accuracy. What we found is reported in Fig. 3.14C and D. In the first part of Experiment 1, WF measures were not significantly different neither comparing the different ISI, nor comparing the different conditions, (two way R.M. ANOVA, F(1,11) = 0.892, p = 0.365, for factor "ISI"; F(1,11) = 2.297, p = 0.158 for factor "condition"). Regarding the second part of the experiment, WFs reflected only a general effect of the sequence of stimuli on discrimination performances (two way R.M. ANOVA, F(1,11) = 7.628, p = 0.018, for factor "condition"; no significant differences in the factor "ISI", F(1,11) = 1.819, p = 0.205), which was mainly driven by a difference between performances with the short ISI in the same versus different sequence conditions (Holm-Sidak Multiple Comparison Procedure, t(11) = 2.768, p = 0.012).

3.3.3 Experiment 2 – visual duration discrimination

In the second experiment, we tested whether modifying the ISI has a similar influence on the discrimination of visual intervals as it does on the auditory intervals (at least for 100 ms intervals). We presented intervals marked by Gaussian blobs equiluminant to the background, in order to minimize the perception of apparent motion. To manipulate whether the two durations are processed by the same or by two different networks, here we presented stimuli either at one or two spatial locations, thus capitalizing on the segregation of visual receptive fields (Hubel & Wiesel, 1962) and localized processing of temporal properties revealed by spatially selective adaptation (Johnston et al., 2006; Burr et al., 2007; Fornaciai et al., 2014; Fornaciai et al., Under revision). The second part of the experiment follows the logic of the second part of Experiment 1: we presented the location of the markers either with a similar or a different sequence.

Materials and methods

Subjects. 12 subjects participated in the experiment (9 females, age ranging from 19 to 30 years), after giving an informed written consent. With the exception of two of the authors (M.F. and E.M.) who participated in the experiment, all the subjects were naive to purpose of the study, and were rewarded for their time with 6 GBP/hour. All the participants reported to have a normal or corrected-to-normal vision. Experimental procedures were conducted according to the protocol approved by the ethics committee of the

University of Birmingham, and are in line with the declaration of Helsinki.

Apparatus. The experiment was performed in a quiet dark room. Observers sat in front of a monitor screen at a distance of about 60 cm, with head stabilized by means of a chinrest. Stimuli were generated using the routines of the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for Matlab (version r2013a, The Mathworks), and presented on a NEC Multisync CRT monitor, with a resolution of 800x600 and a refresh rate of 60 Hz. Visual stimuli were green Gaussian blobs (sigma = 1 deg), presented for one monitor frame (16,6 ms) on a red background (luminance = 13.44 cd/m^2). The luminance of the green stimuli was chosen specifically for each subject according to the their perceived equiluminance with the red background (with red kept at a constant luminance level). Perceived equiluminance was determined before each experimental session by means of a flicker-fusion procedure (Livingstone & Hubel, 1987), and then the green channel luminance in the main experiment was adjusted according to the value that corresponded to the minimum impression of flicker (average matched luminance = 16.35 cd/m^2). Stimuli were presented in two possible positions: 5 deg to the left or to the right of the central fixation point, 6 deg above the center of the screen. Responses were collected by means of a standard keyboard.

Procedure. In each trial, a sequence of two intervals (each one bounded by two visual flashes) separated by an ISI, was presented, while participants kept the fixation on a small black circle at the center of the screen. The first interval (standard) was kept constant, with a duration of 300 ms, while the second one was varied from trial to trial, with durations ranging from 150 to 550 ms. As in the previous experiment, we tested two ISI values: 250 ms (short) and 750 ms (long). In the first part of Experiment 2, we presented the two intervals either in the same or in different position, according to a 2x2 design (ISIs and positions). All the combinations were randomly interleaved across each experimental session. In the second part of the experiment, each of the two durations was bounded by flashes presented in two different locations. The positions of the first two markers were kept always the same across each experimental session, while the markers of the second interval could be presented as either the same sequence of flashes as the first interval, or as an inverse sequence. ISIs and difference sequences were interleaved again according to a 2x2 design. Participants were instructed to indicate, at the end of each trial, which intervals seemed longer, pressing the appropriate key on a keyboard. Figure 2E shows a summary of the different conditions (positions and sequences) tested in Experiment 2.

Results

Fig 3.15 shows the results of the two conditions of Experiment 2. First, regarding PSEs (Fig. 3.15A and 3.15B), we found a pattern of results very similar to the ones obtained in the auditory experiment: in the short ISI conditions, participants systematically overestimated the first interval (one-sample t-test against the physical reference duration (300 ms): t(11) = 4.213, p = 0.001, t(11) = 5.459, $p < 0.001$, t(11) = 5.082, $p < 0.001$, t(11) = 4.142, p = 0.001). Even with the long ISI in the different position condition, PSEs were slightly, but significantly, overestimated compared to the reference physical duration (t(11) = 2.862, p = 0.015; no differences in the same position condition, t(11) = 2.121, p = 0.057, and in the second part of the experiment, t(11) = -0.733, p = 0.479 and t(11) = -1.102, p = 0.294). However, in all the conditions, PSEs with the short ISI were significantly different from the long ISI conditions (Holm-Sidak multiple comparisons: t(11) = 3.362, p = 0.003; t(11) = 2.469, p = 0.022; t(11) = 4.635, $p < 0.001$; t(11) = 5.394, $p < 0.001$; respectively for same and different positions and for same and different sequence).

As in Experiment 1, since we found robust ISI effects on participants' PSE, across all the conditions tested, we used Weber Fractions as measure of precision (Fig. 3.15C and D). Regarding the first part of the experiment, what we found is a significant effect of ISI on discrimination performances in the short ISI condition, (t(11) = 3.166, p = 0.005), when the markers of the two durations were presented in the same position, while no effect was evident in the different position condition (t(11) = 1.248, p = 0.227). Average Weber Fractions in the second part of the experiment (Fig. 3.15D), on the other hand, showed no effects of ISI. Instead, we found only a difference between the two conditions (same versus different sequence), especially with the short ISI (t(11) = 2.549, p = 0.019).

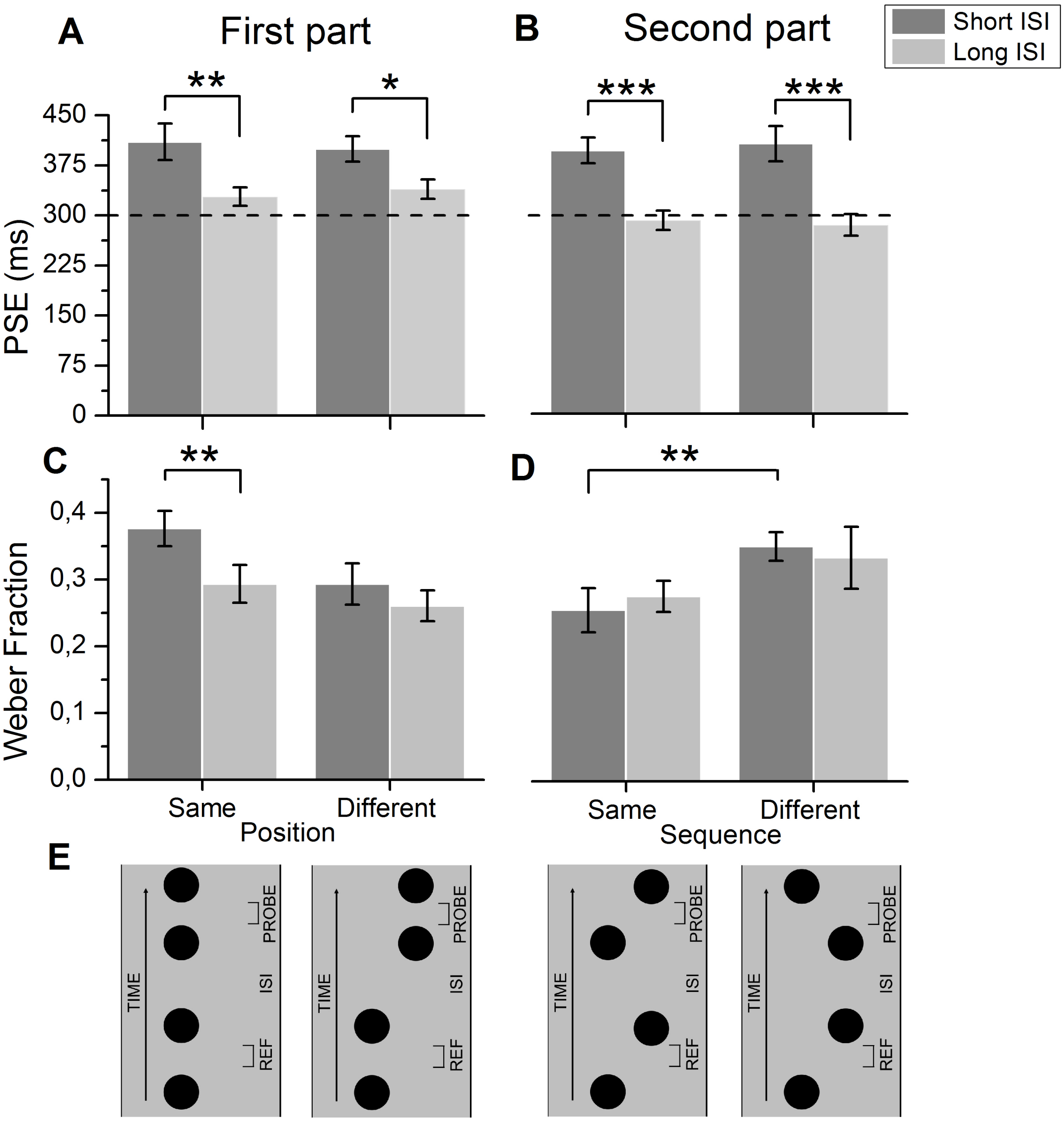

Figure 3.15: **Results of the visual duration discrimination experiment.** Participants' average perceived duration in the first (A) and second (B) part of the experiment. Similarly to the auditory experiment, we found a significant overestimation of the first interval specifically in the short ISI condition, irrespective of the position of the stimuli. (C and D) Average participants' performance (WF) in the two part of Experiment 2. (C) In the first part, in which visual intervals could be bounded by markers presented either in the same or different position, we found a significantly worse discrimination performances (higher WF) in the short ISI condition, only when stimuli were presented in the same location. (D) In the second part, in which we presented stimuli according to either the same or different sequences of positions, we only found a general increase in WFs in the latter case, more pronounced with the short ISI. (E) Experimental conditions in the visual experiment. From the left to the right: conditions in the first part of Experiment 2, where the marker of the two intervals could be presented either in the same spatial position, or with the second interval displaced in a different location; conditions in the second part of the experiment: the two intervals could be bound by visual flashes presented according to the same sequence of positions, or with the second interval presented according to a different sequence. Note that the pictures are displayed in black and white, but the actual colors were red (background) and green (stimuli). Error bars represent the S.E.M. Asterisks represent significance of the multiple comparisons analysis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

3.3.4 Experiment 3 – control: size discrimination.

According to Karmarkar & Buonomano (2007), the effect of ISI should be specific for time perception. Indeed, the authors also tested this effect in a frequency discrimination task, showing that judgments about other features of the stimuli are not impaired by a too short ISI. Here, according to this idea, we attempted to test whether this impairment is specific for time or whether it might be observed in a different task. To keep stimuli as similar as possible to the ones used in Experiment 2, we exploited a size discrimination task, using the same visual stimuli and presenting only the two middle markers.

Materials and methods

Subjects. 8 subjects participated in the experiment (5 females, age ranging from 18 to 29 years), after giving an informed written consent. With the exception of two of the authors who participated in the experiment, all the subjects were naive to purpose of the study, and were rewarded for their time with 6 GBP/hour. All the participants reported to have a normal or corrected-to-normal vision. Experimental procedures were conducted according to the protocol approved by the ethics committee of the University of Birmingham, and are in line with the declaration of Helsinki.

Apparatus. The experimental setup used for Experiment 3 was the same as Experiment 2.

Procedure. The general procedure of Experiment 3 was very similar to that of the first part of Experiment 2, except that in this case only two stimuli were presented (the two middle markers), and the size of the Gaussian envelope on the visual stimuli was modulated. While the size of the first stimulus (reference) was kept constant across the trials, with a width of the Gaussian envelope fixed at 1 deg, the second one (probe) was modulated in size from trial to trial, with width varying from 0.51 deg to 1.51 deg. Stimuli could be presented either in the same or in different positions (with the first one always on the left of the fixation point), and separated by either a short or long ISI. Participants were instructed to indicate, at the end of each trial, which stimulus seemed bigger in size, pressing the appropriate key on a keyboard.

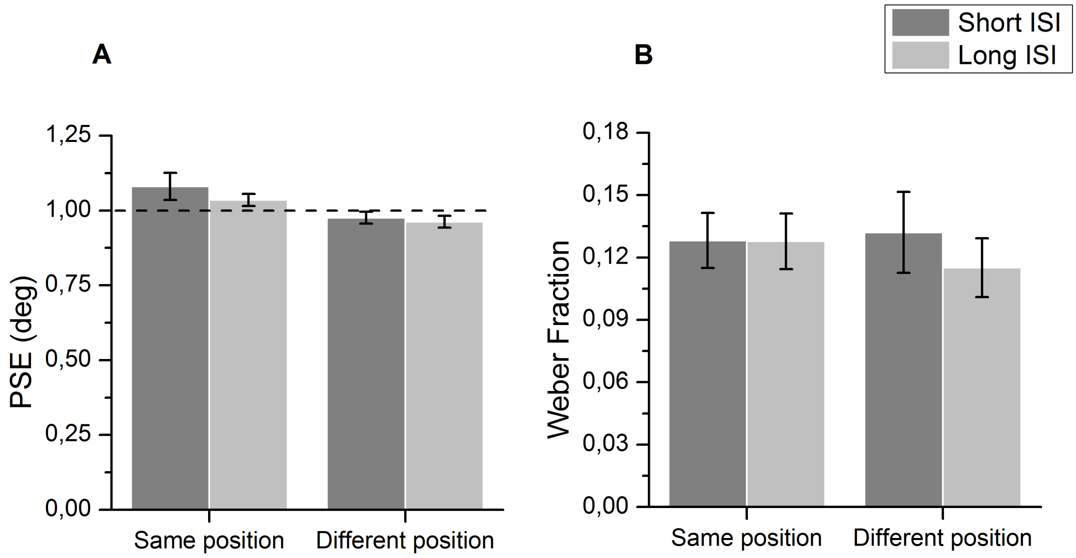

Figure 3.16: **Results of Experiment 3 concerning visual size discrimination.** Participants' average JNDs (A) and PSE (B) in the two part of Experiment 3. In this experiment we did not observe any systematic difference in participants' discrimination performances, neither in terms of precision nor in terms of accuracy. The dashed line marks the point of objective equality (physical reference size).

Results. As showed in Fig. 3.16, participants' performances, both in terms of precision (WFs, panel B) and perceived size (PSEs, panel A) are very similar across the different conditions tested. Namely, we did not observe any significant difference in WFs and PSEs between short and long ISI conditions, both for stimuli presented in the same or in different positions.

3.3.5 Discussion

While the large variety of clock-counter models (Creelman, 1962; Treisman, 1963) probably remains the most prominent frameworks in time perception research, recent theories tried to go beyond the idea of a centralized clock, suggesting the possibility that the perception of time might be deeply rooted into sensory processing, swapping the stopwatch metaphor with more biologically-plausible mechanisms (Miall, 1989; Matell & Meck, 2004; Mauk & Buonomano, 2004). Interestingly, Buonomano and colleagues, beside the several neural networks simulations supporting their theory, tested directly the prediction of the SDN model studying the effect of the inter-stimulus interval between two auditory durations on discrimination performances finding that precision in such task depends on the time between the two stimuli, thus supporting the idea that temporal information might be extracted from neuronal networks dynamic, which needs a proper "reset" time between two intervals, in order to return to their initial state.

However, other investigations further suggested a strict limitation of the SDN model, at least in the auditory modality, for very brief durations. Namely, Spencer and colleagues (2009), investigating the disruptive contextual effect on duration discrimination performances provided by distractor stimuli – which was previously showed by Karmarkar & Buonomano (2007) – found that such effect is no longer apparent when the base duration of the intervals is increased from 100 ms to 300 ms. While the effect found with the "reset task" in Karmarkar & Buonomano's paper provided an intriguing evidence supporting their framework, the later result constrained the model to only very brief durations.

Here we attempted to replicate Buonomano's results concerning the effect of ISI on auditory time perception, using a longer duration range, and then to extend their findings to the visual modality. If the constraints to the SDN model found by Spencer et al. (2009) extends beyond the specific contextual effect investigated in their study, so we should expect a similar patter of results testing the effect of ISI on duration discrimination performances with longer base durations. However, such limitation of the model might not be a generalized rule concerning the processing of durations in general, but it might depend on the relative resolution of the sensory modality tested. Indeed, previous investigations concerned mostly duration discrimination in the auditory modality – a sensory modality renown for its superior temporal resolution, compared to other (Hirsh & Sherrick, 1961; Welch & Warren, 1980; Morein-Zamir et al., 2003; Vroomen & de Gelder, 2004). So, we further tested the effect of ISI on visual duration discrimination, in which a different intrinsic temporal resolution might shift the constraints of the sensory interval processing and the effects predicted to the SDN model even to longer durations.

Results from Experiment 1, where we tested auditory duration discrimination, indeed suggested that increasing the base duration to 300 ms attenuated the effect of ISI on performances. Analyzing Weber Fractions as a measure of participants' precision in the duration discrimination task, we did not find any significant impairment in performances, neither with intervals bounded by sounds with the same frequency or with different frequencies. Such lack of effect closely resemble the results of, Spencer et al., (2009), who noted that increasing the base duration from 100 ms to 300 ms, weakened the distractor effect on auditory perceived duration. While these results, do not ultimately challenge the SDN model, they further suggest that the machinery for time processing in the SDN model might be exploited only for very brief intervals.

Similarly, in the second part of the auditory experiment, where we modulated the individual markers of each duration, we did not observe any systematic influence of ISI on duration discrimination performances, but only a general increase in WFs in the different sequence condition, which was particularly marked with the short ISI. Why does comparing the duration of two intervals marked by two opposite sequences of sounds was so difficult? On the one hand, it could be simply related to an increased task difficulty, pointed out by the increase in WF with both short and long ISI. On the other hand, another possibility concerns the perceptual effect of auditory stream segregation (Bregman & Campbell, 1971). According to this effect, a rapid succession of sounds with different frequencies could be perceived either as a single stream of sounds changing frequency, or as a superimposition of two auditory streams with different frequencies, depending on the frequency difference and on the rate of the stimuli (Snyder & Alain, 2007). Moreover, temporal order judgments about the temporal positions of two consecutive sounds are strongly impaired by the perception of two segregated streams, while performances are usually much better when only one stream is perceived (Warren et al., 1969; Bregman & Campbell, 1971). We did not ask participants to indicate whether the succession of stimuli appeared to resemble one or two streams, but it is possible that in the inverse sequence the two different sounds might be segregated in two streams, just like two rhythms superimposed, which could have lead to an impaired perception of sounds' temporal order and hence to a more difficult discrimination of the two durations, due to a higher temporal uncertainty about the boundary of the intervals (Rolke & Hofmann, 2007).

Interestingly, switching sensory modality and testing the ISI effect with visual intervals, markedly changed the results. Indeed, what we found seems overall in line with the predictions of the SDN model. In the first part of Experiment 2,in order to involve the very same network, or two different networks in the processing of the two durations, we presented the markers of the visual intervals in the same or in two different positions, according to the general idea that visual neurons – from early to high-level stages – process information through spatially-localized receptive fields, and so different position on the retina are likely processed by different networks (e.g. Hubel & Wiesel, 1959; Hubel & Wiesel, 1962). Moreover, it has also been shown that visual mechanisms involved in the processing of event time are spatially-localized (Johnston et al., 2006; Burr et al., 2007; Fornaciai et al., 2014; Fornaciai, Arrighi, Burr, Under review). What we found is that a short ISI (compared to the longer one) actually impaired discrimination performances, when stimuli were presented in the same position, but not when the markers of the two intervals occupied two different positions in space – which is consistent with the idea that sensory networks need a certain "reset" time between two durations. Importantly, we found such effect even using a duration range around 300 ms, which instead has proven to be free from such distortions in the auditory modality.

Regarding the second part of the visual experiment, what we found seems fairly in line with the auditory one: while presenting intervals bounded by markers presented in different positions (both according to a similar or difference sequence of positions) did not lead to effects strictly related to the ISI, we found a general decrease in temporal sensitivity when stimuli when defined by markers presented according to different sequences – with a more pronounced difference from the same sequence condition when the ISI was short. Possibly, such impairment could be only due to an increased difficulty in comparing intervals defined by stimuli with different features.

Why does a too short ISI impair visual duration discrimination even with a relatively long duration range? Possibly, the explanation for this difference might reside in the different sensory processing of that signals undergo in the two modalities: while auditory signals normally carry temporal information with a finer resolution, visual signals are processed with a relatively more sluggish process that makes the visual modality normally less sensitive to the temporal dimension – a feature that has been clearly demonstrated for instance by the temporal ventriloquist effect (Morein-Zamir et al., 2003; Vroomen & de Gelder, 2004). Hartcher-O'Brien et al (2014), have shown that the inferiority of visual duration judgments could be reversed by providing stimuli with low signal-to-noise ratio, making auditory duration judgments equally precise or even worse than vision ones. Such reversal shows that limitation in precision of temporal estimates depends on the noise of the output of sensory processing, not by an intrinsic higher resolution of the sensory modality considered. In light of this finding, the differential effect of ISI on auditory and visual stimuli thus suggests that the limits of temporal processing with normal signals depends on the relative temporal resolution of the outcome of sensory processing which is dictated by the noise of the sensory stimulus and, critically in normal stimulation conditions, by the dynamics of sensory processing, which in the SDN model is the speed with which the activation in the neuronal network changes as a function of the input. Sluggish responses lead to a more discriminable pattern of activation over long time course, whereas rapid changes make the representation more prone to cross talk and thus more limited in the range of representable duration. Our results is in line with this account showing that with normal stimuli auditory processing is more responsive and thus works for a range of duration that is shorter than vision.

As an alternative explanation for the effect in visual modality, another process that might lead to variation in precision in a timing task is attention, by means of effects such the Inhibition of Return (IOR) or the negative cueing effect provided by transient attention (Yeshurun & Levy, 2003). However, on the one hand, IOR was mostly tested in detection and reaction times paradigms (see Klein, 2000 for a review), while several studies failed to report an IOR effect when the task required discrimination (e.g. Terry et al., 1994; Tanaka & Shimojo, 1996). Lupiáñez et al. (1997) further reported that IOR effects could be observed also in discrimination tasks, but with a much slower time-course – which is not suitable to explain our results. On the other hand, several studies pointed out a specific effect of transient attention on temporal tasks: directing the transient attention by means of an exogenous cue might impair the performance in temporal order judgment tasks, due to a reduced temporal resolution – an effect that is thought to be related to a differential inhibition/facilitation of the magnocellular and parvocellular systems (Yeshurun & Sabo, 2012). In our experiment, the transient shift of attention triggered by the onset of a marker might have impaired the temporal resolution upon which the duration discrimination judgments are based. However, such effect has been reported to be consistently weaker when stimuli are isoluminant with the background (Yeshurun, 2004), and so it is not a likely candidate to explain our results.

Moreover, also Cicchini & Morrone (2009) investigated the relation between attentional modulation and duration discrimination performances, but while they found a robust duration underestimation when the markers of the test interval were presented in two different positions instead of only one, they did not find any difference in precision across all the conditions tested. These results further suggest that while attention might modulate perceived time, precision in a duration discrimination task does not systematically depend on it.

So, while Spencer and colleagues (2009) reported that attentional effects might alternatively explain their findings, this possibility seems less suitable for our results. Moreover, neither masking effects (see Enns & Di Lollo, 2000) nor local interactions between the markers of the first and second intervals, could account for the decreased sensibility to duration with the short ISI in the same position condition, as showed by the lack of effect in the different sequence condition of the visual experiment.

Besides the effect of ISI on discrimination performances, we also found a robust and systematic effect on perceived duration (PSE) in almost all the auditory and visual conditions, which was not previously reported. This pattern of results is likely due to difference in the procedures performed (we always presented the reference interval first), and it seems consistent with a systematic time-order effect (Jamieson & Petrusic, 1975). Indeed, time-order effects in duration perception are critically dependent on stimulus' length and ISI, and several studies reported positive time-order errors (overestimation of the first interval) especially with short ISIs, short durations and weak stimuli, while negative (overestimation of the second stimulus) errors are more common with longer stimuli and longer ISIs (Ellis, 1973; Hellström, 1985; Hellström, 2003; Hellström & Rammsayer, 2004; Hellström & Cederström, 2014).

Finally, our control condition concerning a different, non-temporal, dimension, such size perception, showed that the observed impairment in discrimination precision caused by a short ISI is a peculiar feature of time perception, since no effects have been observed in the visual size discrimination task (Experiment 3). This result is consistent with the findings of Karmarkar & Buonomano (2007) in the auditory domain, which showed that ISI has no effect on frequency discrimination performances.

3.3.6 Conclusion

Overall, our results support the idea of an intrinsic processing of temporal information directly at the level of sensory processing, as proposed by the State-Dependent Network model, especially for sub-second durations. Intriguingly, while on the one hand we provided other evidence that the effects predicted by SDN model are limited in the auditory modality to very short durations, in the range of 100 ms, on the other hand we showed that in the visual modality the ISI effect could be observed even with a longer duration range (300 ms), suggesting that the constraints of the SDN model might depend on the temporal resolution of the different sensory modalities.

Chapter 4

The perception of number

4.1 Introduction

We constantly deal with a very rich environment, in which we must select very rapidly the relevant information and ignore the irrelevant stimuli. Beside the pivotal importance of spatial information – in order to understand and navigate through the environment – and temporal information – fundamental for many behavioral processes, another crucial feature that characterize the sensory world is the number of objects in the visual scene. How many objects are out there? This question represents another fundamental step that the brain must accomplish to guarantee survival. Indeed, the ability to rapidly estimate the approximate amount of objects – for instance, the amount of fruit on a tree, or the amount of predators/preys in a field – is evident across a large number of animal species (Butterworth, 1999; Hauser et al., 2000; Pepperberg et al., 2006; Dehaene et al., 2008). This capacity represents an important evolutionary advantage, providing the ability to rapidly choose zones in the external environment with more food, more preys – and hence more chances to catch one of them – or to discriminate which group of competitors is more numerous.

Differently from other animal species, humans have the ability to use symbols to represent numbers, and to manipulate them in order to perform mathematical operations like sums, subtractions, multiplications and so on, as well as counting items using number words introduced by language. However, what humans share with most of the animals, is the ability to make a reasonably accurate estimate of how many objects are in a given visual scene, even without the possibility to count them – an ability that is evident also in newborn and infants well before the acquisition of linguistic and mathematical abilities (Whalen et al., 1999; Xu & Spelke, 2000; Coubart et al., 2014).

Such ability to represents in an imprecise (not exact like counting) fashion a given amount of items, seems allowed by a core "Approximate Number System" (ANS), that would be present as an innate ability from the birth, and might support the later acquisition of formal mathematical abilities (Feigenson et al., 2013; but see also Libertus et al., 2013).

However, besides the abilities supported by the ANS – which comes into play with relatively numerous sets of items – a different process seems involved with very small numbers, approximately up to four items: numbers in the so-called subitizing range are estimated very rapidly, and virtually without any error – an ability that anyway needs attentional resources to work (Burr et al., 2010; Burr et al., 2011).

When the sets of items – say, for instance, clouds of dots, which represent a classic stimulus used in numerosity perception experiments – contain more elements, the estimation is slower, and more prone to errors (Jevons, 1871; Kaufman & Lord, 1949). However, what happens when you add so many items to the set that they become individually undistinguishable? When the image becomes blurred as a sort of texture, is the number of items encoded anyway as a "number", or other mechanisms come into play to mediate the representation? Indeed, another useful parameter that defines highly cluttered images is the density of objects. Several authors claimed that there is no need to hypothesize a specific system for the processing of numbers, while the approximate amount of objects in a scene might be derived by other lower-level properties of the image, such texture density. However, this possibility is still subjects to debate, and other authors suggested the existence of specialized mechanisms to perceive numerosity.

4.1.1 Numerosity as a primary perceptual attribute

Some recent studies, particularly exploiting the adaptation technique suggested that the answer to the question "what is number?" is that it is actually a primary perceptual attribute, similarly to many other more obvious attribute like contrast, orientation, color and so on. To support this idea, it has been shown that numerosity is strongly susceptible to visual adaptation: adapting for several seconds to a large array of stimuli (usually black/white dots) and presenting subsequently a stimulus with a smaller number of elements, lead to a robust underestimation of such stimulus (Burr & Ross, 2008). Conversely, adapting to a small amount of items cause the subsequent viewed stimulus to be perceived as significantly more numerous. Adaptation represents a very common form, throughout the visual processing pathways, of experience dependent-plasticity, in which our percepts are "recalibrated" according to the recent history of stimulation, in order to tune the sensory systems to the recent sequence of stimuli and optimize the use of the limited resources of the system (Boynton, 2004; Kohn, 2007; Solomon & Kohn, 2014). This observation, together with the fact that estimation of numbers approximately follows the Weber's Law – that is, thresholds increase proportionally as the number of items increases – led Burr & Ross (2008) to propose that numbers might be a primary visual attribute. What does it mean? It means that across a large range of conditions, sets of items are perceived and experienced as a particular amount of items, irrespective of other intervening and related features.

However, several authors challenged the idea of numerosity as a primary perceptual attribute. First, Durgin (2008) proposed that the adaptation effect used as evidence of a specific mechanism for number occurred via the mechanisms for texture density. Indeed, a number of studies have shown clear adaptation effects on texture: after adaptation to a coarse texture, the subsequent image appear to be finer, and vice versa (Blakemore & Campbell, 1969; Durgin, 1995) – an effect that has been demonstrated also with different kinds of texture, such "sandpaper" textures (Anstis, 1974) or random-dot patterns (Durgin, 1995). Moreover, it has also been shown that varying the area of the patch over which the elements are presented – and so dissociating the actual number of elements from their density on units of area – causes adaptation effects that are determined by the numerosity in a given area, instead of the total amount of items in the field (Durgin, 2008). This observation, together with the effects of texture adaptation, supported the idea of an adaption mechanism mediated by the low level information about the density of elements.

However, regarding the latter observation, we need to consider that adaptation is spatially selective. Otherwise, the effect of the adapted and the un-adapted stimulus used as a comparison in the typical discrimination paradigm would be both distorted by adaptation, and hence no net effect would be discernible. So while Durgin's (2008) results confirmed this spatial selectivity, they did not point specifically toward a texture-density mechanism.

Besides Durgin, other authors have championed the idea of a texture-density mechanism mediating numerosity perception (Dakin et al., 2011; Tibber & Dakin, 2012; Morgan et al., 2014). However, several experimental observations seem to further support the idea of numerosity as a primary perceptual system – not excluding the involvement of a texture density mechanism, but defining the conditions under which the system switches from one regime to the other.

4.1.2 Serially-presented numerosities, cross-modal and cross-format adaptation effects

Recently, Arrighi and colleagues (2014) investigated the processes of numerosity adaptation from a different point of view. Instead of the classic clouds of dots used to test perceived numerosity, in which several items are presented simultaneously, they devised a paradigm where numbers were presented serially, as a series of briefly presented disks, asking participants to estimate their number: instead of manipulating the spatial dimension of the stimuli, they manipulated the temporal dimension. What they found with such stimuli is strikingly similar to the previous observation: adapting to a slow presentation rate (2 Hz – which is the counterpart of the adaptation to a small set of items) caused an overestimation of the number of disks subsequently presented. On the other hand, adapting to a fast sequence (8 Hz – a large amount of stimuli), caused an underestimation of the subsequent stimuli (Fig. 4.1). Moreover, these adaptation effects were again spatially selective, affecting only the series of stimuli presented in the adapted position, further suggesting a purely perceptual nature for the adaptation effect, rather than cognitive.

Arrighi et al. (2014) also tested the cross-modal effect of adaptation: using auditory and visual stimuli. They showed that adapting to a stream of sounds could change the apparent numerosity of a stream of visual flashes

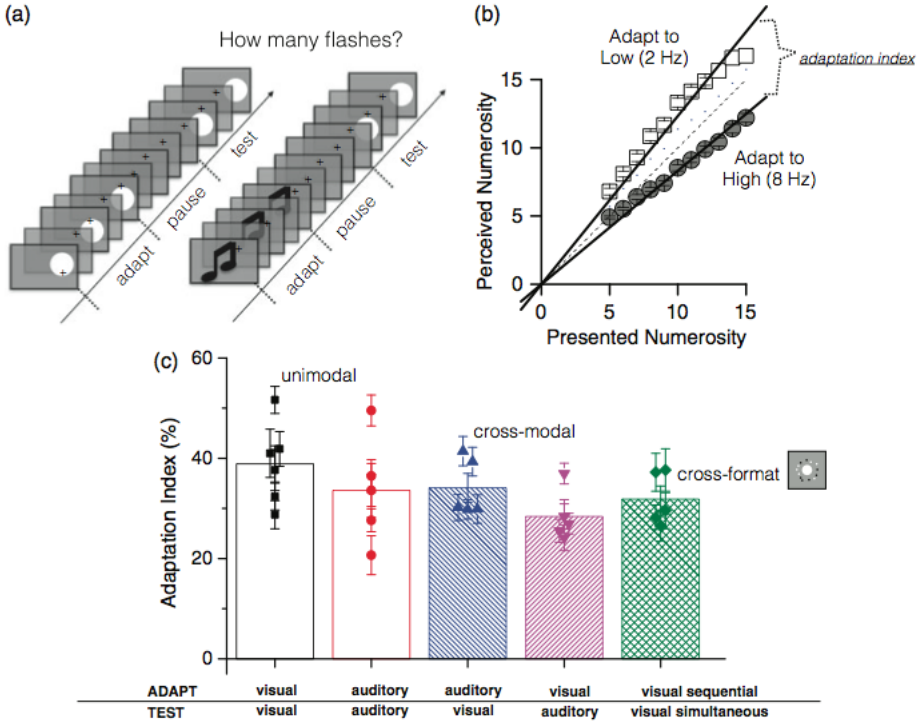

Figure 4.1: **Cross-modal, cross-format numerosity adaptation.** (a) Examples of stimuli and procedure used for the serial number adaptation; (left) visual uni-modal; (right) auditory-visual cross modal. (b) Sample results for the visual uni-modal adaptation task. Adapting to low numbers (2 Hz; open squares symbols) produced an overestimation of numerosity, while adaptation to high numbers (8 Hz; gray circles symbols) produced an underestimation. Data were well-fitted with a linear regressions (represented with lines) constrained to pass throughout zero. The strength of the adaptation effect (adaptation index) was calculated as the difference in slope of the regression lines. (c) Mean adaptation indexes for the different experimental conditions: uni-modal visual and auditory adaptation; cross-modal auditory-visual and visual-auditory adaptation; "cross-format" adaptation (adaptation to a serial presentation, test with simultaneously presented sets of stimuli). Bars show average data, symbols represent single subject data. Adapted from Arrighi et al. (2014).

and, conversely, adapting to visual flashes changed the apparent numerosity of a series of sounds, with the magnitude of the effect completely comparable to the previous within-modality experiment. Finally, Arrighi and colleagues went on to show that adaptation occurs independently of the format used to display numbers. They tested the effect of adaptation to a stream of flashes on "classic", simultaneously presented, spatial arrays of dots, finding again an effect of under- and over-estimation (depending on the relative numerosity) of about the same extent of the previous experiments (Fig. 4.1c).

Overall, these results strongly suggest that the perception of numerosity is indeed a capacity completely unrelated to texture perception; texture-density processes, indeed, cannot explain neither the effect of serially-presented stimuli, nor the cross-modal and cross-format effects.

4.1.3 The difference between numerosity and texture-density regimes

Recent psychophysical and imaging evidence further supported the dissociation between processes for the perception of number and texture-density, showing that number is a perceptual attribute quite distinct from texture. One of the clearest demonstrations of such dissociation between the perception of the number of items and texture perception was the so-called connectedness effect, discovered independently in the same year by He and colleagues (2009) and Franconeri and colleagues (2009). They found that connecting pairs of dots in random-dot patches with task irrelevant lines caused a systematic underestimation of the number of elements: particularly with brief presentation times, connected dots patches give a robust impression of a lower numerosity compared to a similar pattern of un-connected dots with the same numerosity (Fig. 4.2). According to He et al.'s data, each connection (they tested one or two connected pairs) had the same effect as removing one dot – just as each pair had been segregated and "counted" as a single object.

What is the first and most direct difference caused by adding lines to the dots patches? Adding the lines should increase the amount of "stuff" in the area (and hence the density), which, in turn, if the number of items would be calculated by means of a density mechanism, should increase the perceived numerosity, instead of decreasing it. So, overall, such results pointed out that what actually counts for perceived numerosity are the items identified as perceptual objects, not the total amount of "stuff" as should be expected by a texture-based mechanism.

Moreover, another study further investigated whether such connectedness effect could be generalized to any task (i.e. also motor tasks), or could be limited to perceptual judgments (Milne et al., 2013). In this study, the authors presented four dot-patches, either connected or isolated, and asked participants

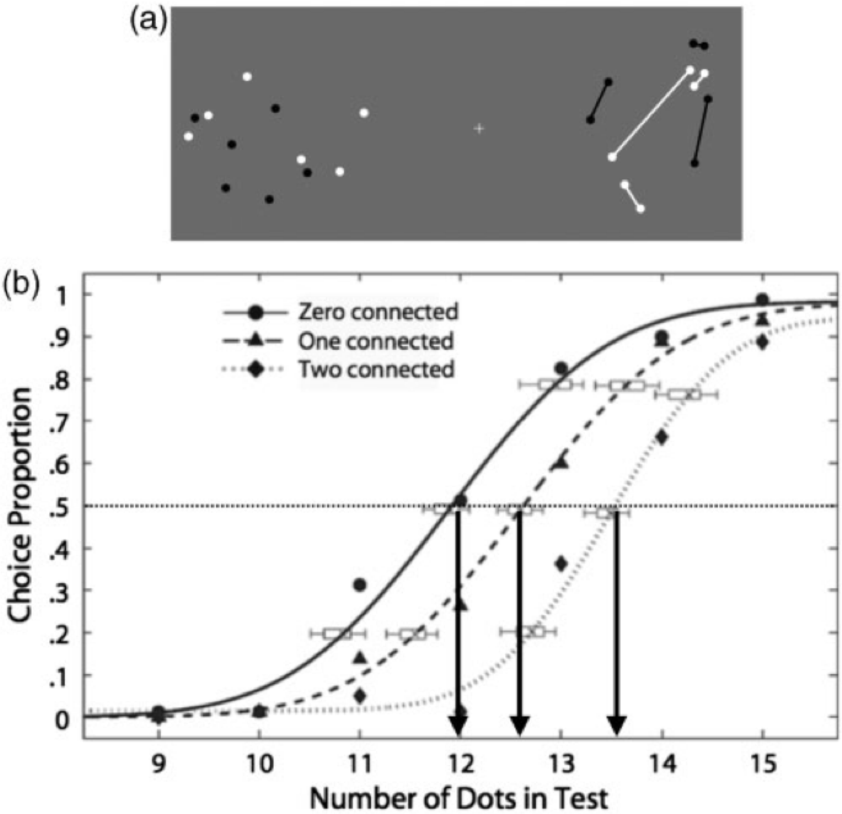

Figure 4.2: **Effect of connectedness on perceived numerosity.** (a) Example of stimuli configuration illustrating the effect of connectedness on the apparent number of elements – the patch on the right appear to be less numerous (adapted from Anobile et al., 2015b). (b) Example of psychometric functions of one of the subject tested by He and colleagues (2009), in which the proportion of choices "more numerous" (referred to the test stimulus compared to the standard stimulus containing 12 unconnected dots). As showed by the three functions, each corresponding to a different level of connectedness, connecting one or two pairs of dots cause a leftward shift of the psychometric, indicating that the numerosity of the test patch has to be increased to perceive it as numerous of the unconnected standard stimulus – a result that indicates underestimation of the patch containing the connected dots. Adapted from He et al., 2009.

either to judge them or make an arm movement in order to reach them. Analyzing perceptual judgments and the dynamics of reaching movements, they found that whilst numerical estimation was subject to the previously found underestimation of the number of elements, the dynamics of motor actions was completely unaffected by the connectedness effect. Such lack of effect for motor actions renders unlikely that connectedness might affect perceived numbers at the low-level stages of the visual hierarchy, while the hypothesis of a higher-level locus appear to be more appropriate for this effect.

Other results showing the influence of item arrangement on numerosity estimation suggested that symmetry could also affect perceived numerosity. Apthorp & Bell (2015) showed that a symmetric arrangement of dots cause a significant underestimation of their numerosity – a result that appear to be closely related to the connectedness literature. Indeed, even if it is not completely clear, and not as straightforward as the simple connection of pairs of dots, the symmetric patterns of dots were underestimated possibly because they appear to be more "connected" compared to random-dot patches, with the dots in symmetric positions tended to be perceptually joined.

But if numerosity is actually a primary perceptual attribute, exactly what attribute of a stimulus represent the primary property taken into account in determining numbers? He and colleagues (2015) investigated in depth this issue, exploring the conditions modulating the apparent numerosity of a set of items (dots). Across a range of presentation durations and numerosities, and with different tasks, they showed that not only connecting the dots led to underestimation, but also enclosing pairs of items with arbitrary and irregular forms led to the same effect. On the other hand, other kinds of grouping such color similarity did not lead to any change in perceived numerosity. The authors concluded that the primitive units counted in numerosity perception are strongly influenced by topological invariants, such connectivity and segmentation provided by the inside/outside relationship. Moreover, besides these psychophysical results, they also measured brain activity using functional MRI, showing that the representation of number in the intraparietal sulcus is strongly influenced by the topology of the stimuli (see next Section for further discussion on the neural underpinnings of the number sense).

However, numerosity and texture-density mechanisms are not mutually exclusive, and between the extreme conditions where the elements can be individually identified and counted, and where the elements are arranged as a uniform and blurred texture, there are many conditions where the difference between number and density could be less immediate, and so whether the judgment is based on one system or on the other might depend on other image characteristic. To investigate the transition from numerosity to texture-density, Anobile

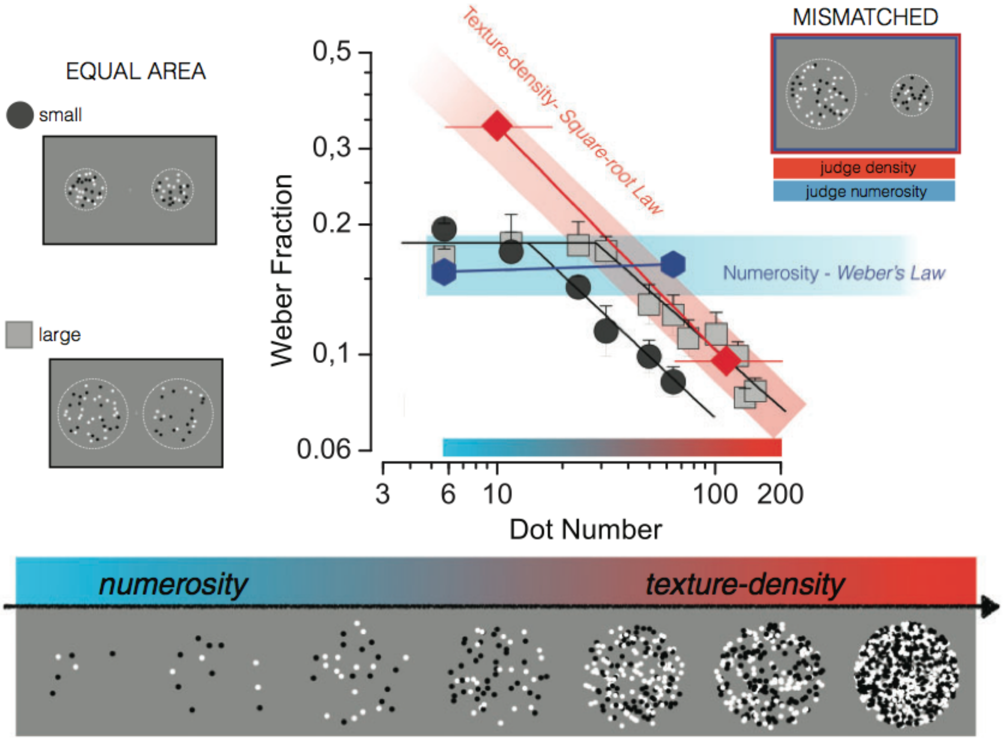

Figure 4.3: **Different sensory thresholds for numerosity and texture-density discrimination.** The graph shows Weber fractions as function as the numerosity of the stimuli, for two different areas of test and probe patches. Dark gray circles indicate the case where test and probe stimuli were confined in a small area (8 deg), whilst light gray squares indicate the case where the area of the stimuli was larger (14 deg). Blue hexagons represent thresholds for numerosity discrimination using stimuli with unequal area; red diamonds represents thresholds for density discrimination using stimuli with unequal area. Adapted form Anobile et al., 2014.

and colleagues (2014) systematically measured Weber fraction – an index of subjective precision in a given perceptual task – over a large range of conditions, with different numerosities and densities. The usual assumption is that Weber fractions for numerical estimates should be constant: indeed, according to the Weber's Law, as the numerosity increases, the estimation error should increase proportionally – leaving their ratio (i.e. the Weber fraction) constant. However, using different stimulus areas (150 deg^2 versus 54 deg^2), they found a different pattern of results: for the lower range of numerosity tested, Weber fractions were constant, but for higher numerosities, after a critical point, they started to decrease with the square root of numerosity. Interestingly, the critical point in which Weber's law was replaced with a square-root law depended on the area of the stimulus: the critical numerosity was lower for the smaller patch, while higher for the larger area – but the density at the turning point was the same (0.3 dots/deg^2). These results led the authors to propose the existence of two different regimes, one for number, which follows the Weber's law, and the other for texture-density, following a square-root law (Fig. 4.3).

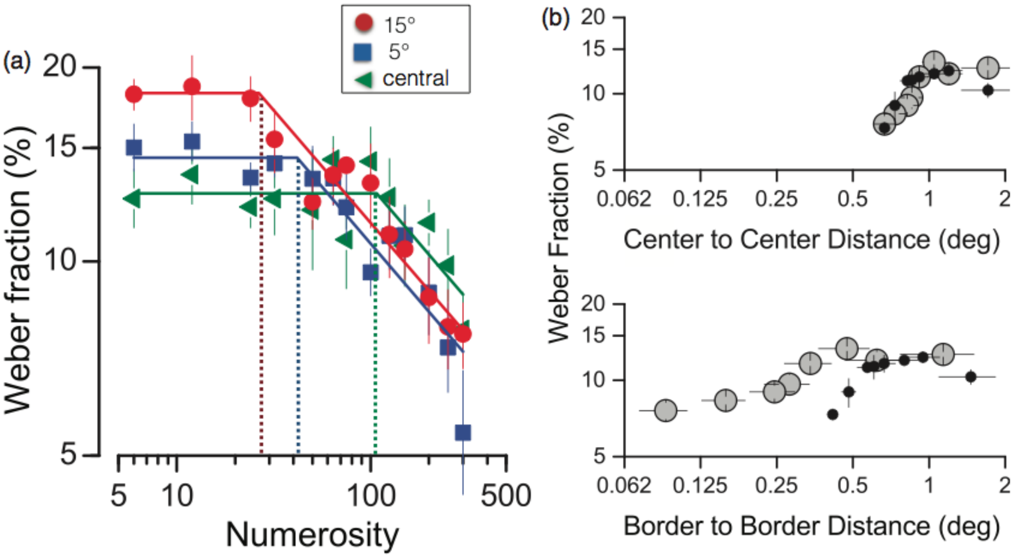

Figure 4.4: **Critical transition point between numerosity and texture-density regimes.** Weber Fractions for numerosity discrimination of dots patches centered at three different eccentricities (15 deg red circles, 5 deg blue squares, central presentation, green triangles). The critical transition points between numerosity (described by the Weber's Law, flat zone) and texture-density regime (described by a square-root law, descending zone) depended on eccentricity (dotted lines). (b) Weber fractions for numerosity plotted as a function of average center-to-center distance between dots (upper panel) or plotted as function of the average border-to-border distance (bottom panel). Large gray circles refer to stimuli patches comprising large dots (diameter 0.58 deg), small black filled symbols refers to small dots (diameter 0.25 deg). Adapted from Anobile et al., 2015a.

In a subsequent study, Anobile and colleagues (2015a) further investigated the mechanism that determines the transition from numerosity to density regimes. According to the authors, the transition resembles in some ways the phenomenon of *crowding*: the deleterious averaging of cluttered stimuli that impairs the recognition of single items in the ensemble (Bouma, 1970; Parkes et al., 2001; Pelli, 2008; Levi, 2008). Similarly, when items are too packed within a given area, they cannot be easily segregated as single items, and hence they become uncountable. Under these circumstances, texture-density mechanisms might come into play. According to this idea, the transition between the two regimes should depend on eccentricity, with the critical numerosity needed to switch to density decreasing as the eccentricity increases. Moreover, the transition to texture-density should depend on the center-to-center rather that edge-to-edge spacing. Indeed, what Anobile and colleagues (2015a) found, testing stimuli located at different eccentricity, is that the critical numerosity systematically decreased as the eccentricity increased (Fig. 4.4 (a)). Moreover, varying the size of the individual stimuli, they further observed that the transition depended clearly on the center-to-center separation, and not on the edge-to-edge spacing (Fig. 4.4 (b) and (c)).

While these results clearly point out some similarity between crowding and the switching from number to texture regime, the two phenomena present some differences. For instance, one major difference concern the paradigm typically used in studies investigating crowding: in such paradigm, participants usually have to identify a given feature of a target stimulus (i.e. form, orientation), whereas in numerosity task participants have to estimate/discriminate the amount of items – which do not require to identify the objects to be enumerated. Indeed, despite the similarity between the two processes, Anobile et al.'s data also show that the limits governing the transition are of a finer spatial scale compared to the spatial limits of visual crowding.

4.1.4 Neural mechanisms and substrates for numerosity processing

As discussed above, many psychophysical results obtained across several studies clearly point toward the existence of a dedicated mechanism for the perception of numbers and quantities – that is, a "number sense" – which seems different and separated from texture perception. But what are the neural mechanisms supporting this ability? What are the neural substrates of the number sense? Since the possibility to discriminate between different numerical quantities is not a uniquely human ability – rather, it is a crucial capacity for the survival, common across many animal species – several studies investigated the neural underpinnings and the mechanisms of numerosity perception in animals, and particularly in primates.

Evidence suggesting the existence of "number neurons", were first provided by Nieder and colleagues (Nieder et al., 2002; Nieder et al., 2006) testing monkeys with a delayed match-to-sample paradigm (i.e. memorize and compare the numerical magnitude of a sample stimulus) and recording cells activity in some key brain areas. What they found is that cells in the intraparietal sulcus (IPS) and the lateral prefrontal cortex (lPFC), behave as "number neurons": the cells responded strongly when presented (in their receptive field) with their preferred number, while their firing rate decreased with the numerical distance of the current stimulus from the preferred number – a dynamics that showed a clear tuning function for numerosity. Neurons in the parietal cortex moreover showed a shorter response onset (by about 30 ms) compared to neurons in the prefrontal cortex – a difference that suggests a hierarchical organization of the processing stream.

Importantly, neurophysiological studies concerning the mechanisms of numerosity perception in animals put great effort into controlling for any possible feature that might confound actual numerosity, such as density, the area covered by the elements, spatial arrangement and shape. Anyway, even controlling and discarding these factors or randomizing them, cells continued to show a clear tuning for the numbers of items presented, without any apparent influence from density or covered area.

Might this tuning for numbers be dependent on training, rather than being an innate function? Some authors indeed proposed such possibility (Roitman et al., 2012), supported by the fact that parietal neurons are particularly susceptible to learning and training (Freedman & Assad, 2006; Roy et al., 2010). However, the selectivity and tuning for numbers in parietal neurons has been demonstrated even testing monkeys that have never previously participated in number discrimination studies (Viswanathan & Nieder, 2013).

While neurons in IPS and lPFC showed selectivity for numbers, other investigations pointed out other kind of neurons that encode numerosity differently. For instance, Roitman and colleagues (2007) found evidence for the existence of another class of neurons in the lateral intraparietal (LIP) area in the monkey brain, which shows a gradient response to number. Such neurons were found to be divided in two types, preferring either low or high numerosities, and encoding numbers monotonically with a response gradient (i.e. increasing or decreasing their activity as function of the amount of items presented). Also for this class of neurons, responses were not defined by other visual attribute of the stimuli, such area or density, but their responses remained strictly related to numerosity. Interestingly, in the hierarchical processing stream, these neurons could be thought as "integrator", providing inputs to the number-selective cells in IPS and lPFC. Moreover, according to some authors (Dehaene, 2009), such neurons might be the mediator of adaptation effect on apparent numerosity. First, their spatially well-defined receptive field can account for the spatial specificity of numerosity adaptation; second, their gradient response to numbers seems consistent with the fact that apparent numerosity monotonically depend on the adaptor numerosity (Burr & Ross, 2008).

Regarding numerosity conceived as a series of events or objects presented along the temporal dimension (as investigated by Arrighi et al., 2014; see paragraph 4.1.2), Nieder and colleagues (2012) further reported another class of "number-neurons" in monkey IPS, which show selectivity for the number of items presented over time. These neurons represent a very likely candidate as the neural underpinning of the adaptation effects reported by Arrighi and colleagues (2014) of numerosity displayed over time.

The studies discussed above investigated the neural underpinnings of numerosity perception in animal, finding evidence of specialized neuronal mechanisms for the encoding of numerical magnitudes and quantities. But what about humans? Several recent studies investigated the existence of similar mechanisms also in human subjects, exploiting functional magnetic resonance imaging (fMRI). Piazza and colleagues (2004), using a fMRI adaptation paradigm, revealed tuning curves for approximate numbers in the human intraparietal sulcus (hIPS), while no selective tuning for numerosity were found in other areas like V1. In a subsequent study, Piazza et al. (2007) further showed that number encoding in human IPS and frontal cortex is notation-independent, finding cross-format habituation with non-symbolic numerosities (i.e. the number of items) and number symbols. Moreover, they also noted that the neural tuning for numbers – measured by means of BOLD (Blood Oxygenation Level Dependent) responses – depends on perceived, and not physical, numerosity. Indeed, exploiting the reduction in perceived numerosity caused by connections between pairs of dots (an effect that dissociate the physical from the apparent amount of elements; see Paragraph 4.1.3), they showed that the minimum change in IPS activity due to adaptation is reached when the perceived numerosity of adaptors and deviants stimuli is most similar. So, this observation further suggests that IPS is actually tuned to the perceptual dimension of numerosity, and not on other low-level properties of the stimuli.

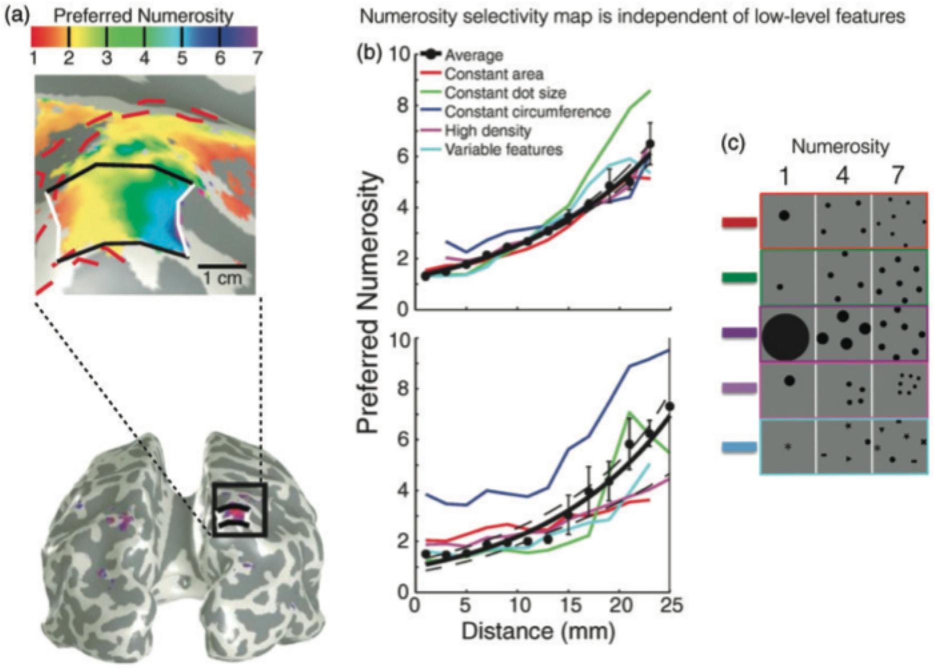

Figure 4.5: **Topographic representation of numerical magnitudes in human parietal cortex.** (a) Representation of the spatially organized numerosity map for small numbers (from 1 to 7) in human parietal cortex. The color code indicates the different parietal zones activated preferentially by a given numerosity. (b) Progression of numerosity selectivity as function of the distance from the markers (white lines) reported in panel (a) of this figure. Different colors represent the different conditions where the low-level properties of the stimuli were controlled for (the color code is explained in panel (c)). Adapted from Harvey et al., 2013.

To study with more detail the organization of number neurons in human IPS, Harvey and colleagues (2013) used high-field 7T fMRI to investigate the topographical organization of number-selective neurons. Interestingly, they found a columnar arrangement for numbers – just like lower-level areas, such V1, show a columnar organization for orientation tuning. Along this topographical organization, neurons responded to their preferred number (Fig. 4.5), while V1 neuronal activity was clearly driven by the absolute magnitude: the higher the number of stimuli, the stronger the response – a result showing again that hIPS responses are not driven by the contrast energy of the stimulus or other low-level attributes.

4.1.5 Conclusion and outline of the next section

Overall, these results from neurophysiological and human imaging, together with the psychophysical evidence discussed in the previous paragraphs, points towards the existence of a specific circuitry for the encoding of numerical magnitudes, while little evidence supports the idea of a more low-level mechanism for numerosity, based on texture-density.

In the next section, we will illustrate and discuss an experiment concerning the effect of numerosity adaptation on connected-dots patches, aimed to further disentangle the debate between numerosity and texture-density accounts. Indeed, the connectedness effect allow for a dissociation of perceived and physical numerosity. So, investigating whether adaptation might tap on the physical or perceived number of elements might provide further evidence about the level in the visual hierarchy of the neural underpinnings of numerosity perception mechanisms.

4.2 Number adaptation operates on perceived, not physical, numerosity

4.2.1 Introduction

In this study we test whether adaptation acts upon perceived or physical number. We measure the effect of adapting to 20 dots, then testing with patches of the same numerosity, either in isolation or connected pairwise. Our results show that the adapter had no effect on the numerosity of unconnected dots, but robustly reduced that of pair-wise connected dots. This shows that adaptation operates on mechanisms for numerosity, rather than more basic visual features, like the number of dots.

4.2.2 Materials and methods

Subjects

Seven subjects (two authors and five subjects naïve to the purpose of the experiment) participated in all experiments. All subjects had normal or corrected-to-normal visual acuity, and gave an informed written consent. Experimental procedures were approved by the Tuscan ethics committee and are in line with the declaration of Helsinki.

Apparatus and stimuli

The experiment was performed in a quiet and dimly illuminated room. Subjects sat in front of a 23-inch LCD monitor (mod. Acer S231HL) subtending 51 x 29 degrees of visual angle, at a viewing distance of 57 cm. Stimuli were generated with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) for MatLab (MatLab 2010b, The Mathworks, Inc.). All stimuli were patches of random-dots, presented within a circular window of 12°diameter. Each dot was 0.4°diameter, randomly black or white. Dot positions were chosen randomly from trial to trial (for adaptors and reference stimuli), respecting the conditions that two dots could not be separated by less than 0.75°. Stimuli with connected elements were calculated offline from a standard dot pattern (generated as described above) by joining iteratively dots with their closest neighbour (minimum line length 0.75°). If any line crossed another or encroached within 0.75°of another dot, the lines were discarded and regenerated iteratively until an acceptable pattern was created. If no uncrossed line combination was possible (which occurred with less than 1% of stimuli), a fresh dot-pattern was drawn and the procedure recommenced.

Adaptors comprised 20 isolated dots, like the unconnected stimuli. Probe stimuli appeared in the same position of the adapter, reference stimuli appeared on the opposite side of the screen and were varied from trial to trial, following a QUEST routine (Watson & Pelli, 1983) homing in on the point of subjective equality of the numerosity of the adapted probe patch, with an added Gaussian jitter of 0.15 log-units. The final estimate of PSE was taken as the median of the best-fitting cumulative Gaussian function to all the data of a particular condition (percentage “more elements than” against test physical numerosity).

Probe stimuli were of three types (Fig 4.7a), depending on condition: (a) 20 unconnected dots; (b) 10 pairs of connected dots; (c) Unconnected dots with numerosity chosen for each subject to appear equal to the perceived numerosity of the connected patch.

General procedure

Trials started with subjects fixating at a small red dot in screen centre. The adaptor stimulus was centred 12.7°left or right of fixation (varying randomly between session), presented for 20 s in the first trial of each experimental session, and for 6 s in subsequent trials to top-up the adaptation. Adaptors were followed by a 500 ms pause, and then probe at the same position of the adapter, together with the reference stimulus at an equal distance on the other side of fixation, were presented for 150 ms. At the end of each trial, subjects indicated which stimulus appeared to contain more elements (guessing if unsure) by pressing the appropriate key. For each subject and for each condition, we first performed a baseline measurement without adaptation. Each subject completed at least two blocks of 40 trials for each experimental condition. The different adaptation conditions were separated by breaks of at least 40-50 minutes.

4.2.3 Results

Figure 4.6 reports average psychometric functions and raw data (obtained pooling the data from the entire group of subjects), plotting proportion of trials in which the test stimulus was judged as more numerous as function of test stimulus numerosities in four conditions: isolated dots baseline and adaptation, connected baseline and adaptation. Results of the isolated dot condition show that numerical estimates after 20-dots adaptation (red diamonds) do not differ significantly from the baseline perceived numerosity (dark red circles), with both centred at approximately 20 dots – the physical reference numerosity, indicating that adaptation does not affect a test stimulus with the same numerosity. Results of the connected-dots condition, on the other hand, show that the perceived numerosity of dots joined by lines is systematically underes-

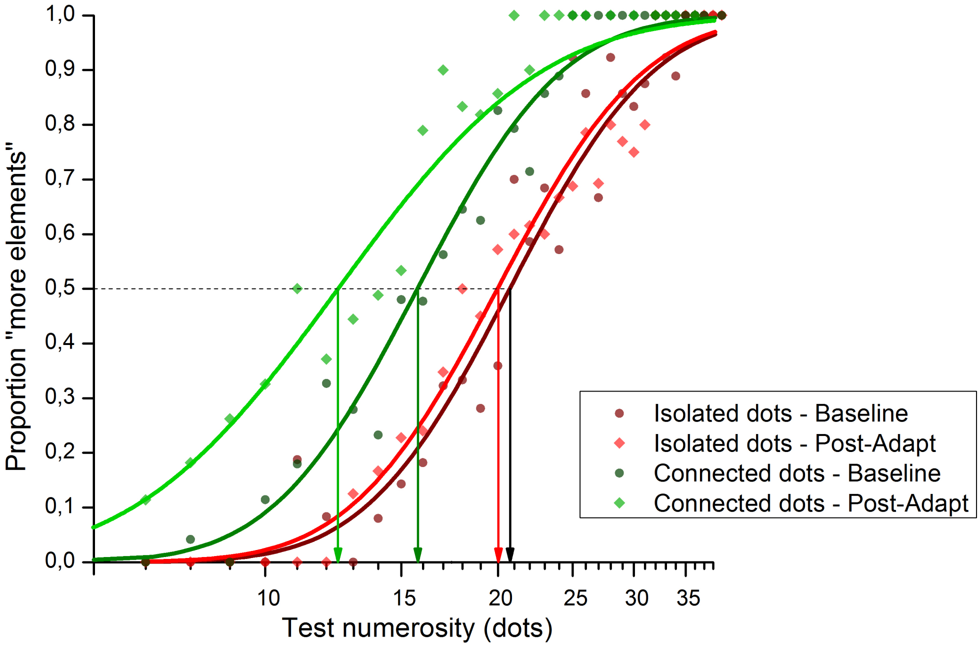

Figure 4.6: **Psychometric curves for isolated and connected dots numerosity discrimination.** Average psychometric functions obtained pooling data of all the participants, showing the proportion of trials in which the test stimulus appeared more numerous than the reference. Dark red and dark green curves refer to the baseline conditions either with isolated dots (dark red) or with dots connected by lines (dark green). Light red and light green curves refer to the post-adaptation performances, for the isolated and connected dots conditions, respectively. The results indicate that while 20 dots adaptation has almost no effect on the perceived numerosity of 20 unconnected dots (dark red and light red curves almost superimposed, and centred on about the physical reference numerosity), connecting the dots cause a robust underestimation of apparent numerosity (rightward shift of the dark green curve), and adaptation further affected connected dots (light green curve shifted even further).

timated – as showed by the rightward shift of the dark green curve. Moreover adaptation caused a further shift of the curve (light green), which indicates an even stronger underestimation compared to the baseline condition.

Fig. 4.7 shows average pre- and post-adaptation PSEs (gray and hatched bar, respectively) for the three main conditions, along with individual data. While a 20-dot adaptor does not affect the perceived numerosity of a 20-dot patch (left-most bars), it does affect perceived numerosity of probes with

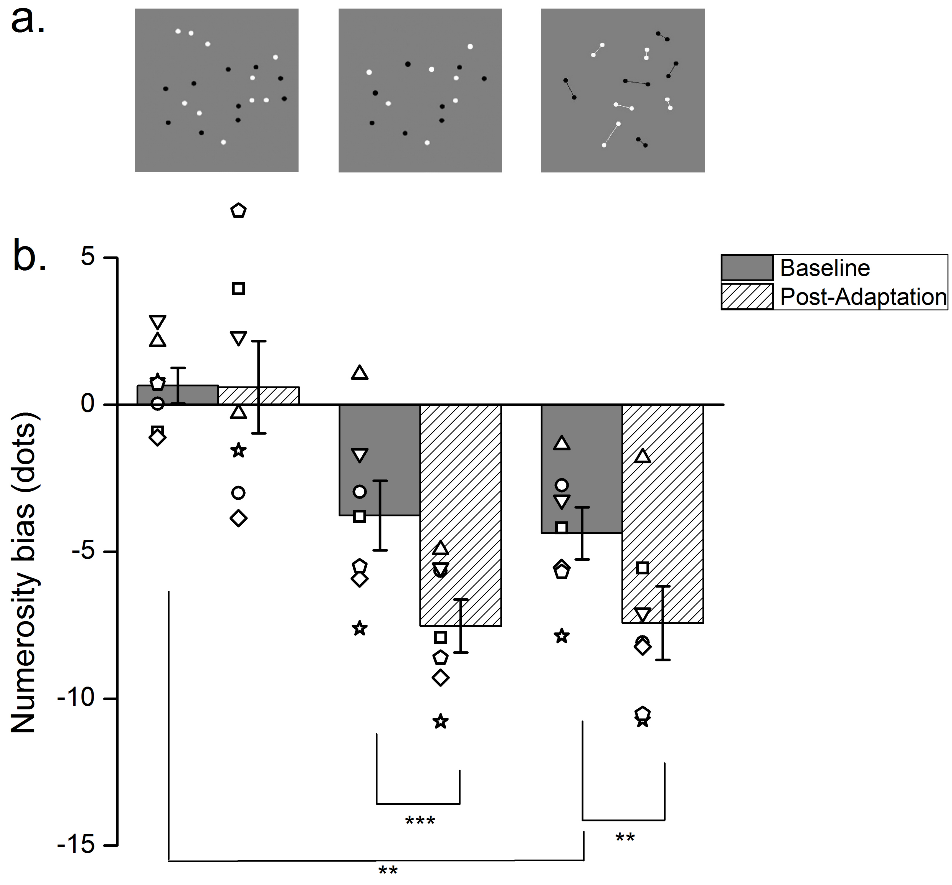

Figure 4.7: **Adaptation Effect for the three conditions.** Numerosity bias of a test patch before and after adaptation (grey and hatched bars), with individual data shown by the symbols (different shape for each subject). Three different conditions are plotted (showed in panel a.): twenty unconnected dots; a reduced number of unconnected dots (matching the numerosity of the connected stimulus); and twenty dots connected by ten lines. Error bars indicate S.E.M., Stars indicate statistical significance (* p<0.05, ** p<0.01, *** p<0.001).

lower numerosities, those matched to the apparent numerosity of the connected stimuli (average numerosity 16.22 ± 1.19: centre bars). The crucial test is for the twenty connected dots (right bars), which are perceived to 15.62 ± 0.89. Adaptation also affected this stimulus, by about 20%, very similar to the effect observed in the matched condition, with stimuli with a physically smaller numerosity compared to the 20-dot adaptors (average numerosity reduction equal to 23%).

We performed a two-way repeated-measures ANOVA with factors "configuration" and "adaptation", with a pairwise multiple comparison analysis run according to the Holm-Sidak method. The analysis first showed that both configuration and adaptation have a significant influence on numerical estimates (df = 2, F = 28.6, $p < 0.001$ and df = 1, F = 21.5, p = 0.004, respectively), and that there is a significant interaction between the two factors (F = 3.9, p = 0.048). The multiple comparisons procedure showed that when dots are connected by task-irrelevant lines their perceived numerosity is underestimated significantly, compared to the unconnected dots (20.65 ± 0.60 versus 15.62 ± 0.89 dots; t(7) = 4.181, p = 0.001). Also the effect of adaptation on connected dots appear to be statistically significant (t(7) = 3.222, p = 0.005), as well as the effect of adaptation on patches with physically smaller numerosity (matched condition, t(7) = 3.968, $p < 0.001$).

Fig. 4.8 shows the individual data for the two main experimental conditions, plotting adaptation effect (change in PSE normalized by baseline) in percentage against baseline PSE. While there is some scatter in the data, the trend is clear: there is little effect of adaptation, on average, on the unconnected dots (red data points), while for connected condition, adaptation had a clear effect for almost all subjects, on average about 20% adaptation.

4.2.4 Discussion

This experiment showed that adaptation a 20-dot patch, which has no effect on the perceived numerosity of a patch of 20 unconnected dots, decreases the perceived numerosity of 10 dot-pairs, to the same extent as it does to a patch whose numerosity matches the apparent numerosity of the connected dots.

Connecting dots is a simple but reliable way to dissociate perceived from physical numerosity: connecting dots not only reduces their perceived numerosity, but also affects the selectivity of the fMRI BOLD response in the number area in the Intraparietal Sulcus (He et al., 2015). These studies suggest that numerosity perception depends on segmentation of the elements in perceptual objects, following several rules such similar shape, orientation, common fate, and connectedness (Wagemans et al., 2012; Anobile et al., 2015b).

Interestingly, there is evidence that connectedness affects only perceived numerosity, but not action. Reaching movements are biased towards the more numerous stimuli (Milne et al., 2013), but this is unaffected by connecting the

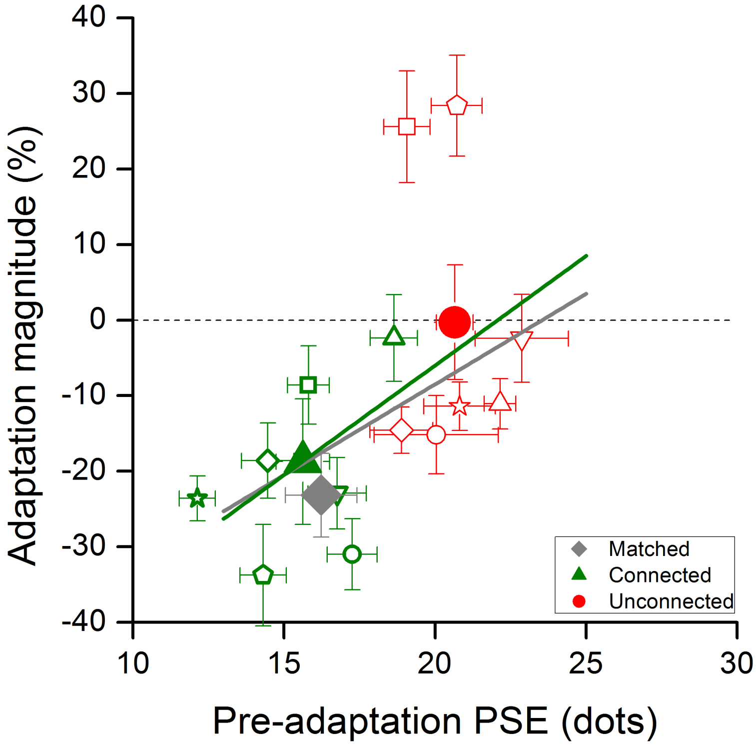

Figure 4.8: **Individual data.** Individual data for the connected (green symbols) and unconnected (red symbols) conditions, plotting change in PSE normalized by baseline against baseline PSE. The large symbols show group averages, the grey diamond the average of the matched, unconnected condition. There is little effect of adaptation on the unconnected dots (average adaptation -0.087% ± 7.75), while adaptation had a clear effect for almost all subjects in the connected condition (average adaptation -20.11% ± 4.64), similar to that of the matched, unconnected condition (average adaptation -23.10% ± 1.67).

elements. Perhaps the mechanism biasing action rests on a more primitive estimate of numerosity, unaffected by segmentation.

That the effects of adaptation are determined by perceived rather than actual numerosity is further evidence that numerosity perception relies on mechanisms distinct from those for texture, or density. Adding the lines that connect the dot-pairs increases rather than decreases the density of the stimuli. If adaptation were based on density, it should occur in the other direction, as adapting to a less dense stimulus causes the test to appear more dense. That our adaptation goes in the other direction is clear evidence that numerosity and density are distinct perceptual processes.

Recently, several researchers have suggested that numerosity and texture may be encoded by the same (or similar) neural mechanisms (Dakin et al., 2011). While this idea is appealing for many reasons, including its economical and parsimonious use of neural resources, there has been little evidence in favour. On the contrary, the evidence points to the existence of very distinct mechanisms, obeying different psychophysical laws, operating at different densities (Anobile et al., 2014), and different dependence on eccentricity (Anobile et al., 2015a). Numerosity mechanisms are favoured by conditions where the individual items can be segregated to allow enumeration: densities cannot be too high, and the limit depends on eccentricity (2.27 dots per square deg for central viewing, 0.8 for 15° eccentricity). When these conditions are met thresholds obey Weber's law. For higher densities, texture-like mechanisms cut in, obeying a square-root law (Anobile et al., 2015b).

The current study provides further evidence for this distinction, and shows that adaptation acts directly on perceptual mechanisms sensitive to numerosity, rather than via surrogate mechanisms such as density. At the moderate densities used here, the effects of adaptation depended not on the density of the stimuli, but on their apparent numerosity (which was less for the physically denser stimulus). This is further evidence for the existence of numerosity mechanisms, quite separate from those encoding density, and evidence that adaptation of numerosity works directly on mechanisms encoding number, not via some other route. Likely candidates for the site of the adaptation are LIP, which has a gradient response to number or, perhaps more likely, IPS, which has been shown to be influenced by the connectedness illusion.

Chapter 5

A common mechanism for space, time and numerosity

5.1 Introduction

With few exceptions, most of the previous work of the past decades on sensory mechanisms has been concentrated on the specific aspects of several different processes dedicated to different features of the sensory world. So studies concerning the mechanisms exploited by the brain to process spatial, temporal and numerical (or, more generally, quantity) information has usually been done independently, generating different stream of research.

However, in recent years, some authors started to wonder about the possibility that the mechanisms for processing different kinds of sensory information might not be so different, and, at least, that they might share some processes or neural substrates. Particularly, the mechanisms for the processing of spatial, temporal and numerical information have attracted much attention, and, even if at first glance they could seem unrelated, they could be more similar than previously thought.

In our everyday lives, we constantly face a very rich environment, full of sensory stimulations and information. Human beings rely mostly on vision as the principal and more important source of information, and across several decades of research, it has been pointed out that our visual sense involves a very complex machinery to select and process relevant information about the external world. The visual world is indeed characterized by many different features, with many of them requiring specialized mechanisms to be processed, at different levels of the visual hierarchy - such contrast, shape, orientation, color, motion, depth, and so on (e.g. Ferster & Miller, 2000; Gegenfurtner, 2003; Burr & Thompson, 2011). However, among the many features of the world, what are the most fundamental visual characteristics that form the ground over which our behavior is controlled, in order to successfully interact with the environment? Probably, if we try to make a list of the most important information that we must take into account for our survival, we will discover that on the leading positions there are spatial, temporal and numerical information (Walsh, 2003; Dehaene & Brannon, 2010). Space, time and number processing indeed allow us to answer several fundamental questions, necessary in order to understand the outside world - that is, what is present in the visual field, where the objects are, when events occur (and for how much time), and how many objects are out there. Interestingly, for most of the perceptual judgments regarding these domains, the brain must encode and process quantities.

Could quantity manipulation in all these three domains be underlayed by a common or at least similar set of processing mechanisms and brain circuitry? It is very difficult to imagine how any animal species could survive without mechanisms for spatial navigation, for time orientation and for elementary numerical elaborations (i.e. decide which set possess more elements) (Deahene & Brannon, 2010) and, indeed, this idea has driven some researchers to hypothesize that space, time and number might be processed by a common mechanism, with a similar metric, implemented in (at least partially) overlapping brain regions. Particularly, Walsh (2003) proposed that these different magnitudes might be processed by common machinery, called Generalized Magnitude System (GMS).

Such a common magnitude system would be involved with the dimensions defined by Stevens (1975) as “prothetic” – that is, dimensions that can be perceptually experienced as “more than” or “less than”. Indeed, space, time and numerical magnitudes are all prothetic dimensions. Moreover, as proposed by some authors (Fischer, 2003; Rossetti et al., 2004), one common link between these dimensions might be action. According to this idea, each dimension could

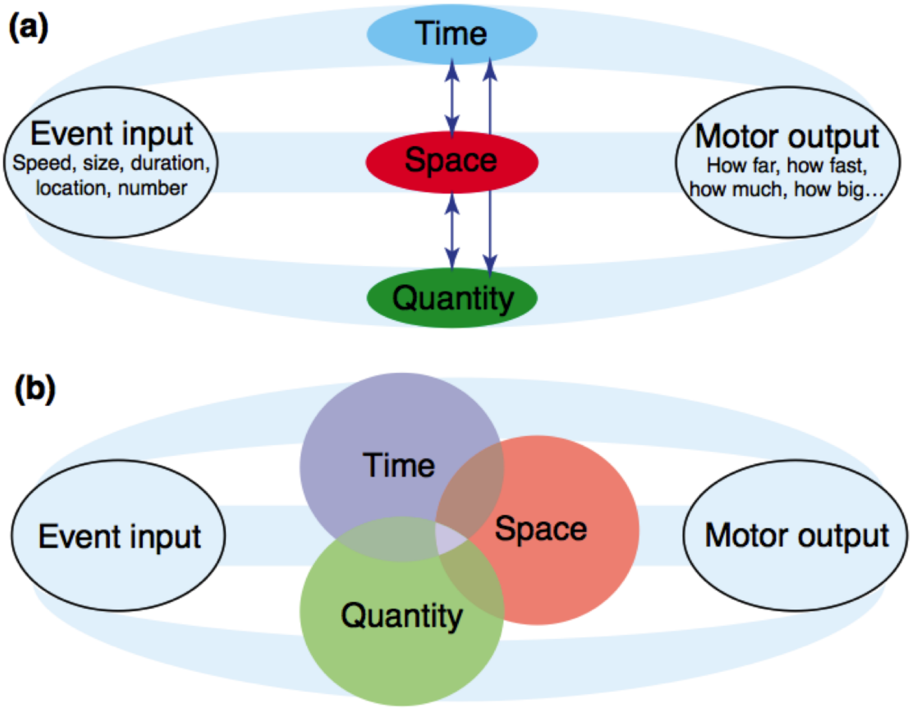

Figure 5.1: **Two schemas for the processing of space, time and quantity.** (a) The three different magnitudes might be analyzed separately, and compared according to their own metrics. (b) On the other hand, space, time and quantity might be processed by a Generalized Magnitude System, with a common metric. Adapted from Walsh, 2003.

be analyzed separately from the other two, and then compared with their own specific metric in order to guide motor outputs. Conversely, according to the ATOM theory, in order to support motor actions, all the three dimensions might be processed exploiting a common metric (Fig. 5.1).

Indeed, such theoretical instantiation proposing a common metric for different magnitudes has been supported by many experimental findings, both at the physiological and perceptual level.

Some hints suggesting a common system for such prothetic dimensions have been noted even a long time ago. Beside philosophers like Kant (Kant I., Critique of Pure Reason) and Pascal (Pascal B., Pensées), who recognized that basic principles like space, time and numbers represent a structure over which all the human experience is build on, connections between time and number and quantities in general were noted since at least 1890 (Walsh, 2003; Walsh

& Pascual-Leone, 2003). Animal studies also pointed out some similarities between different dimensions. For instance, testing the ability of rats in estimating small quantities (i.e. 2-8 tones) and small durations (i.e. 2-8 seconds), Church & Meck (1984) noted that rats' generalization behavior showed that they estimated "4" to be halfway between 2 and 8, both for the numerical and the temporal dimension. Moreover, also from lesion studies have emerged a tendency for specific deficits concerning spatial, temporal and numerical abilities to be combined after particular brain damages (usually involving the parietal cortex), while isolated dysfunctions affecting only one perceptual domain is a far rarer phenomenon (Critchely, 1953). Even if the variability of lesions across different neuropsychological patients is too wide to allow consistent conclusions about the neural underpinning of different sensory abilities, such overlap between impairment in different domains might suggest some indications of overlapping mechanisms.

In the field of human psychophysics, several behavioral observations supported the idea of a common metrics for the processing of spatial, temporal and numerical information. A clear prediction of the ATOM theory is that the interactions between magnitudes should be as a monotonic mapping: "bigger", "faster", "further" in one dimension should correlate with "bigger", "faster", "further" in another dimension, and vice versa (Bueti & Walsh, 2009). One line of evidence supporting this idea comes indeed from similar perceptual effects affecting the representation of different magnitudes and mutual interactions among them. For instance, it has been shown that time could be affected by both space and number: larger stimuli (in size) or more numerous stimuli (i.e. stimuli containing more elements) are perceived to last longer in time, compared to their smaller or less numerous counterparts with the same physical duration (Xuan et al., 2007; 2009; Rammsayer & Verner, 2014; 2015) – a relation that points out a robust link between the representation of time and other magnitudes (see also Chapter 3 for other details about temporal distortions due to the modulation of other features).

Another example of temporal distortions due to the modulation of the spatial domain is the observation that the time experienced in a given environment is related to its scale. For instance, De Long (1981) asked participants to perform a task in an environment which was build as 1/6, 1/12 or 1/24 of its supposed scale, and to stop the task after 30 minutes had passed. Interestingly, he noted that the ratio of time passed to time estimated scaled according to a proportion similarly to that of the environment scale (see also Mitchell & Davis, 1987).

The association and modulation between different dimensions is even more pronounced in children, who very easily associated “more A” with “more B”, even when provided with clear evidence that the two features A and B are not correlated and totally irrelevant for each other. A clear example comes from one study by Stavy & Tirosh (2000), who presented children with two toy trains running along a track. When asked which train moved faster – even if speed was actually the same – the bigger (in size) train was systematically reported as the faster one. This showed that the spatial dimension had a strong influence on speed, which is a dimension intrinsically related to time.

Number has also proven to be a domain tightly linked to the dimension of space, and indeed the possibility to represent numbers in a spatially organized manner (i.e. the mental number line; Dehaene, 2003; Izard & Dehaene, 2008; Zorzi et al., 2006; Zorzi et al., 2012), is a well accepted phenomenon. Related to this, it has been shown that reaction times to numerical stimuli are faster for small magnitudes when the response effector is the left hand, while for big numbers the faster effector is the right hand - an effect that clearly show how numbers are organized along the mental number line from the left to the right (SNARC effect; Dehaene et al., 1993; Fias, 1996). Moreover, the apparent number of objects in a given visual scene can be affected by spatial information, as showed by studies pointing out a modulation of perceived number caused by connections between pairs of elements (He et al., 2009; Franconeri et al., 2009), and particularly showing that the representation of the number of items is strongly affected by topological invariants, such connectedness and the inside/outside relationship (He et al., 2015), as well as by a symmetrical arrangement of the items (Apthorp & Bell, 2015) (see Chapter 4 for a more detailed discussion about this topic).

According to this evidence, the most malleable dimension seems to be time, while space in particular could hardly be affected by time and number. This asymmetry between different dimensions has led to the proposal of a different kind of relationship between them: instead of shared mechanisms processing different magnitudes, time (and possibly number) might be accounted as a metaphor of space – an idea arising from the striking observation that most of the words describing the temporal dimension are the same used for space (Boroditsky, 2000) – that is, more "abstract" dimensions are identified using the concepts of more "concrete" domains. Supporting this idea, much evidence has been provided about the dissociation between space and time, with the latter being affect by the spatial information without a corresponding effect in the other direction (Casasanto & Boroditsky, 2008; Casasanto et al., 2010; Bottini & Casasanto, 2013). However, in a recent study by Cai & Connell (2015), it has been pointed out that observations suggesting an asymmetric relationship between space and time are systematically biased. Indeed, according to Cai & Connell, this asymmetry might be due to the intrinsic resolution of the visual system: while vision has a good spatial resolution, its temporal resolution is relatively poor compared, for instance, to the auditory modality (Morein-Zamir et al., 2003; Vroomen & de Gelder, 2004). Indeed, they showed that changing sensory modality the effect could be reversed: spatial processing in the haptic domain (i.e. judging the length of a stick) could be strongly biased by an auditory temporal duration (i.e. a longer sound leads to the perception of a longer stick).

Other evidence of a similar mechanism for the processing of spatial, temporal and numerical information comes from studies investigating perceptual distortions at the time of saccades. As discussed in Chapter 2, performing saccades is strongly related to robust distortions of the perceived position of brief peri-saccadic stimuli (Fig. 5.2a), which resemble a compression of visual space around the target of the saccade (Ross et al., 1997; but see also Lappe et al., 2000 and Morrone et al., 2005). However, some recent studies also investigated the possibility that the distortions triggered by saccades – and possibly related to the mechanisms for visual stability – might not be limited to perceptual space, but also extend to other magnitudes (i.e. time and quantity). Interestingly, it has been shown that saccades not only cause a distortion of perceptual space, but also of event time (Morrone et al., 2005; Binda et al., 2009). If a pair of bars separated by a 100 ms interval is flashed well before or after a saccade, the interval between them is usually perceived correctly. However, when the stimuli are presented in proximity of the saccadic onset,

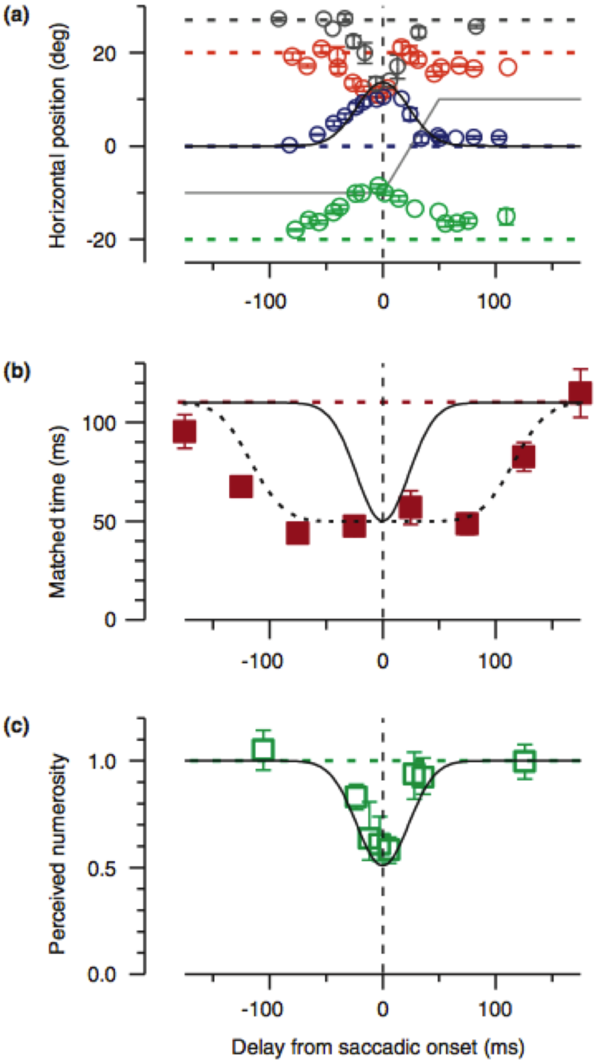

Figure 5.2: **saccades compress space, time and numbers.**The three panels show the results of different studies investigating distortions in the perception of space, time and numerical magnitudes at the time of saccades. In all the panels, the vertical dashed line marks the saccadic onset. The gray line in panel a depict the typical saccade trajectory, straddling 20 deg of visual space. (a) Perceived spatial position of a brief bar presented at four possible locations (circles, indicated by different colors), as function of the presentation time relative to the saccadic onset. Dotted lines indicated localization during fixation. (b) Apparent duration of the 100 ms interval between two briefly flashed stimuli (bars), as function of the presentation time relative to the saccadic onset. The solid curve indicates the predicted time course of the compression effect obtained deconvoluting the spline fit of the data (dashed curve) with the temporal interval between the two markers (100 ms). The predicted time course appears to be more tightly tuned compared to the actual data, but this is because participants' performances were tested using a broad temporal stimulus, that blurs the effect over time. The dashed horizontal line depicts perceived duration during fixation. (c) Perceived relative numerosity of a 30-dots patch flashed perisaccadily, normalized by estimated numerosity during fixation (horizontal dashed line) – i.e. the plot shows the ratio of the numbers of dots in the probe stimulus when presented during fixation to the number presented during saccades. The black solid curve in all the plots are Gaussian functions with the same mean and standard deviation that best fit all the data. Adapted from Burr et al., 2010.

the interval is strongly distorted, compressed approximately to half of its physical duration (Fig. 5.2b). Moreover, such compression of perceived time also follows a very similar time-course compared to the spatial distortions.

So both space and time are severely distorted just before and during eye movements, but what about numerical magnitudes? Further studies also investigated this issue, presenting array of stimuli around the onset of saccades, and asking participants to compare the number of elements of the perisaccadic array with the numerosity of another stimulus presented well before the onset of the eye movement (Binda et al., 2011). Results showed that when stimuli were flashed well before or after the saccade, the numerosity of the test patches was correctly perceived. However, when the array of stimuli were presented near the onset of the eye movement, perceived numerosity was strongly underestimated, to about one half of the physical numerosity, again with a time-course very similar to the distortions of space and time (Fig. 5.2c). Moreover, it has also been shown that such compression of perceived numerosity extends also to abstract quantities, affecting the representation of symbolic numerical magnitudes (Binda et al., 2012).

Beside the psychophysical and behavioral evidence supporting the idea of a common mechanism for space, time and quantity, several studies pointed out similarities between them also exploiting imaging techniques on human subjects. Several imaging studies indeed suggested that temporal, numerical and spatial information, as well as other magnitudes, activate overlapping structures in the parietal cortex (Cochon et al., 1999; Dehaene et al., 1999; Rao et al., 2001; Simon et al., 2002; Pinel et al., 2004). Activation for spatial and temporal tasks is mostly found in the right inferior parietal cortex (rIPC), while numerical task might activate parietal cortex bilaterally. Indeed, while numerosity estimation and comparison tasks are usually associated with right parietal activation, other tasks like exact calculation may activate the left hemisphere – consistently with the fact that such tasks involve the use of language and symbols (Cohen & Dehaene, 1996; Maquet et al., 1996).

Figure 5.3 shows the areas investigated across several studies, and that were found to be involved with predictive remapping processes (see Chapter 2), and with the processing of temporal and numerical information. The processing of these three dimensions involves a fronto-parietal network of different brain

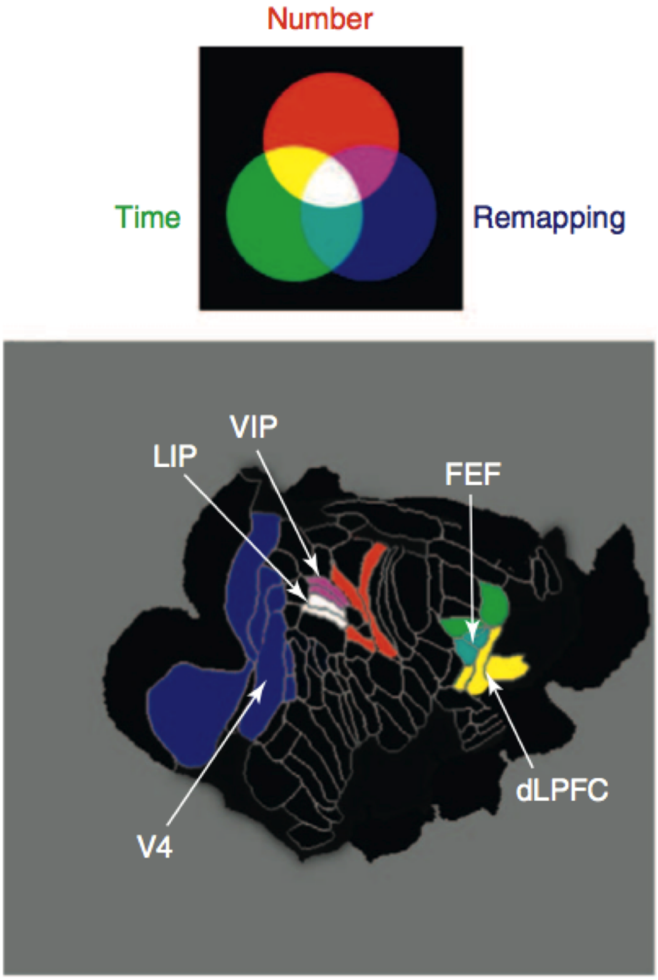

Figure 5.3: **Network of areas involved in the processing of space (saccadic remapping), time and numbers.** The brain map represents a flattened cortical surface of a monkey right hemisphere. Borders define cortical areas identified in the partitioning scheme by Lewis & Van Essen (2000). Different colors highlight areas involved with the processing of different information: red indicates areas involved in the processing of numbers; green indicates areas modulated by temporal durations; in blue are indicated areas involved with the saccadic remapping. The intersecting of two or more of these abilities is indicated by the summation of the corresponding colors (see legend at the top of the panel). Please note that this figure – as reported in Burr et al., 2010 – is intended only as a schematic representation. For the exact localization of the areas and the recording sited mentioned, the reader is referred to the original papers (see Table 5.1). Adapted from Burr et al., 2010.

areas, which seems to intersect at the level of the Lateral Intraparietal (LIP) area, which indeed is thought to play a key role in the processing of different magnitudes.

Time	Leon & Shadlen, 2003	LIP
	Genovesio et al., 2005	46, 8, 9, 6
Number	Sawamura et al., 2002	5, 2
	Nieder & Miller, 2004	dLPFC, PPC
	Roitman et al., 2007	LIP
Space	Duhamel et al., 1992	LIP
	Tolias et al., 2001	V4
	Nakamura and Colby, 2002	V2, V3, V3A
	Umeno and Goldberg, 2001	FEF

Table 5.1: Sources used by Burr et al. (2010) to construct the schematic representation reported in Fig. 5.3. Adapted from Burr et al., 2010.

Interestingly, in a recent study by Kaufmann and colleagues (2008), the authors compared brain activation of adults and children in non-symbolic numerical and spatial tasks, finding evidence of the role of the action system in learning and training with magnitudes, and of overlapping representation of space and number in adults. Namely, what they found is an overlap between adult brain activation for a spatial and a numerical task, located in the posterior superior parietal lobule (PSPL). On the other hand, children showed relatively more activation in the supramarginal gyrus (SMG), the precentral gyrus, and the lateral anterior intraparietal sulcus (IPS). This further suggests that in the developing brain the activity in parietal cortex might be more spread, with spatial and numerical representation sharing neighboring areas, while becoming more overlapping with increasing mastery in such tasks, which could be related to an increased functional connectivity.

Studies investigating brain activity during time perception tasks consistently reported a network of areas including parietal and prefrontal cortex, supplementary motor area, cerebellum and basal ganglia (Macar et al., 2002; Coull et al., 2004; Lewis & Miall, 2006a,b). Interestingly, Lewis & Miall divided the functions involved in temporal perception into “cognitive” and “non-cognitive”, showing that while dorsolateral prefrontal cortex seems involved in both “cognitive” timing and working memory, the parietal cortex is associated with more automatic, “non-cognitive”, temporal processing.

However, some studies did not find time-related activation in the parietal cortex. Such failure in showing activation for temporal task in right parietal cortex could be related to the specific methodology used. Indeed, if other magnitudes are used as control tasks, subtracting the brain activations would result in actually cancelling out the parietal activity.

Beside the results of imaging and neuropsychological studies, overall suggesting overlapping areas for space, time and numerical magnitudes, also Transcranic Magnetic Stimulation (TMS) investigations obtained results in such direction. Indeed, across several studies, it has been shown that stimulation over the parietal cortex impairs performances in tasks involving spatial processing (Ashbridge et al., 1997; Rushworth et al., 2001; Bjoertomt et al., 2002), numerical comparison (Gõbel et al., 2001) and duration discrimination (Walsh & Pascual-Leone, 2003).

5.1.1 Conclusion and outline of the next sections

While much evidence seems support the idea of a Generalized Magnitude System processing spatial, temporal and numerical information by means of a common metric, some aspects of such mechanism are still not completely clear, and so there is still room for debate. In the next sections, we will report two studies aimed to further investigate some features of the hypothetical magnitude system, related to the possibility to observe similar effects of motion adaptation on time and numbers, and regarding the interplay between different features – i.e. the possibility to exploit spatial information to improve judgments in the temporal domain. To do so, first (Section 5.2) we studied the effect of motion adaptation on perceived numerosity (Fornaciai et al., 2015), as it has been previously studied on perceived time (Johnston et al., 2006; Burr et al., 2007; Fornaciai et al., 2014; Fornaciai, Arrighi, & Burr, Under review; see Chapter 3 for a more detailed description of this effect). Indeed, motion represents an ideal candidate to study the intertwining between different magnitudes, given that it is intrinsically defined as a combination of spatial and temporal information (i.e. modulation of spatial information over time), and that it might also contain numerical information, since a moving stimulus could be defined by a certain number of events over time. Moreover, given that motion adaptation has proven to be effective in distorting the representation of both space (e.g. positional aftereffect; Snowden, 1998; Nishida & Johnston, 1999) and time (e.g. Burr et al., 2007), it represents a good test bench for the ATOM theory: if space, time and number actually share some functional mechanisms, so it might be possible that motion adaptation could distort also numerical estimates. Furthermore (Section 5.3), in order to investigate the possibility to exploit spatial information to improve performances in the temporal domain, we studied how temporal order judgments are affected by the availability of additional spatial information (Fornaciai & Di Luca, Under review), which might allow the prediction of both where and when an event is expected to happen, and that in turn may improve the temporal resolution around the predicted timing. Finally, we attempted to test the possibility that the information carried by the asynchrony between the auditory and visual component of an audio-visual event might be exploited by the brain as a cue for visual distance in depth.

5.2 The effect of motion adaptation on perceived numerosity

5.2.1 Introduction

As we introduced in the previous section, several studies found robust mutual influences among the spatial, temporal and numerical dimensions. However, another line of evidence seems suggest similar perceptual distortions caused by adaptation, and particularly for motion adaptation. On the one hand, it has been shown that motion adaptation has a strong influence on the spatial dimension of the stimuli, strongly distorting the perceived position of subsequent stimulus (Snowden, 1998; Nishida & Johnston, 1999; McGraw et al., 2002). On the other hand, several recent studies pointed out that fast motion adaptation (usually 20 Hz) can also affect time, causing a robust compression of perceived duration (Johnston et al., 2006), that seems to be spatially selective in real-world coordinates (Burr et al., 2007). Moreover, Fornaciai and colleagues (Under review), further found that such distortion is also selective for the motion profile used, with simple translational motion being the most effective, while more complex motion profiles (such radial and circular motion) exerting no effect of perceived time.

Following the idea that spatial, temporal and numerical information might be processed by a common magnitude system, here we tested whether motion adaptation could also affect numerical estimates. First, we used sinusoidal oscillating grating, trying to replicate the adaptation-induced duration compression previously demonstrated across many studies. Besides translational motion, we also used a more complex motion pattern, namely circular motion, since it has been shown that the effect of adaptation on time is also selective for the motion profile used. Our prediction is that while translational motion should behave as a strong adapting stimulus for perceived numerosity, the more complex motion should have a weaker effect, mirroring the effects of apparent time. We also devised a series of control experiments in order to control for the possible confounding factors provided by contrast/visibility effects or spatial frequency adaptation.

5.2.2 Materials and methods

Subjects

A total of 30 subjects participated in the study (11 females, age ranging from 21 to 30 years), either in one or multiple conditions, after giving their informed written consent. All the participants were naive to the purpose of the study, with the exception of authors I.T. and M.F., who participated in all the experimental conditions. Experimental procedures were approved by the local ethics committee, and were in line with the declaration of Helsinki.

Apparatus

Visual stimuli were generated with the Psychophysics Toolbox V.3 (Kleiner et al., 2007) for MatLab (version 2010b) running on a PC computer, and presented on a Barco CRT monitor (Barco Calibrator Line), subtending 40 x 30 degrees of visual angle at the viewing distance of about 57 cm. Screen resolution was set to 800 x 600 pixel, and the experiments were performed with the monitor running with a refresh rate of 100 Hz. Test and reference stimuli in the numerosity task were patches of dots randomly arranged inside a circular area with a minimum radius equal to 1 deg and a maximum radius equal to 5 deg, in order to be arranged as an annulus (see below). The position of each dot in both the reference and test stimuli was assigned on a trial by trial basis, with the constrain of a minimum distance from the position of each other dots equal to 0.75 deg. Adapter stimuli concerned two different kinds of motion profiles: a) linear translation and b) circular motion. Translating stimuli were vertical luminance modulated gratings (SF 1 cpd) drifting horizontally, changing drift direction every 2 seconds. Circular motion stimuli were windmill-like rotating gratings, with spatial frequency increasing from 0.5 to 1.2 cpd respectively for the outer and inner border. All stimuli were windowed within an annular mask (inner and outer radius equal to 1 deg and 5.5 deg, respectively), with borders blurred by a Gaussian smoothing (spatial constant equal to 0.15 deg) and were presented with a Michelson contrast of 90%. To avoid the displacement of test dots in zones not covered by the adaptor stimuli, we also arranged the random-dot patches in an annular shape.

Procedure

The main experiment comprised a test phase consisting of the presentation of two stimuli, test (variable numerosity) and reference (fixed numerosity), while participants kept the gaze on a central fixation point. Stimuli were presented for 500 ms, simultaneously on both sides of the central fixation point, with a horizontal eccentricity of 10 deg, calculated as the center-to-center distance between the fixation point and the stimulus. At the end of each trial, subject were required to indicate which stimulus contained more elements. Across each experimental block, the numerosity of the reference stimulus was kept constant at 15 dots, while that of the test was varied by an adaptive QUEST (Watson & Pelli, 1983) routine (with a range of 0.3 Log Units around the reference numerosity). Usually 3-5 sessions (of approximately 30 trial) were run for each condition. The final estimate of PSE was taken as the median of the best-fitting cumulative Gaussian function to all the data of a particular condition (percentage “more numerosity than” against test physical numerosity).

On separate sessions, subjects performed numerosity discrimination after being adapted to fast (20 Hz) or slow (5 Hz) motion, either translational or circular. Each trial started with the presentation of the adaptor stimulus, displayed in the left portion of the screen with a horizontal eccentricity of 10 deg (corresponding to position of the test stimulus described above). The adaptor was displayed for 20 seconds in the first trial, and then for 8 seconds in the subsequent trials. The subsequent test phase, followed the same procedure as the baseline condition described above. The total amount of adaptation-induced distortions on perceived numerosity was defined as the difference between the PSEs of the baseline and the adaptation condition, normalized for the baseline value and converted in percentage.

We also devised a series of control experiments, using uniquely translational motion adaptation, in order to test whether the adaptation effects could be selective for the numerosity range used, or whether masking effects could drive them. However, in all the conditions the main procedure was always the same, except for the difference listed below. First, we tested the effect of adaptation increasing the reference numerosity (50 dots), or reducing it to the subitizing range (3 dots), in two separate conditions. Moreover, we also reduced the adaptor contrast to 50%, to reduce any possible influence on stimulus visibility. Finally, we investigated the role of spatial and temporal frequencies reducing the spatial frequency of the adaptor to 0.5 cycles per degree, and testing also the effects of static (0 Hz) adapters.

5.2.3 Results

Fig. 5.4 shows the results of our main condition, concerning the average effect of translational and circular motion adaptation on perceived numerosity. We reported adaptation effects calculated as change in percentage of the post-adaptation PSE related to the baseline PSE. At first glance it is immediately clear that motion adaptation caused a relatively substantial change in the perceived numerosity of the adapted stimulus, corresponding to, on average, a reduction of 12.1% and 25.5% for translational motion, respectively in the 5Hz and 20Hz condition, and 10.7% and 16.4% for circular motion adaptation. In all the conditions tested, comparing baseline perceived numerosity with post-adaptation performances shows that baseline numerical estimates were significantly different from post-adaptation performances, with both adaptors and both temporal frequencies (One-Way R.M. ANOVA, baseline PSEs versus post-adaptation PSEs in 5Hz and 20Hz conditions; translation: $F(2,14) = 40,017$, $p < 0.001$; circular: $F(2,13) = 13.997$, $p < 0.001$). A Holm-Sidak multiple comparison analysis (versus control [baseline]) further revealed that

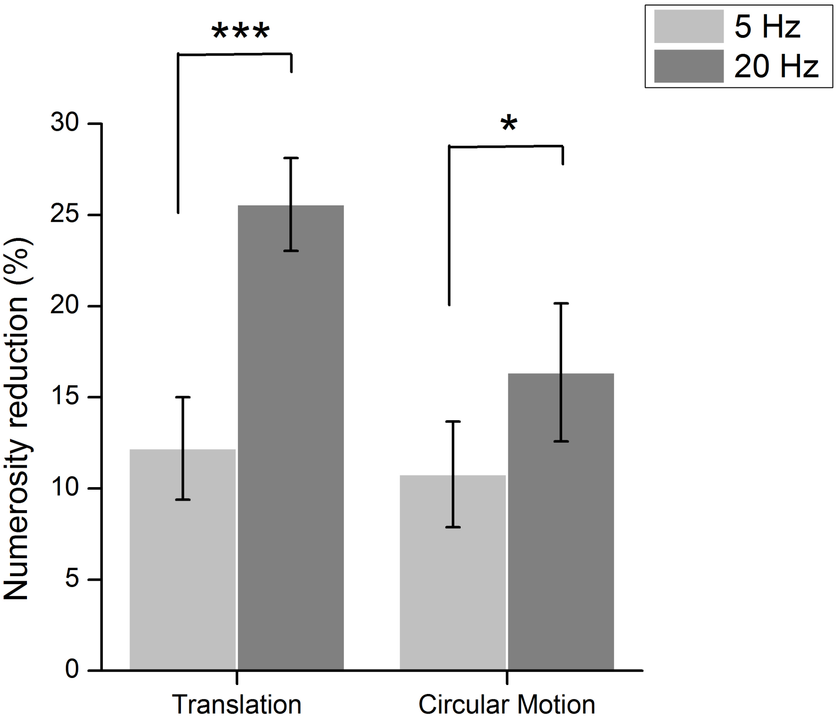

Figure 5.4: **Average effects of translational and circular motion adaptation.** Average effects of motion adaptation on perceived numerosity, calculated by the difference between post-adaptation and baseline PSE estimates, normalized by the baseline. Rightmost bars represent the effects of translational motion, both at 5 Hz (light gray) and 20 Hz (dark gray). Leftmost bars represent the effects of circular motion adaptation. Error bars represents S.E.M. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

baseline PSEs are significantly different from PSEs in both 5Hz and 20Hz adaptation conditions, both for translation (baseline versus 5Hz, t(14) = 4.346, $p < 0.001$; baseline versus 20Hz, t(14) = 8.945, $p < 0.001$) and circular motion (baseline versus 5Hz, t(13) = 3.301, p = 0.003; baseline versus 20Hz, t(13) = 5.231, $p < 0.001$) conditions. However, the effect of 20 Hz adaptation on perceived numerosity was stronger for both translation (two-tailed paired t-test, t(14) = 5.173, $p < 0.001$) and circular motion (t(13) = 2.318, p = 0.0374), compared to 5 Hz adaptation. Moreover, a comparison between the effects of the two motion profiles across the different temporal frequencies (One-Way ANOVA with Holm-Sidak multiple comparison procedure versus control ["translation 20Hz"]) showed that 20 Hz translational motion adaptation led to a stronger numerosity underestimation compared to all the other combinations (Translation 20Hz vs. Translation 5Hz, t(3) = 3.306, p = 0.003; Translation 20Hz vs. Circular 5Hz, t(3) = 3.593, p = 0.002; Translation 20Hz vs. Circular 20Hz, t(3) = 2.237, p = 0.029).

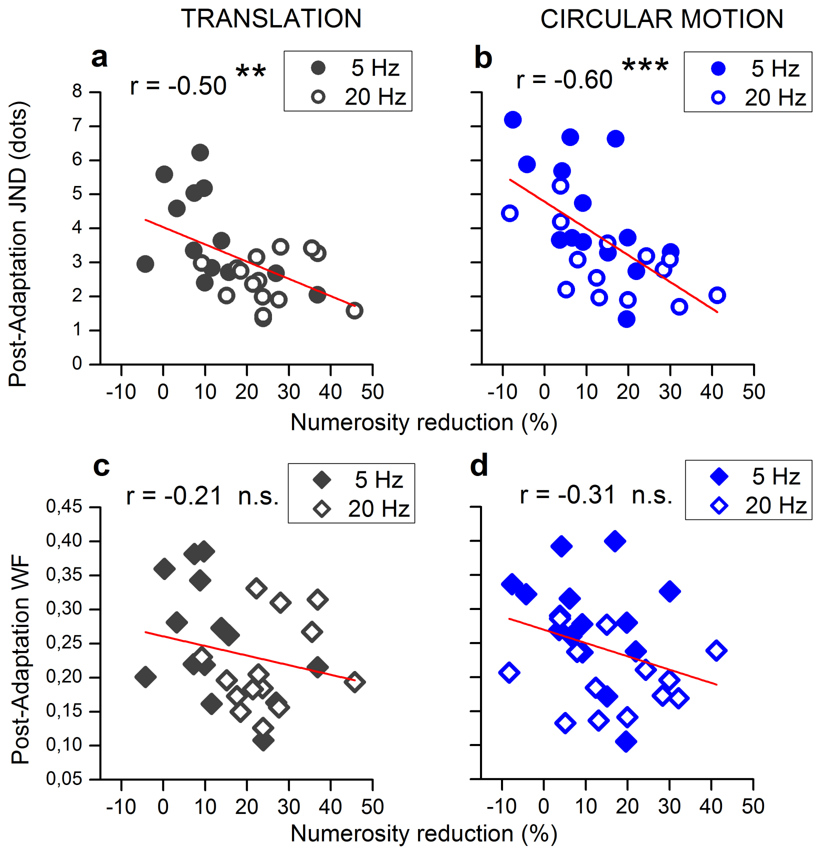

Figure 5.5: **JNDs and WFs as function of the numerosity reduction effect.** (a and b) Post-adaptation measures of sensitivity for numerosity discrimination (JND) plotted as function of the magnitude of numerosity reduction effects for translational (dark grey circles) and circular motion condition (blue circles). (b and c) Post-adaptation Weber Fractions (JNDs normalized by PSEs) plotted as function of the magnitude of numerosity reduction effects for translational (dark grey diamonds) and circular motion condition (blue diamonds). Points represent single subjects. Solid red lines represent linear fits to the data. Pearson correlations: n.s = not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

We also analyzed performance in terms of the precision (JND) in the task. Could the results be related to an increased uncertainty of numerosity judgment after adaptation? In Fig. 5.5a and 5.5b are reported participants' JND (with 5 Hz and 20 Hz conditions pooled together) as function of the magnitude of adaptation effects, respectively for translation and circular motion condition. As showed in the panels, adaption clearly affected precision for numerosity judgments, with JNDs showing a clear trend toward an improvement, as the effect increases: the more the numerosity underestimation, the higher the precision in the discrimination task (slope = -0.05, r = -0.50, p = 0.004 and slope = -0.07, r = -0.60, $p < 0.001$, respectively for translation and circular motion). Moreover, in Fig. 5.5c and 5.5d are reported participants' Weber Fraction (WF = JND/PSE) as function of the adaptation effect. Correlations between WF and the numerosity reduction effect are not significant for both translation (slope = -0.0014, r = -0.21, p = 0.244) and circular motion (slope = -0.0019, r = -0.31, p = 0.106), possibly because the decrease of JND as the effect increases seems to, at least partially, compensate for the decrease in PSE (which reflects the numerosity underestimation effect), leading to a flatter distribution of Weber Fractions – a result similar to what has been observed testing the effect of adaptation on duration perception by Fornaciai and Colleagues (Under review).

Is this effect specific for the numerosity range used? To control for this issue, we devised two additional conditions, where we tested the effect of translational motion adaptation either a bigger numerosity range (50 dots) or a smaller one (3 dots). However, these two conditions gave rise to two opposite predictions: while the effect on the bigger numerosity range should be comparable to the effect observed in the first experiment, adaptation should not affect the smaller one. Indeed, previous studies showed that the subitizing range (the range of numerosity between 1 and 5) is virtually immune to any adaptation effect – unless you divert the attentional resource needed for it to an additional task (Burr et al., 2010; Burr et al., 2011). So, if the effect observed in the first experiment actually concern the representation of approximate number, we should observed a similar effect on stimuli around 50 dots, while no effect should be evident on stimuli in the subitizing range. Fig. 5.6 (a and b) shows the results of these two conditions. While the effect on the 50 dots range is actually of a similar magnitude compared to the previous experiment (13.6% ± 5.95% and 24.2% ± 3.66% reduction in perceived numerosity, respectively for 5 Hz and 20 Hz adaptation condition), and statistically significant (One-Way R.M. ANOVA with Holm-Sidak multiple comparison procedure versus control

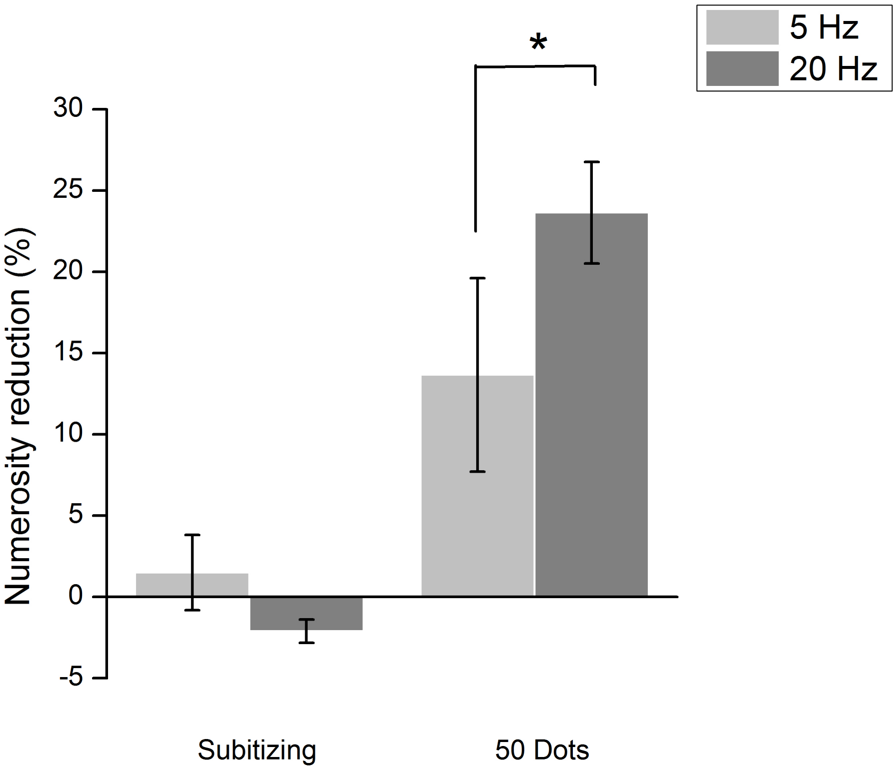

Figure 5.6: **Adaptation effects on numerosity in the subitizing and 50-dot ranges.** The different levels of numerosity reduction effect (percentage difference between post-adaptation and baseline PSE estimate, normalized by. baseline PSE) are reported for the conditions in which we reduced the numerosity range tested to the subitizing range (3 dots; rightmost bars), and in which we increased the range to 50 dots (leftmost bars), for 5 Hz (light grey) and 20 Hz (dark gray) adaptation conditions. Error bars represent S.E.M. * $p < 0.05$.

[“baseline”]; baseline vs. 5Hz adaptation, t(5) = 3.432, p = 0.006; baseline vs. 20Hz adaptation, t(5) = 5.867, $p < 0.001$; two-tailed paired t-test, 20 Hz effect significantly greater than 5 Hz effect: t(5) = 3.395, p = 0.0194), adaption had no effect on stimuli near the subitizing range (One-Way R.M. ANOVA, F(4,2) = 2.330, p = 0.159).

Furthermore, we also asked whether the perceived amount of dots could have been affected by adaptation effect in a more direct way, namely by a simple reduction in visibility (Georgeson, 1985). Even if the lack of effect on stimuli near the subitizing range makes this explanation less likely, we also repeated the translational motion condition (the most effective condition)

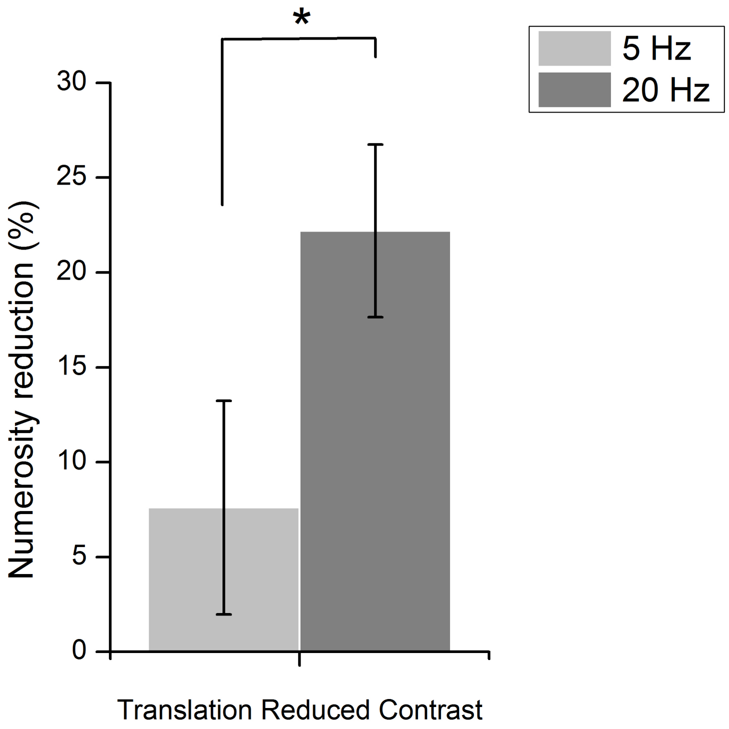

Figure 5.7: **Effect of adaptation with 50% contrast.** Motion adaptation-induced numerosity reduction effects after adaptation to translational motion, in the condition where the adaptor contrast was reduced to 50%. Error bars represents the S.E.M. * $p < 0.05$.

reducing the contrast of the adaptor stimulus to 50%, in order to keep the visibility of the dots patches not compromised by the adaptation. Indeed, as pointed out by Georgeson (1985), if the adaptor has a lower contrast than the test patch, so its effect on test's perceived contrast is small or even negligible. Fig. 5.7 shows the results of the reduced contrast condition. A two-way R.M. ANOVA first showed that 5 Hz adaptation resulted in a weaker effect compared to the previous conditions, which did not reach statistical significance (Holm-Sidak multiple comparison, baseline versus post-adaptation: $t(4) = 2.361$, $p = 0.061$). However, 20 Hz translational motion adaptation affected perceived numerosity in a very similar way as the full contrast condition, showing a significant reduction in perceived numerosity ($t(4) = 4.726$, $p = 0.004$), which is consistently stronger than the 5 Hz effect (paired sample t-test, $t(4) = 2.943$, $p = 0.0211$).

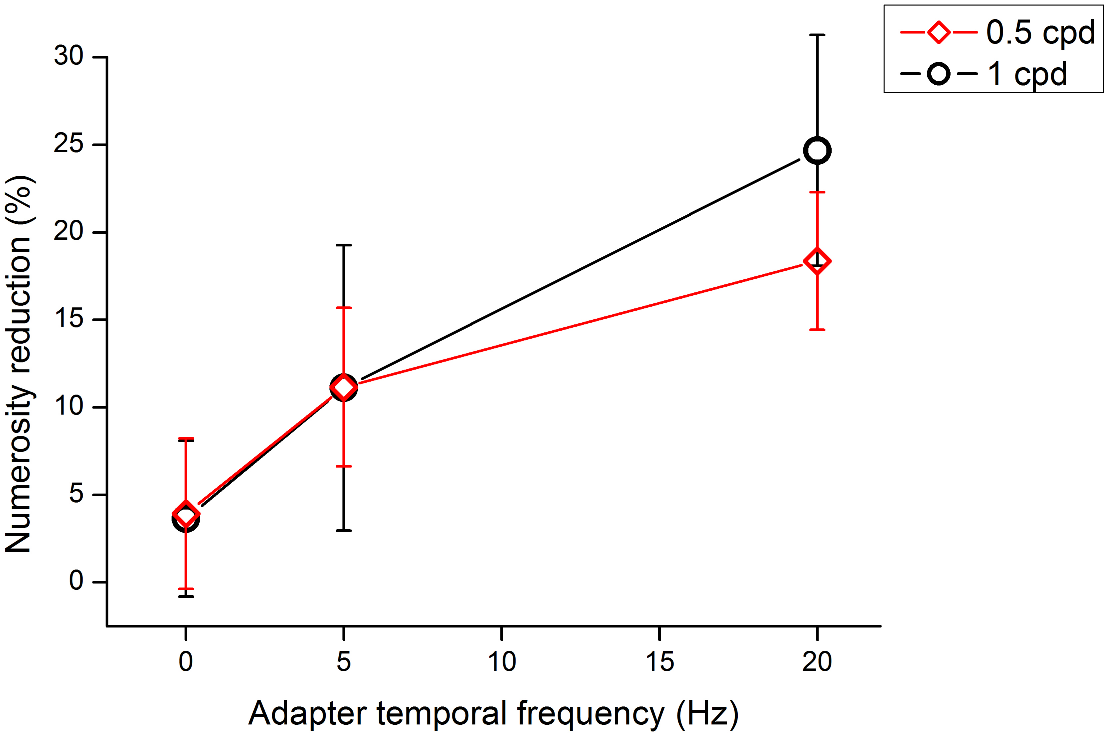

Figure 5.8: **Translational motion adaptation with different spatial and temporal frequencies.** Numerosity reduction effects reported as function of adaptation temporal frequency, for both the condition in which the adaptors were 0.5 cpd (red diamonds) and 1 cpd (black circles) drifting gratings. Note that the 5 Hz and 20 Hz effects reported for the 1 cpd conditions are the same reported in Fig. 5.4. Error bars represent S.E.M.

Finally, our last control concern the possibility that adaptation at the level of spatial frequency (SF) channels might have influenced numerosity estimation, masking some of the dots and causing an underestimation. To be sure that the observed effects have been driven by a spatial frequency adaptation, we repeated the translational motion condition using an even more different adaptor's spatial frequency, namely 0.5 cycles per degree (cpd), since adaptation effects at this level should be tightly related to the specific SF adapted. Moreover, if spatial frequency adaptation contributed in our pattern of results, this should be evident even with a static adaptor. The results of this condition are summarized in Fig. 5.8. What we found adapting to 0.5 cpd translating gratings is indeed consistent with the first experiment, with the 5 Hz and 20

Hz conditions in line with the previous observations (post-adaptation PSEs significantly different from baseline measures: t(4) = 2.615, p = 0.039 and t(4) = 4.119, p = 0.006, respectively), and the static condition (0 Hz) resulting in almost no effect on perceived numerosity (no difference between baseline and post-adaptation PSEs: t(4) = 0.975, p = 0.367). We also measured the effect of a static adapter (0 Hz) using 1 cpd gratings, finding again no effect on perceived numerosity, similarly to the 0.5 cpd grating (numerosity reduction effect = 3.64% ± 4.45%, paired sample t-test: t(4) = 0.792, p = 0.236). Finally, a series of tests confirmed that there is no statistically significant difference between adaptation effects with the 0.5 cpd and 1 cpd adapters, across the different temporal frequencies used to adapt the participants (Mann-Whitney Rank Sum Test for the comparison between 0.5 and 1 cpd in the 0 Hz condition: U = 12, p = 1,000; two-tailed t-tests for the comparisons in the 5 Hz and 20 Hz conditions: t(4) = -0,004, p = 0.996 and t(4) = 0.920, p = 0.384, respectively).

5.2.4 Discussion

In the present work, across several conditions, we investigated the effect of motion adaptation on the apparent number of elements in a visual display. While at first glance the relation between motion and number might seem counterintuitive, our aims were driven by recent theoretical instantiations suggesting the existence of a common generalized system processing information related to different “magnitudes” (ATOM, A Theory Of Magnitude; Walsh, 2003). We exploited motion adaptation because several classic and more recent studies pointed out that such adaptation could affect the representation of both space (Snowden, 1998; Nishida & Johnston, 1999; McGraw et al., 2002) and time (Johnston et al., 2006; Burr et al., 2007). According to the ATOM theory, information about different magnitudes – whether they are spatial, temporal or numerical – might be managed by means of a generalized magnitude system, which would exploit similar mechanism to process information in different domains, possibly located in (at least partially) overlapping neural substrates. Our predictions, according to such theory, concerned the possibility to observe an effect of motion adaptation on numerical estimates similar to what has been previously observed particularly on time perception.

Our results showed that numerical estimates could indeed be affected by motion adaptation, causing a numerosity underestimation similar to the effect on apparent time. First, such adaptation is spatially localized, affecting only stimuli located in the same position of the adaptor stimulus; otherwise, the effect should have extended also to the reference stimuli, cancelling out any net effect. Second, the effects observed on numerical estimates were only in terms of underestimation, while none of the conditions led to overestimation of perceived numerosity – similarly to what has been observed on time. Third, the most effective adaptor – the one causing the larger effect on numerosity – was fast (20 Hz) translational motion, while both slower (5 Hz) translation and complex motion caused smaller effects. However, especially regarding this latter point, results are not completely consistent with effect on perceived time. Indeed, while the motion-induced duration compression is similarly spatially localized (Johnston et al., 2006; Burr et al., 2007), it seems also very specific for the motion profile used to adapt and test perceived duration (Fornaciai et al., 2014; Fornaciai, Arrighi, & Burr, Under review; see Chapter 3). Conversely, our results on numerosity showed that, although to a lesser extent, slower translational motion and complex (circular) motion adaptation also caused an underestimation of perceived numerosity.

Overall, in the first part of the study where we tested translational and circular motion adaptation, we found an effect on perceived numerosity in all the conditions tested. Consequently, this raised a critical question that needed accurate control: given that the representation of the number of elements in the patches is strictly related to the possibility to detect and segregate the perceptual units (Anobile et al., 2015a), could this effect be driven by other lower-level effects affecting the perceived contrast or the visibility of the stimuli? According to this idea, the observed underestimation might be simply due to a reduced spatial resolution, which could have decreased the sensibility (or precision) of numerical estimates or made some of dots disappear due to a reduced visibility.

However, an increased difficulty in performing numerosity judgments – and a consequent reduction in precision – seems unlikely to explain the robust underestimation effects. Indeed, precision in the numerosity task improved after adaptation, with JNDs decreasing as the effect on accuracy (PSE) increased, showing that adaptation not only affected accuracy, but also sensitivity to numerosity. The relation between concurrent changes in JND and PSE is showed more clearly analyzing Weber Fractions – i.e. the measure of sensitivity (JND) normalized by perceived numerosity (PSE). As reported in the bottom panels of Fig. 5.5, both measures decreased as function of the adaptation magnitude, leaving Weber Fractions almost constant, as showed by the reduced slope of the regression line. Interestingly, this pattern of results has also been previously observed on saccades (Morrone et al., 2005) and on perceived time (Fornaciai, Arrighi, & Burr, Under review).

Regarding the possible decrease in perceived contrast that might have been caused by visual adaptation, we also performed two additional experiments to address this issue. First, we tested the effect of motion adaptation on stimuli with two different ranges of numerosity, one lower (3) and one higher (50). If the effect on numerosity observed in the first experiment was actually due to a direct modulation of numerosity representation, so we should expect two opposite patterns of results from these stimuli: while adaptation should affect the 50-dot patches approximately to same extent of 15-dot patches, adaptation should not affect the smaller stimuli. Indeed, several previous studies pointed out that very small numerosity – also termed "subitizing" range - is underpinned by a different mechanism, highly dependent on attentional resources, which is virtually immune to adaptation effect (Burr et al., 2010; Burr et al., 2011; Anobile et al., 2015b). Indeed, this is exactly what we found: while translational motion adaptation affected the 50-dot patches similarly to the previous condition, numerical estimates for stimuli in the subitizing range were always veridical, even after adaptation. Additionally, the lack of effect on such range of numerosity made unlikely the possible role of attentional effects, since the disruption of attentional resources should have made the stimuli prone to adaptation as the more numerous stimuli are.

Moreover, we also tested adaptation effects reducing the contrast of the adaptor stimuli, according to the idea that if the adaptation contrast is lower that test stimulus contrast, so the effect of adaptation on contrast itself would be small or absent (Georgeson, 1985). What we found in this condition seems fairly in line with our account, since fast translational motion caused again a robust effect on perceived numerosity. However, we did not observed a significant underestimation after 5 Hz adaptation, but the lack of effect in this condition might have been driven by an increased variability in participants' performances (see Fig. 5.7).

Finally, we also addressed the possibility that the observed effects might be related to spatial frequency adaptation masking the dot patches. To control for this possibility, we tested adaptation to translating gratings with a more different spatial frequency (0.5 cpd) compared to the size of the dots (0.3 deg), and also with static 0.5 and 1 cpd grating, according to the idea that effects at this level should be specific for the spatial frequencies profile of the stimuli, and should work even with a static stimulus. However, while we did not find any influence of static stimuli, 0.5 cpd gratings moving at 5 or 20 Hz caused a very similar effect on perceived numerosity, ruling out the influence of spatial frequency adaptation.

Overall, the effect seems strictly related to the representation of numbers, with adaptation somehow distorting it as a sort of recalibration. However, why motion adaptation should affect perceived number is not clear. One possibility is related to the predictions of the ATOM theory: spatial, temporal and numerical information might be processed by means of a common generalized magnitude system, sharing mechanisms and neural substrates (Walsh, 2003; Burr et al., 2011). According to such framework, motion – that is, a stimulus tightly related to both space and time – should affect all these three dimensions. And, indeed, this is what has been observed by several studies, comprised the present one (Snowden, 1998; Nishida & Johnston, 1999; Johnston et al., 2006; Burr et al., 2007).

However, while simple uni-directional translational motion has a similar effect on time and number, the effect of complex motion profiles seem less intuitive. Indeed, our results are only partially consistent with previous observation on time perception, due to the fact that the effects on perceived time are far more selective (Fornaciai, Arrighi, & Burr, Under review). Moreover, taking into account previous results about motion adaptation, we must note that even on the spatial dimension different motion profiles cause different effect: complex motion profiles cause stronger motion aftereffects (MAEs; Bex et al., 1999). It is difficult to compare the results concerning MAE with time and numerosity experiments, since methodologies to estimate the magnitude of the effect are different, but overall the emerging picture is that while "simple" motion adaptation can affect space, time and number, the effects of more complex motion profiles changes according to the domain examined.

Do these observation run against the ATOM theory? We think not. However, while the system processing different magnitudes could be the same, what we pointed out is that the mechanisms exploited to process different magnitudes might be partially different. But this difference is not so surprising, because while the ATOM theory assumes a common metric for the three domains, the specific mechanisms for their processing must be different, due to the specific nature of the information processed, and to the specific brain areas involved with different features of the stimuli along the visual processing stream.

5.2.5 Conclusion

Why complex motion has such different effects on different magnitudes is an interesting question that future studies should address. While we do not have a clear explanation for this difference, our data is however clear in showing a robust effect of motion adaptation of perceived number, which is not dependent on low-level effects such contrast, visibility or spatial frequency masking. Our results thus support the idea of a Generalized Magnitude System, but point out that different magnitudes might be processed exploiting different mechanisms, as showed by the different patterns of adaptation effects on different magnitudes.

5.3 Audio-visual simultaneity in a time to come: spatial predictability improves temporal order judgments

5.3.1 Introduction

Among the temporal properties that have been studied, one of the most important is the perception of temporal order, that is, the ability to determine which of two events happened first. This ability is important to ascertain which event is the cause and which one is the effect (Michotte, 1963; Leung 2002; White, 2006). Despite its importance, the perception of temporal order often deviates from veridicality because the perceptual latency of stimuli differs depending on the modalities involved (Allison et al., 1983) and on the properties of the stimuli (i.e., depends on stimulus intensity; Matteson, 1971), but also because perceived timing can be plastic and malleable. Agency, for example, binds events to their generating actions making the two to appear closer in time (Haggard, Clark, & Kalogeras, 2002). Repeated exposure to a constant asynchrony leads to a recalibration of perceived simultaneity that compensates for the physical delays (Fujisaki et al., 2004; Di Luca et al., 2009). Recalibration leads to the erroneous perception of sensory events preceding their causes once the asynchrony is removed (Cunningham et al., 2001; Heron et al., 2009).

Temporal information can be exploited not only to understand the causal structure of present events, but also to predict future ones, speed up reactions, and allow a precise control on the outcome of our actions (Thomaschke & Dreisbach, 2013). In everyday life, a surprisingly large portion of our perception is shaped by prior expectations (Summerfield & Egner, 2009; Seriès & Seitz, 2013) and the necessity of prediction in sensory processing becomes evident especially in time-critical behavior where sensory delays are one of the factors limiting performance (Wolpert & Miall, 1996; Wolpert & Flanagan, 2001; Roth et al., 2013). For instance, when it comes to crossing a road, the brain uses sensory information and prior knowledge about motion to predict whether (and when) a car will intercept us. Such predictions are continuously updated according to the changing dynamics of the environment. Recent investigations (Roth et al., 2013) further suggest that the mechanisms for recalibrating sensory predictions about external events is located in the cerebellum – a crucial substrate also for the processing of motor-related information and for the optimization of self-action perception (Blakemore et al., 2001; Bastian, 2006; Synofzik et al., 2008). These predictive abilities are highlighted in the remapping of visual receptive fields prior to eye movements (saccades; Duhamel et al., 1992). Predictions are driven by a copy of the motor command sent to the sensory cortices: during the preparation of a saccade, a portion of visual neurons shifts the location of their receptive field toward the predicted position after the movement. Such process anticipates the changes in incoming visual information, possibly allowing for the maintenance of visual stability (Sommer

& Wurtz, 2006; Wurtz, 2008).

Some studies suggest that predictability could facilitate motor performance, but not perceptual discrimination. Such difference in facilitation is possible because the mechanisms for perception and action function in different ways (Goodale & Milner, 1992; Goodale & Westwood, 2004). For example, Thomaschke & Dreisbach (2013) found that facilitation of speeded responses was effector specific, disappearing if the response required a different action. Such selectivity implies that temporal predictions are exploited at a post-perceptual stage (i.e. the motor stage). On the other hand, studies involving judgments of spatial properties have shown that uncertainty and predictability can influence perception. Heron and colleagues (2004), for example, demonstrated that an anticipated sound could bias the perceived position of a visual bounce, but only with degraded visual positional information. They further showed that also the contrary holds: while temporally well-defined sounds can strongly bias visual information, broader auditory stimuli have less of an effect. In other words, sensory uncertainty can modulate the extent by which different signals are combined and integrated. Spatial uncertainty has also been shown to influence perceived duration of visual events: the apparent duration of a brief sequence showing a stimulus moving along a discontinuous trajectory is underestimated when the sequence is reversed – an effect called the Memento effect. This effect was initially thought to be related to the segmentation of the reversed path in multiple visual events, instead of a single, smooth, motion trajectory (Liverence & Scholl, 2012). However, recently, it has been shown that the temporal underestimation of the reversed sequence is not linked to a segmentation process, but it depends on an increased spatial uncertainty (and thus reduced predictability) of the motion path (Meyerhoff et al., 2015).

Predictions and expectation not only can improve motor performance or bias the judgment of spatial and temporal features, but they could modify the content of perception. For instance, it has been shown that expectations can bias bistable stimuli toward one interpretation (Goolkasian & Woodberry, 2010), they can promote stimuli during binocular rivalry (Chopin & Mamassian, 2012), they can break the continuous flash suppression revealing the suppressed stimulus (Costello et al., 2009), they can facilitate object recognition (Eger et al., 2007), and they even can bias the interpretation of stimuli that are normally stable (Adams et al., 2004). The distorting effect of expectation has been also demonstrated in the temporal domain. For example, Pinto and colleagues (2015) recently showed that expectations might speed up awareness. The accelerated entrance into awareness might be due to the faster verification of predictions with expected stimuli, an effect consistent with the principles of the Predictive Coding theory (Rao & Ballard, 1999; Friston, 2005; Friston & Kiebel, 2009).

Since expectations and predictability play an important role in several sensory and motor tasks, here we wanted to investigate whether predictions about incoming events can also play a role in the perception of temporal order of multisensory events. The idea that the predictability provided by motion information might improve the sensibility to difference in timing of audio-visual stimuli is certainly not new (Van Eijk et al., 2008; Vroomen & Keetels, 2010; Cook et al., 2011; de Boer-Schellekens et al., 2013). Particularly, de Boer-Schellekens and colleagues (2013) showed that while sensitivity to audio-visual temporal order is generally impaired in subjects with autism spectrum disorder, some stimuli lead to better performances compared to other. Particularly, using clapping-hands stimuli results in a better precision of temporal order estimates – a result also consistent with a previous study by Stekelenburg & Vroomen (2007). According to the authors, such advantage in temporal order judgment precision for clapping-hands stimuli is due to the higher amount of visual predictive information, provided by the motion of the hand toward the other, compared to other audio-visual pairs such speech stimuli (lips movements and sounds) or artificial flash-beep pairs, where motion information is limited or completely absent. So, the availability of predictive motion information might allow a more precise definition of the temporal relation between the two events.

However, the critical comparison between predictable versus un-predictable stimuli in the work of de Boer-Schellekens and colleagues (2013) and of Stekelemburg and Vroomen (2007) comprised two separate factors: predictability and movement of the visual stimulus. Thus, the better precision for the clapping-hands could be a combination of motion information and temporal predictability. Indeed, apparent motion has been shown to positively affect the performance of temporal order judgments, improving the precision and speeding-up the processing of visual information even when predictability was not possible (Kwon et al., 2014). To the best of our knowledge, the role of predictions has not been analyzed separately from motion information.

To test the net effect of predictions provided by visual anticipatory information without the influence of motion, we used moving stimuli for both the predictable and the unpredictable information. We asked participants to judge the temporal order of two events: a change in trajectory of a moving stimulus ("visual event") and the onset of a brief sound ("auditory event"). In the predictable condition we cued the position of the change in trajectory so to test whether the brain is capable of exploiting the positional cue by combining it with motion information to anticipate not only where but also when the visual event is bound to happen.

5.3.2 General methods

Subjects

A total of 39 participants (28 females; age ranging from 18 to 30 years old) took part in the experiments after giving their written informed consent. They were compensated for their time with course credits or with a payment of 6 GBP/hour. All participants had normal or corrected-to-normal visual acuity and reported to have good hearing. With the exception of the author M.F., all other participants were naïve to the purpose of the experiment. Experimental procedures were conducted according to the protocol approved by the stem ethics' committee of the University of Birmingham, and are in line with the declaration of Helsinki.

Apparatus

Experiments took place in a quiet dark room. Observers sat in front of a monitor screen at a distance of about 60 cm, with head stabilized by means of a chinrest. Stimuli were generated using the routines of the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) and they were presented on a CRT monitor (NEC MultiSync FP2141sb) with a resolution of 800x600 and a refresh rate of 120 Hz. Sounds were played by speaker (Fostex PM0.4n) located on the left side of the monitor.

Procedure

Participants performed a Temporal Order Judgments (TOJ) task, discriminating which of two sensory events – a “visual event” and an “auditory event” – happened first in time. Each trial started with a white circle (1.5 deg diameter) moving horizontally either toward the right or toward the left. The starting coordinates of the circle (6.67 deg ± 0.5 deg horizontally and 3 deg ± 0.5 deg vertically with respect to the center of the screen) and its speed (average speed = 8.7 deg/sec ± 0.5 deg/sec) were random. After a variable amount of time (0.950 s ± 0.108 s), the circle reversed movement direction, and kept moving with the same speed until disappearing at the starting point. The change in motion trajectory represented the “visual event”. This event was paired to an “auditory event” – a brief sound (duration = 8.3 ms, frequency = 1000 Hz, intensity 65 dB), with a Stimulus Onset Asynchrony (SOA) randomized across trials (-350, -150, -100, -66, -33, 0, 33, 66, 100, 150, 350 ms, where negative values represent audio-leading stimuli pairs). Subjects were instructed to press a button after the circle had disappeared, indicating which of the two events happened first. Pressing the response key triggered the start of the next trial after 650 ms. Subjects completed 6 or 8 blocks of 44-88 trials (according to the condition) and were free to take a break between blocks. Before the experiment, participants completed a training session, comprising 45 trials (SOAs = -350, -250, -120, -100, 0, 100, 120, 250, 350 ms), to ensure that the task was properly understood.

The effects of different kinds of visual predictive information were investigated in three experiments.

5.3.3 Experiment 1

In Experiment 1, we tested the effects of cues on TOJ performance. The experiment comprised two parts, performed in separate sessions with the same order across participants. In the first part of the experiment we compared responses on basic TOJ task (un-cued trials) with randomly interleaved trials in which a visual cue indicated the exact position of the visual event (cued trials). The cue was a pair of black dots (“Dots cue”, diameter = 0.35 deg; inter-dots distance = 2 deg), aligned to the position in which the target circle will undergo the change in trajectory. The cues were present at the start of the trial and remained on the screen until the end of the sequence. In the second part of the experiment we tested the effects of conflicts between the cue and the position of the reversal. Trials with correct cueing were randomly interleaved with trials having cues suggesting an earlier or later position on the path on the moving circle, which corresponded to a 50 ms anticipation or delay.

Results

Fig. 5.9 shows the results of Experiment 1, both in the first and in the second part. Participants' perceived simultaneity was not consistently affected by the presence of the cues (paired sample t-test on PSS un-cued vs. dots $t(8) = 0.1$, $p = 0.94$). PSS values showed a slight reduction in all the cued conditions of the second part of the experiment (simultaneity was obtained with more anticipated auditory stimuli), but the comparisons did not reach statistical significance (paired sample t-test on PSS between un-cued and the three cued conditions, all p-values > 0.133) and the difference in cue displacement did not have an influence (one-way r.m. ANOVA on PSS in the three cued conditions $F(2,8) = 0.03$, $p = 0.97$). TOJ discrimination performance in un-cued trials was lower than in cued trials, as shown by a reduction in the JND of 34.8 ms ± 15.9 ms (Wilcoxon Signed Rank test, $p = 0.004$). Cues led to an improvement in participants' performance even when displayed displaced from the actual point of inversion (paired sample t-test on JND against un-cued: early $t(8) = 2.6$, $p = 0.015$; on-time $t(8) = 3.1$, $p = 0.006$; late $t(8) = 2.2$, $p = 0.026$). The difference in cue displacement did not affect performance (one way Repeated Measure ANOVA on JND in the three cued conditions $F(2,8) = 0.3$, $p = 0.74$).

5.3.4 Experiment 2

In Experiment 1, we found that information about where an event in bound to occur improves the performance in judging when the event occurs in relation to another event. In Experiment 2, we attempted to modulate the spatial reliability of the cue by changing the shape and dimensions of the dots to determine whether this modification influences the precision of temporal judg-

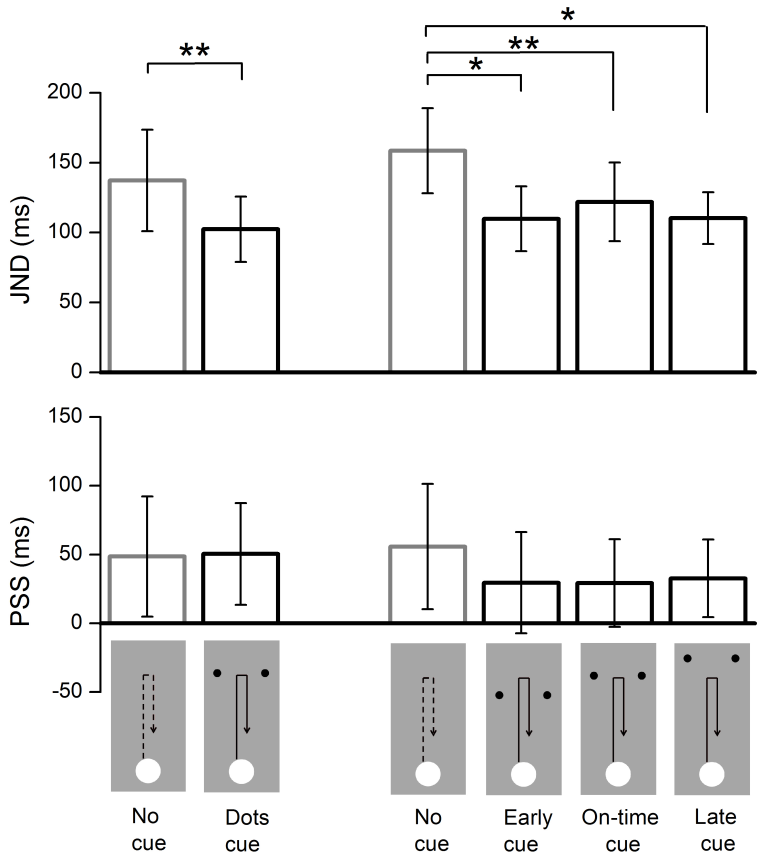

Figure 5.9: **Average JNDs and PSSs from Experiment 1.** (top) Average JNDs in Experiment 1. The left of the panel reports JNDs for the first part of the experiment, where performance in the TOJ task was tested with un-cued and cued trials (Dots cue). JNDs were significantly reduced in the cued condition. The right side of the panel reports JNDs from the second part of Experiment 1. Here, JNDs were significantly reduced irrespectively of whether the cue was in the correct position or slightly deviating from it. (bottom) Average PSSs in Experiment 1 (leftmost and rightmost respectively for the first and second part). We did not observe any significant change in subjects' PSS. Asterisks refers to the results of the statistical tests: * $p < 0.05$, ** $p < 0.01$. Error bars are S.E.M.

ments. Subjects' performance in the TOJ task was compared in interleaved conditions where either there was no cue or where different kind of visual cues were presented.

The experiment comprised two parts, performed in separate sessions and with different observers. In the first part of the experiment, we interleaved un-cued trials, with cued trials ("Dots cue" as in Experiment 1), and trials with a solid black bar ("Solid cue"). The area of the bar (1 deg^2) is larger than the one of the dots (0.096 deg^2 each), and whereas at the time of the change in trajectory the center of the circle coincides with the segment connecting the dots in the “Dots cue” condition, in the “Solid cue” condition the moving circle touches the side of the bar at reversal. The additional visual feature created by the contact should allow a finer estimate of the impact position. The second part of the experiment was equivalent to the first part except that we presented two ellipses (“Ellipses cue”) instead of the “Dots cue”. The area of the ellipses (0.56 deg^2 each) differs only slightly from the one of the bar. However, the ellipses do not enter in contact with the moving stimulus. The bigger area of the ellipses could increase the uncertainty about the correct position of the visual event (Heron et al., 2004) leading to a smaller improvement in performances compared to the Dots cue.

Results

Fig. 5.10 shows the results of Experiment 2. In the first part, the results replicate the improvement in time-order discrimination with the presentation of the Dots found in the first experiment (paired sample t-test on JND; un-cued vs. dots, t(18) = 4.0, $p < 0.001$), and we also found a comparable performance improvement with the Solid cue (paired sample t-test on JND; un-cued vs. solid, t(18) = 2.3, p = 0.018). Performance does not differ in the two cued conditions (paired sample t-test on JND dots vs. solid t(18) = -0.1, p = 0.89).

In the second part of the experiment, performance was again better with the Solid cue presentation (paired sample t-test on JND; un-cued vs. solid, t(7) = 3.3, p = 0.006), while no significant improvement was observed with the Ellipses cue (paired sample t-test on JND; un-cued vs. ellipses, t(7) = 1.2, p = 0.139). We also calculated the improvement in precision with each of the cues compared to baseline un-cued condition. On the one hand, the Bar cue provided an average improvement of about 28.7 ms (± 9.2), while with the Ellipses cue the improvement was about 10.6 ms (± 9.7). However, the difference between Bar and Ellipses cues did not reach statistical significance (paired sample t-test, t(7) = 1.9, p = 0.051). In neither part of the experiment we found a systematic difference in PSS across conditions (data not shown; one-way r.m. ANOVAs, p-values > 0.44).

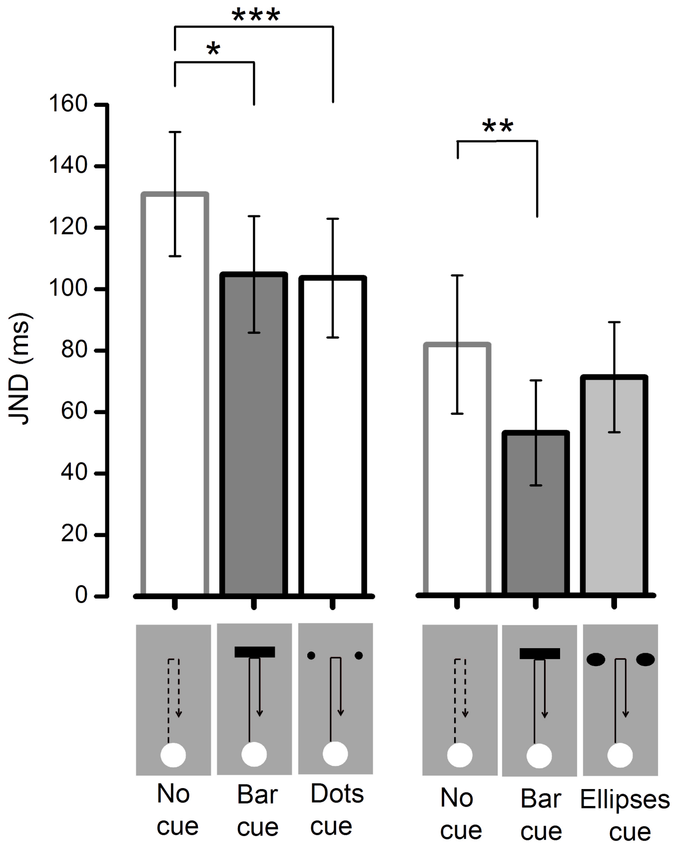

Figure 5.10: **Average JNDs for the two parts of Experiment 2.** (left) JNDs for the first part of Experiment 2, where we compared participants' performance in the TOJ task with un-cued trial, trials with the Solid cue, and Dots cue as in Experiment 1. Here both Solid and Dots cue exerted a significant influence of participants' JND, improving the precision in the task compared to the un-cued condition. (right) Second part of Experiment 2. Here we introduced a larger version of the Dots cue to increase the uncertainty (Ellipses cue). In this case JNDs were again reduced in the Bar cue condition, compared with the un-cued condition, but no significant change were observed in the Ellipses cue condition. Asterisks refers to the results of the statistical tests: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Error bars are S.E.M.

5.3.5 Experiment 3

In the third experiment, we investigate the reasons behind the performance increase we found. One possibility is that the cue provides a better estimate of the location of the event at the time of the auditory cue. An alternative possibility is that participants use the shrinking distance between the moving dot and the cue to predict the time of the reversal and use this prediction to improve the precision of their temporal estimate. To separate these two possibilities we present the cue only shortly before the reversal, thus still allowing the estimate of the location, but decreasing the ability to predict the reversal time. In every cued trial, one of the two cues (bar or ellipses) could appear, but only for a limited amount of time (140 ms prior to the visual event, and disappeared 140 ms after it).

Results

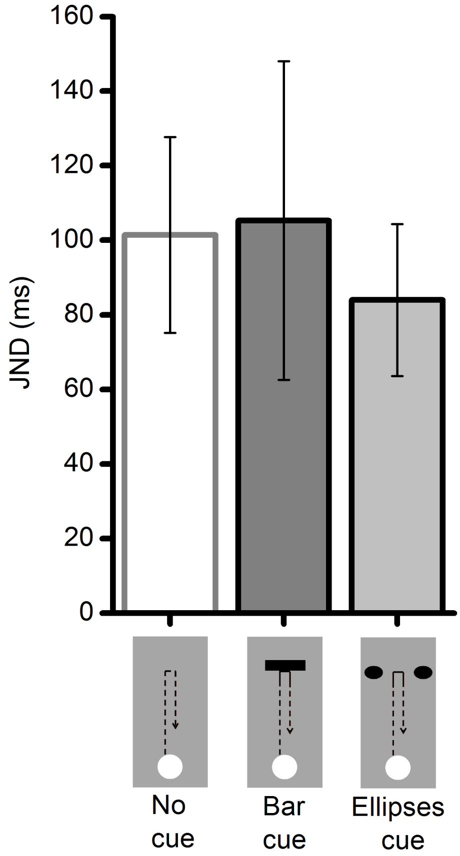

Figure 5.11: **Average JND in the three conditions of Experiment 3.** In the third experiment we intermixed un-cued trials with Bar and Ellipses cue, but presenting such cues only 140 ms before the visual event, so reducing the available time to perform a prediction about the location and timing of this event. Doing so, we observe no significant changes in participants' JNDs. Error bars are S.E.M.

In experiment 3 (Fig. 5.11), where cues were presented only for 280ms, we did not observe a significant difference between the un-cued vs. Solid cue and between the un-cued vs. Ellipses cue (paired sample t-test on JND between un-cued and solid t(5) = -0.2, p = 0.43; ellipses t(5) = 1.7, p = 0.071). Moreover, we did not find any systematic change in participants PSSs (data not shown; one way r.m. ANOVA, F(2,5) = 0.783 p = 0.48).

5.3.6 Discussion

The aim of this work was to investigate whether knowledge about where an event is expected to occur could improve the precision in judging when the event takes place. We tested this hypothesis by asking participants to perform a TOJ task that required judging the temporal order of a visual event (the motion reversal of a stimulus) and an auditory event (a beep), and we then compared the performance achieved with and without spatial cues indicating the location of the visual event. Moving stimuli seem particularly suitable for testing the predictability of the events, given the ability of the visual system to anticipate the future positions of moving objects (Rosenbaum, 1975; Peterken et al., 1991), to extrapolate their trajectories (Nijhawan, 1994; Krekelberg & Lappe, 1999; Murakami, 2001), as well as determining time-of-arrival (Ellingstad & Heimstra, 1969; Schiff & Oldak, 1990). Indeed, the visual system continuously predicts and extrapolates future positions of moving objects, possibly to compensate for neural transmission delays across the visual processing hierarchy (Nijhawan, 2008).

A paradigm involving visual movement has been used by Roseboom and colleagues (2009) to test the effect of multiple events on the window of audio-visual simultaneity. In this study, the authors presented a circle moving toward a static object, colliding with it and then returning to its initial position. Participants had to judge the simultaneity of the (visual) collision event with an auditory beep. They compared this baseline sequence with another condition in which the moving stimulus was accompanied by another preceding or succeeding moving object, actually creating an additional visual event. What they found is that the additional event had a modulatory effect on the temporal window of integration and it also enhanced sensitivity of simultaneity judgments.

In this study, the trajectory and timing of the moving stimulus were always predictable. Conversely, in our un-cued trials, participants could extrapolate the trajectory of the moving object but they could not determine in advance the position or the timing of the trajectory change that constitutes the visual event. Such lack of predictability (which is typical of most stimuli used in TOJ tasks) limits the precision with which temporal order can be judged. We expected to find an increase in precision if the brain can successfully exploit spatial information to perform a prediction about the expected time-of-arrival of the event from the cued position of the change in trajectory. This prediction is reinforced by the recent results of Kwon et al. (2014), which showed how an apparent-motion stimulus leads to better audiovisual TOJ performance than a simple flash. Results indicated changes in both PSS and JND suggesting a faster processing of visual information and a greater temporal resolution with apparent motion (Kwon et al., 2014).

Such an improvement in TOJ discrimination performance with a spatial cue was confirmed from the data of Experiment 1. Results also indicate that the reliability of the cue can modulate this effect: increasing the spatial uncertainty (Experiment 2) reduced the advantage provided by predictions. This is consistent with the findings of Heron and colleagues (2004), who found that uncertainty governs the extent of multisensory interactions. In a spatial alignment task, they tested the effect of a concurrent sound on the perceived "bounce" position of a moving stimulus with respect of two spatial markers, modulating the uncertainty both at level of the visual and the auditory signals. What they found is that by increasing the size of the positional markers – and so increasing their spatial uncertainty – the sound became more effective in biasing the perceived bounce position. This result fits our observations, particularly with the results of the second part of Experiment 2 where increasing the area of the visual cue ("Ellipses cue") strongly reduced the improvement in judgments' precision.

Our results also indicate that reducing the available time to make the prediction disrupts the observed facilitation (Experiment 3). From time-of-arrival estimation studies (Peterken et al., 1991) we know that reducing the visible time increases the error in temporal predictions. Our results thus indicate that an accurate prediction about the timing of the visual event is necessary for the improvement in performance, and that this improvement is not driven only by a better estimate of the position of the visual event at the time of the auditory event.

Regarding the perceived simultaneity of the events (PSS), we expected that cues would induce a facilitation in the processing of the visual event (as in Kwon et al 204). Instead, we did not observe consistent changes in perceived timing of the visual event driven by the spatial cues. There was a non-significant 26 ms tendency that perceived simultaneity is reached with the auditory event slightly anticipated when subjects had to cope with displaced cues in a block (Experiment 1, second part). The direction of this tendency suggests that with uncertainty there might be a perceptual anticipation irrespectively of the presence and direction of the displacement in the trial. The lack of robust effects on PSS with displacements, however, might be related to the small discrepancies used (which corresponded to 50 ms of anticipation/delay).

A differential role of predictive and postdictive processing of visual information on temporal order and simultaneity judgments was recently investigated by Van Eijk et al. (2010). This study showed that the PSS in a TOJ task is shifted toward the region of audio-visual asynchronies continuum where the observer's sensitivity is higher. Interestingly, to modulate the differential sensitivity for audio-leading or visual-leading asynchronies, they modulated the availability of predictive and postdictive information, using stimuli that resembled a Newton's cradle toy, and covering the leftmost or rightmost part. What they found is that removing predictive information by covering the initial part of the stimulus (where the first ball is released to impact with the other balls) leads to an increased sensitivity for visual-leading asynchronies caused by the availability of postdictive information, and also to a shift of the PSS toward the visual-leading region of the continuum. On the contrary, leaving visible the predictive information about the dynamic of the first ball and covering only the final part of the stimulus caused an opposite increase in sensitivity for audio-leading asynchronies and a shift of the PSS toward such part of the asynchrony continuum. How to explain that in the present study we did not observe significant changes in PSS, but only an increased precision in temporal order discrimination? Indeed, as pointed out in a subsequent study (Kohlrausch et al., 2013) the observed changes in sensitivity toward audio- or visual-leading asynchronies and the concurrent shift in PSS found by Van Eijk et al. (2010) could be attributed to an implied causal interpretation of the events. Such causal interpretation is unlikely to have happened in our experiment because with our Dots and Ellipses cues the moving stimulus never actually touched the cue. Accordingly, PSS did not differ between the un-cued and cued conditions. The only condition that was consistent with an interpretation based on a causal structure was the Bar cue condition in Experiment 2, but since it did not lead to a change in perceived simultaneity it might have also provided only weak cues. So, our results seem reflect only a facilitation driven by visual predictive information, not by cognitive influence of causal interpretation.

Another possible account for our results is based on the possibility that the improvement in precision merely reflects the ability to detect the event because of additional spatial information. According to this idea, even a cue flashed briefly before the visual event should improve the discrimination of temporal order, and the lack of effect observed in Experiment 3 (where the cues were presented only for a limited amount of time) might be related to the transient onset of the cues themselves rather than to the insufficient time to make a prediction. Transient stimuli are known to elicit burst of activity in the magnocellular pathways (Livingstone & Hubel, 1987; Maunsell et al., 1999) and to strongly attract a transient exogenous shift of attention toward them (Posner, 1980; Müller & Rabbitt, 1989). On the one hand, the sudden presentation of the cue just before the visual event could somehow saturate the magnocellular system, nullifying the improvement in discriminability of the onset of the subsequent transient visual event. However, the magnocellular activity is renown to be very fast, peaking shortly after stimulus onset (21-25 ms) and for a very short time (50-80 ms after the onset) (Maunsell et al., 1999). Given that our cues in the transient condition were presented 140 ms before the visual event, the activity elicited by the transients was not likely to saturate the magnocellular responses for a period sufficient to interfere with the onset of the visual event. On the other hand, it has been shown that the exogenous attraction of visual attention, while it can enhance spatial processing could reduce temporal resolution (Yeshurun & Levy, 2003; Yeshurun, 2004; Rolke et al., 2008) and affect performances in temporal order judgment tasks

(Hein et al., 2006; Nicol et al., 2009). Such impairment is thought to be related to a preferential facilitation of the parvocellular system and a concurrent inhibition of the magnocellular one, leading to the enhancement of spatial but the impairment of the temporal resolution of events (Yeshurun & Levy, 2003; Yeshurun, 2004; Yeshurun & Sabo, 2012). So, the attraction of transient attention driven by the sudden onset of the cue might have impaired the ability to judge the temporal order of the stimuli. However, we did not observe any significant difference in performance in the cued conditions of Experiment 3 compared to the un-cued baseline – a result that does not support an explanation based on such attention-related impairment in temporal resolution, and gives more credibility to the idea that spatial predictions improve performance in temporal order judgment.

How does this facilitation enabled by temporal prediction work? A possible explanation for its functioning can be sort by applying the Bayesian framework (Knill & Richards, 1996; Mamassian et al., 2002). We first observe that a moving stimulus is expected to continue moving in the same direction and with constant speed (representational momentum; Freyd & Finke, 1984). This could be captured in the Bayesian framework as an a-priori expectation that the stimulus will continue its spatiotemporal trajectory in the future. The information carried by the cues, moreover, can also be seen as the a-priori probability of where the event is about to happen. In the two experiments we modulated the presence and reliability of the prior distribution of the cues on a trial-by-trial basis. Because the posterior distribution reflects the contribution of any a-priori probability and of the sensory evidence, the presence of a cue indicating the position of the event improves the localization precision for the moving stimulus, thus improving temporal order precision. Even when the cue does not indicate the position accurately (cue displaced in earlier or later positions) the prediction appears to still be facilitated. At a first glance, such result appears to be inconsistent with the predictions of a Bayesian model, as it does not match with any change in perceived timing of stimuli as evidenced by the PSS. However, one should consider that – especially with attention focused mostly on the timing of the events – the spatial and temporal dimensions could be treated as independent and so providing a spatial cue with a bias could have in fact changed the perceived location of the moving circle by speeding up or slowing down its motion, without actually affecting the timing of the movement reversal. Increasing the cue's area using the Ellipses cue (in the second part of Experiment 2) clearly reduced cue effectiveness. In this case, prior probability distribution was possibly too broad to consistently influence the posterior distribution. On the other hand, we did not observe any difference using the Solid cue, compared to the Dots cue condition (in the first part on experiment 2), suggesting that the area of the cue is not the critical factor, whereas it is the precision of the point in which the event will take place, which with the bar cue is the juxtaposition of the circle to the rectangle. Besides the Bayesian framework, what are the possible neural bases for the observed improvement triggered by predictions about future events? Several frameworks take into account the role of predictability and expectations in shaping our sensory experience.

A very influential framework that takes into account predictions about sensory events is the Predictive Coding framework (Rao & Ballard, 1999; Friston, 2005). According to this theory, feedback projections from higher-level cortical areas carry predictions about the possible causes of the external events, while feed-forward projections to higher cortical areas transmit error signals related to violations of the predictions. As such, the encoding of sensory signals is mostly based on predictions about the incoming events. Several evidence support this view, particularly showing that expectations and predictions could facilitate and speed up the entry of visual information into awareness (Pinto et al., 2015). For instance, expectations might be created automatically by learning the statistical distribution of the stimuli during the experiment. Indeed, it has been recently shown that this process can actually take place, leading to the advantage of facilitating judgment of the more frequent stimuli (Gekas et al., 2015). Namely, Gekas and colleagues (2015) observed that participants' detection performance in a visual search task was affected by the probability of encountering a stimulus at a location: detection rates (and false alarms) increase for stimuli presented in the most frequent location. Such changes in performance show that when predictions match the actual sensory signal they can successfully facilitate the behavioral outcome – with predictions in this case allowed by the increased probability of targets being present in some locations rather than others. At the physiological level, some clues about how such facilitation may take place are provided by studies investigating the effect of probabilities and predictability on neuronal responses. Interestingly, it has been shown that predictability can modulate neuronal responses in the Superior Colliculus, a brain region known to be crucial for many spatial functions and particularly to the deployment of spatial attention (see Krauzlis et al., 2013 for a review), eliciting stronger responses for stimuli presented in the more probable locations (Basso & Wurtz, 1997; 1998). The fact that this influence is observed in areas critical for the allocation of attentional resources seems to suggest that the attentional modulation driven by expectations might be involved in the improvement at the behavioral level observed across several psychophysical experiments. Our results of an improved performance with cues are qualitatively consistent with this view, however the lack of a change in PSS due to conflicting cues and the lack of acceleration in the perception of the cued visual event are at odd with the foundations of the Predictive Coding framework.

Spatial attention has also been shown to modulate the extent of multisensory interactions, narrowing or enlarging the temporal window of integration (TWI). Donohue and colleagues (2015), for example, showed that directing the spatial attention to the stimuli biased the proportion of "stream" responses in a cued version of the stream/bounce illusion task (Sekuler et al., 1997; Shimojo & Shams, 2001) – which actually indicates a narrower TWI and thus a more selective integration of visual and auditory information for the impression of a bounce. On the other hand, they also showed an increase in "simultaneous" responses, when it comes to judge the simultaneity of the events, which is consistent with a widening of TWI. This shows that spatial attention possibly plays an important role in modulating the integration of visual and auditory signals, and that this influence depends also on the goals of the task. In our case, even if no changes are evident in the participants' perceived simultaneity, it might be possible that an enhanced allocation of attention toward the cued location (and hence also to the timing of the visual event) narrowed the TWI of the visual and the auditory event, reducing the variability and increasing precision.

Finally, since we provided anticipatory cues in the spatial domain, while the task concerned the temporal dimension of the events, the observed interaction between the two dimensions, which concerns the effect of spatial information over time-order judgments, seems consistent the ATOM theory, according to which spatial and temporal (as well as numerical) information might be processed by a common and partially overlapping mechanism (see Section 5.1; Walsh, 2003). As fundamental features of the sensory world, space, time and number might be encoded in a common map, possibly represented in the parietal cortex (Burr et al., 2010). Thus, influencing one of these features should be reflected in a change in the other two, and, indeed, several studies highlighted several commonalities and similar effects in space and time (De Long, 1981; Mitchell & Davis, 1987; Xuan et al., 2007; Rammsayer & Verner 2014; 2015) and similar distortions caused by, for instance, eye movements (Morrone et al., 2005; Burr et al., 2010; Binda et al., 2012). From such point of view, it seems reasonable to think that by providing a cue that indicates the spatial location of an event we also provided a cue about its timing: if the two dimension are encoded with a similar metrics, so spatial information should be suitable to be exploited in order to improve judgments in the temporal dimension. However, the lack of a change in PSS with conflicting spatial information seems at odd with this idea. One possibility is that in this specific task, in which participants focused on the temporal dimension, the integration of spatial information with the perceived timing could be somehow weakened, reducing the influence of fine spatial details (such the slight displacement of the cue) on perceived simultaneity.

5.3.7 Conclusion

Our results show that predictions and expectations about future events play a role in the perception of the order of two sensory signals. Moreover, the facilitation in time-order judgments allowed by the positional cue suggests a link between spatial and temporal information, with the spatial cues exploited to improve temporal perception, and that the two dimensions might be encoded with a similar metrics.

5.4 Audio-visual asynchrony as a cue for object size and visual distance in depth

5.4.1 Introduction

Can we tell the size of an object basing our judgment only on its retinal projection? In most of the cases, the answer to this question is no, especially for very distant objects. The moon is a typical example: our perceived size of the moon is so inaccurate that can be easily subject to illusions, such the "moon illusion", due to variation in its perceived distance (e.g. Kaufman & Kaufman, 2000). And what about perceived distance? While perceived size is thought to be based on extra-retinal cues, the distance of an object in depth is usually considered to be entirely dependent on visual information, such on occlusion, or binocular disparity.

Apart from the purely visual cues that our brain exploits to calculate the distance of an object or an event, other information is out there that could, in principle, help our judgment. For instance, how can we roughly calculate the distance of lightning, followed by its corresponding thunder? – We can just take the difference in the time of arrival of light and sound, and divide it by the ratio of their speeds. Indeed, light and sound have very different speeds, so due to this difference, especially for very far events, they arrive at an observer with different latencies: for audio-visual events near the observer, the discrepancy in arrival time between light and sound is very little, but for increasing distances become significant. So given that we can roughly calculate the distance of thunderstorm only from the difference in time of arrival of light and sound, might this difference – that is, the audio-visual asynchrony – be taken into account by the brain to disentangle the distance at which an audio-visual event takes place?

What follows from this idea, is that if you have a cue about the distance of an object, you might infer its size. This is exactly what was first tested by Jaekl and colleagues (2011): could asynchrony between the auditory and visual information bias the perceived size of an object? Presenting audio-visual stimuli, they measured the matched size of an asynchronous pair with a synchronous one, modulating audio-visual asynchrony. What they found is that the perceived size of a visual stimulus is increased by the addition of a certain asynchrony to a concurrent sound, compare by an identically sized object presented with a synchronous sound – a result that is consistent with the fact that if two objects has the same retinal projection but different distances in depth, the farther object must be bigger in size. Particularly, asynchrony in terms of audio-leading pairs (an unlikely condition in the natural environment) had no effect on perceived size, while in audio-lagging conditions the effect increased from 20 ms delay toward 60-80 ms, and remained stable (or even slightly decreased) for bigger asynchronies. So, these results showed that audio-visual asynchrony might be exploited by the brain as a cue for distance in depth, and this, in turn, induces also a bias in the perceived size of an object.

Moreover, Jaekl and colleagues (2015) tested more directly the hypothesis concerning the role of audio-visual asynchronies in depth perception, trying to ascertain if such cues could bias or improve sensitivity to visual distance in depth. First, they tested the hypothesis presenting three-dimensional dots clouds accompanied by sound with varying asynchronies, and asking participants to judge their relative distance. Interestingly, what they found is that dots cluster accompanied by asynchronous sounds (with sounds delayed compared to visual onsets) are systematically perceived as more distant compared to dots clusters accompanied by leading sounds. They also confirmed that a delayed asynchronous sound is actually sufficient to bias perceived distance. And they found that when it comes to judging the distance of audio-visual stimuli with ecologically congruent sound delays, precision for distance judgments is increased. Importantly, they also demonstrated that the audio-visual delays used in the experiments were not detectable by participants, ruling out the possibility that the results might have been driven by a response bias, thus concluding that audio-visual asynchronies might represent a genuine cue for visual distance in depth.

The possibility that the difference in time of arrival between the auditory and visual component of a multisensory event might provide useful information about the distance of an event was also tested not only on humans, but also on frogs (Halfwerk et al., 2014). Male Tungàra frogs are used to emit sounds from ponds to attract females and repel rival males. The vibrations created by the production of these sounds also create ripples in the surface of the water, so that giving rise to another signal that may be used to infer the distance of other animals, and creating the opportunity to compare the arrival time of sounds and ripples that are simultaneously generated. Indeed, what Halfwerk and colleagues (2014) found, is that male frogs' behavior depends on the distance of other animals, with emitted sounds' intensity increasing as the distance of other males increases. Moreover, they went on showing that frogs can discriminate between near and far rivals basing on unimodal cues, such changes in sounds and water ripples, as well as cross-modal cues, such the difference in time-of-arrival of sounds and their simultaneously-generated water ripples, strongly supporting the idea that timing difference between signal in different modalities might actually be exploited to judge the distance of multisensory events.

Overall, these results seem in line with the idea of a common mechanism for the processing of spatial, temporal and numerical information (Walsh, 2003; Burr et al., 2011). Indeed, according to this idea, different kinds of information would be processed with a similar metrics, making them available to be exploited also in different contexts – such as using timing information to improve spatial judgments.

Here we wanted to further test whether cues for visual distance in depth might be extracted from the difference in time-of-arrival of visual and auditory information, on human participants. In the first experiment we attempted to replicate the finding of Jaekl et al. (2011), which showed that audio-visual asynchrony can bias the perceived size of an object, via a change in its perceived distance. In the second experiment, we went on to test whether audio-visual asynchrony could be exploited by the brain to judge the distance of an event, providing a meaningful and useful cue to visual distance, similar to what has been showed by Jaekl et al., 2015. If audio-visual asynchrony actually represents a cue for the distance in depth, manipulating the onset difference between sound and light of a multisensory stimulus, we might observe a bias in estimating the distance of this audio-visual event in a 3D environment. Finally, in the third experiment, we investigated again the role of audio-visual asynchrony as a depth cue using a simpler experimental setup, but with increased uncertainty about the depth dimension. Namely, we used 2D stimuli viewed monocularly, asking subjects to judge the perceived distance of audio-visual pairs. Our hypothesis is that, if asynchrony is actually a useful cue to disambiguate the distance of an audio-visual event in depth, so it may be possible that in conditions of great uncertainty – such judging monocularly the distance in depth of a 2D stimulus, asynchrony may bias the judgment even more than in 3-dimensional conditions.

However, we did not succeed in replicating either Jaekl et al.'s (2011) or Jaekl et al.'s (2015) findings, and we did not observe any effect even in the 2D condition, where we expected a more robust effect due to the increased uncertainty. Conversely, we found hugely variable performances across participant, and no clear indication of a perceived size modulation due to audio-visual asynchrony.

5.4.2 Experiment 1 – The influence of audio-visual asynchrony on perceived size

Subjects

Four observers participated in the experiment (1 female, age ranging from 25 to 28 years). With the exception of one of the authors (M.F.) who participated in the experiment, all the observers were naive to the purpose of the study. All the participants reported having normal hearing and normal or corrected-to-normal vision. Experimental procedures were approved by the local ethics committee, and are in line with the declaration of Helsinki.

Apparatus and stimuli

All the experimental measures were performed in a quiet and dark room. Observers sat in front of a monitor screen at a distance of 3 m, with head stabilized by means of a chinrest, with the right eye occluded by a dark-colored patch. Stimuli were generated using the routines of the Psychophysics Toolbox (Kleiner et al., 2007) for Matlab (version r2010b, The Mathworks), and presented on a CRT monitor (Barco Calibrator Line), with a resolution of 800x600 and a refresh rate of 60 Hz. Stimuli were dark gray circles (luminance = 1.06 cd/m^2) presented on a black background (luminance = 0.17 cd/m^2). Auditory stimuli were 1000 Hz tones presented by means of two speakers located beside the screen. Responses were collected by means of a standard keyboard.

Procedure

Participants performed a visual size discrimination task. Trials started with subjects fixating at a central fixation point. After a 500 ms delay, the fixation point was extinguished, and the first circle (probe) was presented for 200 ms at the center of the screen, followed, after an inter-stimulus interval of 1000 ms, by the second circle (reference). Reference stimulus' size was kept fixed (diameter equal to 1.3 deg), while the probe stimulus' size was varied by an amount randomly chosen from trial to trial (size difference ranging from -0.1 to +0.1 in 0.01 increments, with respect to the reference's diameter). Both stimuli were accompanied by a concurrent irrelevant sound, always synchronous for the probe stimulus, or with a variable delay (from 0 to 120 ms in 20 ms increments) for the reference stimulus. The possible audio-visual asynchronies were chosen only in terms of an auditory delay, since we wanted to test only realistic combinations of audio and visual information. Subjects were instructed to indicate which stimulus appeared to be bigger in size, pressing the appropriate key at the end of each trial.

Results

Figure 5.12 summarizes the results of the size discrimination task. Overall, PSEs showed a systematic underestimation of the reference's size, irrespective of the audio-visual asynchrony. Indeed, we did not find any systematic variations in perceived size linked to the different audio-visual asynchronies (one-way repeated measure ANOVA on factor “asynchrony”: $F(3,6) = 1.223$, $p = 0.340$).

5.4.3 Experiment 2 – Audio-visual asynchrony as a cue for visual stimuli in a 3D environment

While the first experiment concerned perceived size, in Experiment 2 we aimed to test directly the role of audio-visual asynchrony in depth perception, using 3-dimensional stimuli. With three different task, we investigated whether and how audio-visual asynchrony might bias perceived depth and perceived size of the stimuli, controlling for the discriminability of audio-visual asynchronies.

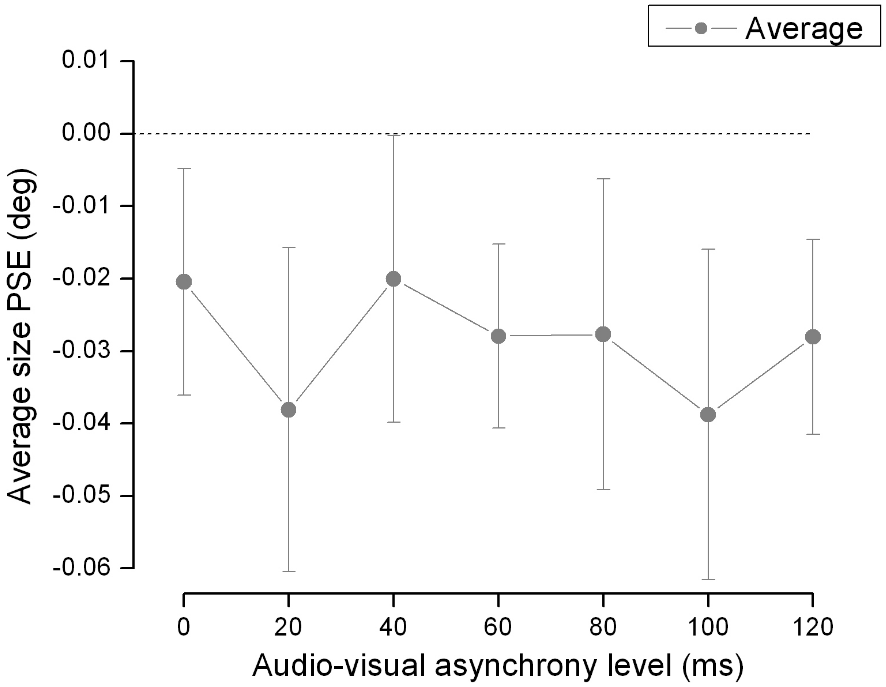

Figure 5.12: **Average size difference PSEs.** The average size-difference PSEs as function as the audio-visual asynchrony. The dotted line indicates the point of objective equality. Data shows that participants systematically underestimate the size of the reference stimulus – a result that is actually on the opposite direction compared to what was expected. Audio-visual asynchrony does not exert any significant effect on perceived size. Error bars represent S.E.M.

Subjects

Three observers participated in the experiment (1 female, age ranging from 25 to 28 years). With the exception of the author (M.F.) who participated in the experiment, all the observers were naive to the purpose of the study. All the participants reported to have normal hearing and normal or corrected-to-normal vision. Experimental procedures were approved by the local ethics committee, and are in line with the declaration of Helsinki.

Apparatus and stimuli

Visual stimuli are briefly flashed filled circles, presented inside a 3-dimensional environment resembling a “cage” or a “corridor”, built by several dots arranged radially in depth. Both the dot cage and the visual stimuli are modulated in stereo difference, in order to provide 3D perception through red-blue anaglyph glasses. Stimuli were presented on a sound-permeable screen of 224 cm by 140 cm, by means of a projector running at 60 Hz. The position of the subjects was fixed 130 cm away from the screen. Sounds are brief 1000 Hz, 70 dB tones, presented by means of speakers located behind the screen, in positions congruent to the location of the stimuli on the screen. Stimuli were bright disks with radius equal to 10 deg, modulated in stereo difference in order to appear inside the 3D environment, at about 200 cm from the participants.

Procedure

The experiment started with subjects fixating at the central fixation point. After a delay (500 ms), the first audio-visual stimulus was presented on the left side of the fixation point, with a horizontal eccentricity of 60 deg. After a variable Inter-Stimulus Interval (ISIs = 667ms, 700ms, 733ms, 767ms), the second stimulus appeared on the right side of the fixation point. Both visual stimuli were accompanied by a sound. One of the two audio-visual stimuli (randomly chosen from trial to trial) presented a given asynchrony between the visual and auditory components (10ms, 30ms, 50ms, 70ms, 100ms, 150ms, 200ms, 250ms – always visual first), while the other one is synchronous. Subjects were instructed to judge the stimuli according to three different tasks. In the main task (depth task), subjects had to indicate, at the end of each trial, which stimulus seemed more distant in depth. In the second task, we asked participants to indicate which stimulus seemed bigger in size (as in Jaekl et al., 2011). Finally, we asked participants to judge audio-visual asynchronies, indicating which audio-visual pair seemed more asynchronous. Stimuli and procedure were the same in all the different cases, and the order of the tasks was counterbalanced across participants. Participants completed at least 160 trials for each condition.

Results

What is immediately evident inspecting Fig. 5.13 is that, on average, depth and size judgments were almost at chance level. JNDs for all the three condition were quite high, spanning almost the entire range of the asynchronies,

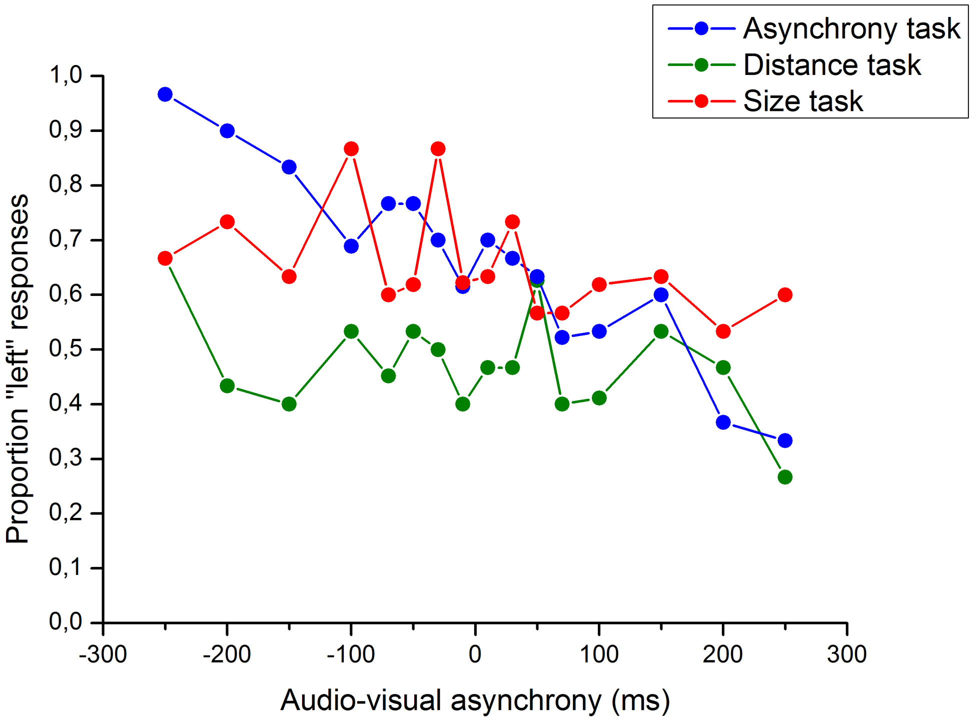

Figure 5.13: **Average results of Experiment 2.** Lines depict the probability of response "left" as function of the asynchrony, in the three different tasks. Note that negative asynchronies represent the cases where the asynchronous audio-visual pair was presented on the left; positive values correspond to the opposite case, where the asynchronous pair was presented in the right hemifield. The blue line describes data from the asynchrony task, where participants were asked to discriminate which audio-visual stimulus seemed more asynchronous (y-axis label correspond in this case to the probability of "left more asynchronous" responses). The green line reports data from the distance task, where participants discriminated which stimulus seemed more distant in depth (proportion of "left more distant" responses, in this case). The red line report data from the size task, where participants judged the apparent size of the stimuli ("left bigger", in this case). While the blue line describing asynchrony data shows that participants could discriminate different asynchronies, in the other cases response probability appear to be almost at chance level.

suggesting that uncertainty in judging the stimuli was so high that even in the most extreme cases the participants did not perceive any clear modulation of such features. Interestingly, even for synchrony judgments the average JND was quite high. This could suggest that in our setup – possibly due to the nature of the stimuli – even discriminating the audio-visual asynchrony was somehow difficult for the participants. A trend suggesting a better discrimination performance for the asynchrony task seems evident in the figure, but, however, comparisons did not reach statistical significance, but only showed a trend indicating that participants performed better in judging the asynchrony of the stimuli, while judging the distance or size was more difficult, as reflected by an average JND of 219 ms, compared to 291 ms and 316 ms, respectively for asynchrony, distance in depth and size judgments (paired sample t-tests, asynchrony vs. distance, t(2) = -2.784, p = 0.054; asynchrony vs. size, t(2) = -2.685, p = 0.057).

5.4.4 Experiment 3 – Cues for visual distance in a 2D environment

In both Experiment 1 and 2 we failed to find a relation between the asynchrony between auditory and visual component of a multisensory event and its perceived size or distance in depth. While the difference in time-of-arrival of auditory and visual information was hypothesized to convey a useful cue to size and distance, our paradigm did not show any systematic modulation. Previous studies, on the contrary, found instead clear evidence suggesting that asynchrony might concur to disambiguate these two features, both in human (perceived size, Jaekl et al., 2011) and in frogs (distance, Halfwerk et al., 2013). Our lack of effect, on the one hand, might be due to a reduced sensibility to such changes related to our specific experimental setup, since the expected effect is very small. On the other hand, it could be that other cognitive factors might play a role, overcoming this kind of information. For example, it is possible that knowing the real size of the testing room, or the fact that stimuli were not actually modulated in depth, forced the brain to discard information about audio-visual asynchrony.

In Experiment 3 we tried again to test the effect of audio-visual asynchrony on perceived distance in depth, using a simpler setup, but with increased uncertainty. Namely, we presented a group of flickering dots accompanied by a series of sound modulating audio-visual asynchrony, in a 2-dimensional environment.

Subjects

Three observers participated in the experiment (1 female, age ranging from 25 to 28 years). With the exception of one of the authors (M.F.) who participated in the experiment, all the observers were naive to the purpose of the study. All the participants reported to have normal hearing and normal or corrected-to-normal vision. Experimental procedures were approved by the local ethics committee, and are in line with the declaration of Helsinki.

Apparatus and stimuli

Subjects sat in a dimly illuminated room, in front of a Barco CRT monitor, subtending 40x30 degrees of visual angle (at a distance of about 57 cm). The experiment was performed monocularly (with right eye closed by means of a small cardboard patch, mounted on a plastic glass structure, worn by subjects for the entire duration of the experiment). Stimuli were presented in the context of a simulated "corridor" or "cage", built by means of a dots pattern displaced radially from a squared background (the "bottom" of the environment). Viewing was monocular and there was no stereo modulation, so the dots structure was only a 2D simulation of the 3D structure used in Experiment 2. Subjects were instructed to consider the stimuli as presented in the middle of a 3D corridor, even if a real 3D structure was not actually perceived.

Visual stimuli were five-dot arrays (dots diameter = 0.5 deg). Initial dots positions were randomly assigned, with dots placed in a circular area with 2 deg radius, and with a minimum distance from each other equal to 0.2 deg. All the dots in the patch were presented simultaneously, for a total duration of 216 ms, but with positions recalculated every 83 ms. The auditory component was a stream of five 1200 Hz tones, with a duration of 16 ms each, separated by 30 ms pauses, presented by means of two speakers located on both sides of the monitor screen. Dots were presented in parallel, while sounds were presented serially, but altogether audio-visual stimuli give the impression of a flickering/fluttering whole.

Procedure

Trials started with the subjects fixating at the central fixation point, located in the middle of the simulated corridor. After a delay of 500 ms, the first audio-visual stimulus was presented on the left side of the screen. After a variable ISI (667ms, 700ms, 733ms, 767ms), the second stimulus was presented on the right of the fixation point. In each trial, one of the two audio-visual stimuli was always synchronous, while the other one (the left or the right one, randomly chosen from trial to trial) presented a variable asynchrony (10ms, 30ms, 50ms, 70ms, 100ms, 150ms, 200ms, 250ms – always visual first). At the end of the trial, subjects had to indicate which stimulus seemed closer in depth to him ("distance" task), or which audio-visual stimulus seemed more asynchronous ("asynchrony" task). These two tasks were performed in separate sessions, with the order counterbalanced across participant. Responses were collected by means of a standard keyboard, and pressing a key triggered automatically the start of the next trial.

Results

Figure 5.14 shows the results of third experiment. Similarly to the previous experiment, what we found is a very large average JNDs values, suggesting that also in this case judging the relative distance in depth was very difficult, and participants could not reliably discriminate which one was closer or farther in depth – as evidenced by an average JND spanning almost the entire range of asynchrony (363 ± 639 ms). On the other hand, participants could reliably discriminate audio-visual asynchronies, and performance on this task was significantly better compared to the distance task (average JND = 159 ± 31 ms; paired sample t-test asynchrony vs. distance task, t(2) = -4.255, p =0.0255).

5.4.5 Discussion

In this study we investigated the possibility that the asynchrony between the visual and auditory components of an audio-visual event might be exploited by the brain as a cue for the distance of the event itself. Indeed, the difference in speed of the two kinds of information creates a disparity in the time of arrival of the two signals, which increases with the distance (with visual information arriving always first). This possibility was raised by some authors, which tested it first in the form of biases in perceived size (Jaekl et al., 2011) and then directly in the form of a bias on perceived distance (Jaekl et al., 2015). Moreover, the strategy of using sound delay as a cue for distance has also

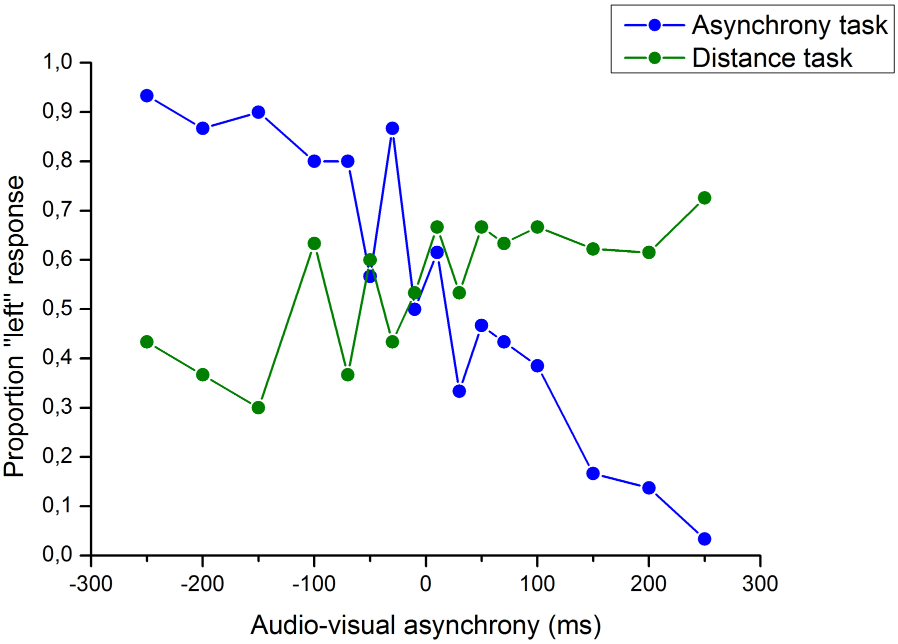

Figure 5.14: **Average results of Experiment 3.** Lines depict the probability of response "left" as function of the asynchrony, in three different tasks. Note that negative asynchronies represents the cases where the asynchronous audio-visual pair was presented on the left (while the stimuli presented in the right hemifield were synchronous); positive values correspond to the opposite case, where the asynchronous pair was presented in the right hemifield. Blue line describes data from the asynchrony task, where participants were asked to discriminate which audio-visual stimulus seemed more asynchronous (y-axis label correspond in this case to the probability of "left more asynchronous" responses). The green line reports data from the distance task, where participants discriminated which stimulus seemed more distant in depth ("left more distant", in this case). As showed by the trends indicated by the proportion of responses, while participants seem to have performed the asynchrony task with a reasonable degree of precision, performances in the distance task are almost at chance level (response probability never reaches 1 or 0, even with the larger values).

been demonstrated in animals (frogs), which can use these cues to modulate their behavior (Halfwerk et al., 2014). These are very intriguing effects, that shed light on the numerous strategies that might be exploited by the sensory systems to determine the properties of the external events. Moreover, the possibility of exploiting timing cues to judge purely spatial features seems in line with theories proposing a common mechanism for the processing of spatial, temporal and numerical information (Walsh, 2003; Burr et al., 2011). So, the aim of this study was first to replicate previous results concerning the biases on perceived size and distance, and then to extend them to circumstances where the uncertainty about depth is even more accentuated.

However, we did not succeed either in replicating previous results, or in demonstrating such effect in a more uncertain context.

Our lack of effect could be due to various reasons. Firstly, regarding Experiment 1 (perceived size), the experimental setup was not optimized as well as the setup used in the original paper (Jaekl et al., 2011): the distance from the screen was only 3 m, instead of 5 m, and this could be a limiting factor considering the very small variation in size (increments of 0.01 deg). Indeed, variations in steps of 0.01 deg correspond to a change of 0.6 arc min, which is actually about the effect size found by Jaekl and colleagues (2011). Second, the room was not completely dark, due to the presence of some LED lights (e.g. from speakers, keyboard and from the PC itself). Third, subjects were aware of the actual distance from the screen, because they were introduced in the experimental room with the lights turned on, and possibly of the fact that stimuli were always presented without any real variation of the distance in depth. Overall, it might be possible that our stimuli and our paradigm were simply not suitable to detect such a small effect. On the contrary, Jaekl and colleagues' experimental setup was much more controlled: they performed the experiment in a completely dark room, taking participants into the room blind-folded to prevent them from knowing the real distance from the screen. Moreover, in their setup, the distance from the screen was 5 m – a distance that could more easily allow modulation of the size of the stimuli maintaining the variations in visual angle very small. Particularly, at our 3 m distance any given size change in terms of visual angle was represented by a smaller change in terms of the number of pixels, compared to a 5 m distance, possibly making the task even more difficult. So it seems that if the participants know the room in which the experiment is performed and the actual distance from the screen (and probably know those stimuli are presented without any real variation in their distance in depth) differences in time-of-arrival of the visual and auditory information might simply not be taken into account to judge the size of an object.

Consistent with the negative result of Experiment 1, in which we failed to replicate previous observation by Jaekl and colleagues (2011) about a modulation in perceived size, even in this further experiment we did not find any clear indication that audio-visual asynchrony might be exploited by the brain to disambiguate distance in depth. Statistical comparisons showed only a non-significant trend, indicating that while participants could judge the asynchrony of the audio-visual stimuli, judging distance in depth and size was much more difficult. The lack of significance may be due to the small number of participants tested, and while this does not allow us to draw strong conclusions about the possibility that asynchrony cues might be exploited to judge the relative distance in depth or the size of an audio-visual event, it does suggest that the effect, if it exists, could be very small, and limited by the fact that in the absence of any real variation in distance or size the task could be too difficult for the participants. However, the fact that we tested only a very small group of participants is a limiting factor that does not allow drawing robust conclusions from our data. On the one hand, it appears to be an effect not so simple to replicate, but, on the other hand, since it is not a huge effect, it might be evident only testing a large amount of naive subjects.

Regarding the third experiment, exacerbating the uncertainty about the depth dimension of the stimuli, we expected to somehow increase the magnitude of the effect. However, we did not observe any effect of asynchrony on perceived distance. Actually, distance judgments about our stimuli were almost at chance level, which indicates that a modulation of perceived distance was not perceived at all by our participants. Again, the small group size does not allow us to conclude that the effect does not exist, but only that with our experimental setup we were unable to find it.

We conclude that, while the possibility to exploit asynchrony cues to judge the size or the distance of a visual event is a very intriguing sensory feature, it could be limited to only a small range of circumstances, while more often purely visual (e.g. stereopsis) or cognitive factors (e.g. knowing the actual distance from the screen) might overcome the auditory timing cues.

Chapter 6

General discussion

Given that some of the perceptual mechanisms exploited by the brain to process spatial, temporal and numerical information are still not completely understood, and that a unitary account for their mechanisms, as suggested by theories proposing common mechanisms, still misses some details to be complete, we investigated the sensory mechanisms for the processing of such features, as well as their possible intertwining and the mutual interactions among them. Across the experiments presented in this work, we tried to provide some new evidence to reinforce the understanding of different perceptual mechanisms, testing the predictions of several recent theories about the domains of spatial, temporal and numerical perception.

In the first part of this research (Chapter 2) we took into account the processes responsible for visual stability. Keeping vision stable is probably one of the most complex and important tasks that the brain has to perform, in order to achieve a correct representation of the external world and provide us with a continue and seamless visual experience. Studying the patterns of eye movement distortions in totally inexperienced subjects we demonstrated that such distortions in perceived space – which lead to the erroneous localization of stimuli briefly presented around the onset of an eye movement – is genuinely linked to saccade execution, and mostly determined by the ocular-motor parameters. Moreover, we also demonstrated robust effects of perisaccadic mislocalization even testing expert participants in a different task promoting an increased variability and reducing the stereotyping of the eye movements dynamic. We thus ruled out the possibility that the pattern of saccadic effects usually observed might have been driven by other processes, such practice and the level of automaticity of eye movements, due to the large number of saccades performed by subjects in such highly unnatural experimental paradigms. Overall, we propose that the pattern of perisaccadic distortions, resembling a compressed spatial metric, is an index of the underlying processes related to predictive remapping and visual stability. Additionally, it may be possible that this compressive effect – which has been demonstrated in the spatial, temporal and numerical dimension – might be related to the updating of a generalized magnitude system across eye movements (Burr et al., 2011).

Regarding time perception (Chapter 3), we first investigated the possibility of extending previous results concerning the effect of motion adaptation on apparent time to more complex motion profiles, like radial or circular motion. Motion-induced adaptation has been previously demonstrated using almost only simple, uni-directional, translational motion (Johnston et al., 2006; Burr et al., 2007). Our data, however, showed that the adaptation-induced duration compression does not generalize to more complex motion profiles, and not even to motion stimuli containing two motion directions at the same time –remaining limited to the case in which only one direction is displayed at any given time, during both adaptation and test phases. These results strongly constrain the theoretical instantiations trying to address the effect of adaptation on time proposing a very low-level neural locus (Johnston et al., 2010). Indeed, it has been proposed that such distortions in apparent time might be driven by a change in the temporal dynamics of magnocellular neurons, as early as in the lateral geniculate nucleus. Our results, however, run against this point of view, given that all the different stimuli used in our experiment were locally very similar, and undistinguishable at the early level of visual processing. The difference between simple and complex motion profiles arise from the need, in the latter case, to integrate local signals across large portions of the stimulus into a global percept (Morrone et al., 1995; Burr et al., 1998), which could be done only at relatively high level in the visual stream (area MST in the monkey, hMT+ in the human brain; Duffy & Wurtz, 1991; Duffy & Wurtz, 1997; Morrone et al., 2000). While our results do not allow us to precisely address why different motion profiles cause such different effects on apparent time, theoretical frameworks and model of time perception should take into account also such robust difference in order to better describe the machinery of time perception.

We also directly tested the predictions of a recent theory about the perception of time – that is, the State-Dependent Network (SDN) model (e.g. Buonomano & Maass, 2010), in order to figure out whether such model could work similarly in different sensory modalities. Interestingly, our results suggested that one of the critical predictions of this theory – the need of a minimum interval between two consecutive durations (inter-stimulus interval, ISI) in order to encode them properly (Karmarkar & Buonomano, 2007; Buonomano et al., 2009) – differs in different modalities. Indeed, while the disruptive effect of a too short ISI in the auditory modality seems limited to very short durations, around 100 ms, in the visual modality the effect seems to be present even at longer durations, around 300 ms. Our results overall support the SDN theory, and add further details to the model: while the machinery of the SDN model is thought to be exploited only for short durations, we propose that the working range of the proposed processes is not a fixed property, but depends on the temporal resolution of the specific sensory modalities. While in the presence of a more precise temporal resolution (e.g. in the auditory modality) the mechanisms of the SDN model could be limited to durations around 100 ms (Spencer et al., 2009) – possibly due to the faster dynamics of auditory networks – in the sluggish visual modality the model's working range could extend to longer durations, due to the intrinsic lower temporal resolution of the visual system.

In the context of numerosity perception (Chapter 4), we further investigated the possibility that number might represent a primary perceptual feature, as suggested by adaptation studies revealing the effects of numerosity adaptation (see Anobile et al., 2015b for a review). We exploited a visual illusion whereby a pattern of dots joined by task-irrelevant lines is robustly underestimated, in order to dissociate physical (the actual amount of dots, irrespective of the lines) from perceived numerosity (the apparent amount of elements in the connected-dots patch) (He et al., 2009; Franconeri et al., 2009; He et al., 2015). Results showed that while 20-dot adaptation does not affect a stimulus containing the same amount of unconnected dots, it does affect the connected patches, causing a further underestimation. Our findings thus suggest that numerosity adaptation operates on perceived and not physical numerosity, making unlikely that such effect might be driven by lower level processes, such texture-density adaptation. Overall, our data support the idea of number as a primary perceptual feature.

Finally (Chapter 5), we investigated the possibility of a generalized magnitude system processing spatial, temporal and numerical information by means of common mechanisms and with a common metrics, as proposed recently with the ATOM theory (Walsh, 2003; Burr et al., 2011) and related accounts (Dehaene & Brannon, 2010). First, we tested whether the effects of motion adaptation – previously demonstrated on the representation of space and apparent time – might be also observable on numerical estimates. Indeed, if the three features are actually processed by a common system, they should be similarly affected by processes like adaptation. What we found, testing the effect of translational and circular motion adaptation on numerosity discrimination, is a robust effect of underestimation, to a similar extent of the effect observed on apparent time (for translational motion). Moreover, the observed pattern of results seem to be independent from lower-level effects, such reduction in perceived contrast and visibility, and spatial frequency adaptation. However, even circular and slower translational motion have proven to be effective in distorting numerical estimates, while such conditions do not yield any effect on perceived time (Fornaciai et al., 2014; Fornaciai, Arrighi, & Burr, Under review). According to these results, we propose that while space, time and number might actually be processed by a common magnitude system, the specific mechanisms for their processing might be different – as showed by the different patterns of adaptation effects on temporal and numerical estimates.

We also tested the intertwining of different magnitudes in the form of the influence of information in one domain to judgments in another. In particular, we tested how the availability of spatial information interacts with judgments in the temporal modality – i.e. in an audio-visual temporal order judgment task. Indeed, if space and time are encoded with a similar metrics, information in one domain should be suitable to be used for improving judgments in the other domain. We found that the precision of temporal order estimates could be improved by the availability of spatial information. We conclude that the brain is capable of exploiting spatial cues to make predictions about stimulus timing. Such improvement can be seen as the a-priori probability of where and when an event is about to happen, which in turn can lead to an additional estimate of the stimulus timing that is integrated with the incoming sensory information. Moreover, even if in the previous study exploiting motion adaptation we found evidence suggesting two different mechanisms for time and number, the effect showed here suggest anyway that different magnitudes could be encoded with a similar metrics, in order to be exploited (if needed and useful) for improving behavioral outcomes regarding judgments on other dimensions.

6.1 Concluding remarks

Overall, the emerging picture from the current work, is that a common cortical system for the processing of spatial, temporal and numerical information is more than just a possibility. Indeed a growing amount of evidence seems to clearly support this possibility, although many features and details are still unknown. The crucial challenge for future works will be to carefully examine the features of such system, in order to shed light on what mechanisms are actually shared to subserve the representation of space, time and number, and what are conversely mechanisms unique for the specific domains - in other words, how much is extended such common system, and to what extent it subserve the perceptual and motor systems.

The studies presented in this work, while dealing often with specific aspects of different perceptual domains and different magnitudes, overall appear to converge in suggesting that the three perceptual judgements share common *metrics*. Indeed, while each domain surely possesses its own processing mechanisms, when metrical judgements come into play there might be a great overlap. The crucial purpose of such hypothetical generalized magnitude is to obtain, process and integrate information from different senses, as well as the output of different specific mechanisms, and the only way to do this is by means of a common code for different information.

The results presented here, while speaking in favour of the possibility of a common metric among different magnitudes, need however follow up studies to clarify several aspects. In particular future studies should investigate to what extent information in one domain could be exploited in different domains, and the extent of cross-dimensional interactions.

We think that the study of this magnitude system is crucial for a deeper understanding of human perception. Moreover, a more extensive understanding of such system could have important implications for the understanding of how the abilities concerning different magnitudes develop together from the childhood to the adult age, and how a dysfunction in one system might be predicted, or even alleviated, studying and intervening on the other two domains.

Such research program gets the perspective on human sensory systems on a broader framework: no longer concentrated only on the specific mechanisms, but tackling the problems from a wider point of view. According to that, while with our studies we only scratched the surface of a more complicated problem, we think that understanding the common sensory mechanisms for the processing of space, time and number represents a very exciting scientific challenge, that surely worth further and extensive research.

References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature neuroscience*, *7*(10), 1057–8. doi: 10.1038/nn1312

Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body. *London and Edinburgh Philosophical Magazine and Journal of Science*, *5*, 373–374. doi: 10.1080/14786443408648481

Alais, D., & Carlile, S. (2005). Synchronizing to real events: subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(6), 2244–7. doi: 10.1073/pnas.0407034102

Alais, D., & Cass, J. (2010). Multisensory perceptual learning of temporal order: audiovisual learning transfers to vision but not audition. *PloS one*, *5*(6), e11283. doi: 10.1371/journal.pone.0011283

Alder, T. B., & Rose, G. J. (2000). Integration and recovery processes contribute to the temporal selectivity of neurons in the midbrain of the northern leopard frog, Rana pipiens. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, *186*(10), 923–937. doi: 10.1007/s003590000144

Alexander, I., Cowey, A., & Walsh, V. (2005). The right parietal cortex and time perception: back to Critchley and the Zeitraffer phenomenon. *Cognitive Neuropsychology*, *22*(3-4), 306–315. doi: 10.1080/02643290442000356

Allan, L. G., Kristofferson, A. B., & Wiens, E. W. (1971). Duration discrimination of brief light flashes. *Perception & Psychophysics*, *9*(3), 327–334. doi: 10.3758/BF03212659

Allison, T., Wood, C. C., & Goff, W. R. (1983). Brain stem auditory, pattern-reversal visual, and short-latency somatosensory evoked potentials: Latencies in relation to age, sex, and brain and body size. *Electroencephalography and Clinical Neurophysiology*, *55*(6), 619–636. doi: 10.1016/0013-4694(83)90272-9

Andersen, T. S., & Mamassian, P. (2008). Audiovisual integration of stimulus transients. *Vision Research*, *48*(25), 2537–2544. doi: 10.1016/j.visres.2008.08.018

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology*, *19*(4), 452–458. doi: 10.1016/j.conb.2009.06.008

Anobile, G., Cicchini, G. M., & Burr, D. C. (2014). Separate Mechanisms for Perception of Numerosity and Density. *Psychological Science*, *25*(1), 265–270. doi: 10.1177/0956797613501520

Anobile, G., Cicchini, G. M., & Burr, D. C. (2015b). Number As a Primary Perceptual Attribute: A Review. *Perception*. doi: 10.1177/0301006615602599

Anobile, G., Turi, M., Cicchini, G. M., & Burr, D. C. (2015a). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision*, *15*(5), 4.

Anstis, S. M. (1974). Size adaptation to visual texture and print: evidence for spatial-frequency analysis. *The American journal of psychology*, *87*(1-2), 261–7.

Apthorp, D., & Bell, J. (2015). Symmetry is less than meets the eye. *Current Biology*, *25*(7), R267–R268. doi: 10.1016/j.cub.2015.02.017

Arbuthnott, G., Ingham, C. A., & Wickens, J. (2000). Dopamine and synaptic plasticity in the neostriatum. *Journal of Anatomy*, *196*(4), 587–596. doi: 10.1046/j.1469-7580.2000.19640587.x

Arnold, D. H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. *Vision Research*, *45*(10), 1275–1284. doi: 10.1016/j.visres.2004.11.014

Arrighi, R., Orsi, S., & Burr, D. C. (2010). Adaptation of speed and duration of biological motion stimuli. *Perception ECVP abstract*, *39*, 151.

Arrighi, R., Togoli, I., & Burr, D. C. (2014). A generalized sense of number. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1797), 20141791–20141791. doi: 10.1098/rspb.2014.1791

Ashbridge, E., Walsh, V., & Cowey, A. (1997). Temporal aspects of visual search studied by transcranial magnetic stimulation. *Neuropsychologia*, *35*(8), 1121–1131.

Atsma, J., Maij, F., Corneil, B. D., & Medendorp, W. P. (2014). No Perisaccadic Mislocalization with Abruptly Cancelled Saccades. *Journal of Neuroscience*, *34*(16), 5497–5504. doi: 10.1523/JNEUROSCI.4773-13.2014

Avillac, M., Ben Hamed, S., & Duhamel, J.-R. (2007). Multisensory Integration in the Ventral Intraparietal Area of the Macaque Monkey. *Journal of Neuroscience*, *27*(8), 1922–1932. doi: 10.1523/JNEUROSCI.2646-06.2007

Awater, H. (2004). Perception of Visual Space at the Time of Pro- and Anti-Saccades. *Journal of Neurophysiology*, *91*(6), 2457–2464. doi: 10.1152/jn.00821.2003

Ayhan, I., Bruno, A., Nishida, S., & Johnston, A. (2009). The spatial tuning of adaptation-based time compression. *Journal of Vision*, *9*(11), 2–2. doi: 10.1167/9.11.2

Ayhan, I., Bruno, A., Nishida, S., & Johnston, A. (2011). Effect of the luminance signal on adaptation-based time compression. *Journal of Vision*, *11*(7), 22–22. doi: 10.1167/11.7.22

Barnes, R., & Jones, M. R. (2000). Expectancy, Attention, and Time. *Cognitive Psychology*, *41*(3), 254–311. doi: 10.1006/cogp.2000.0738

Basso, M. A., & Wurtz, R. H. (1997). Modulation of neuronal activity by target uncertainty. *Nature*, *389*(6646), 66–69. doi: 10.1038/37975

Basso, M. a., & Wurtz, R. H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience*, *18*(18), 7519–34.

Bastian, A. J. (2006). Learning to predict the future: the cerebellum adapts feedforward movement control. *Current Opinion in Neurobiology*, *16*(6), 645–649. doi: 10.1016/j.conb.2006.08.016

Batchelor, A., Madge, D., & Garthwaite, J. (1994). Synaptic activation of metabotropic glutamate receptors in the parallel Fibre-Purkinje cell pathway in rat cerebellar slices. *Neuroscience*, *63*(4), 911–915. doi: 10.1016/0306-4522(94)90558-4

Bausenhart, K. M., Rolke, B., & Ulrich, R. (2008). Temporal preparation improves temporal resolution: evidence from constant foreperiods. *Perception & psychophysics*, *70*(8), 1504–1514. doi: 10.3758/PP.70.8.1504

Bechlivanidis, C., & Lagnado, D. a. (2013). Does the "Why" Tell Us the "When"? *Psychological Science*, *24*(8), 1563–1572. doi: 10.1177/0956797613476046

Berridge, M. J., Bootman, M. D., & Roderick, H. L. (2003). Calcium signalling: dynamics, homeostasis and remodelling. *Nature reviews. Molecular cell biology*, *4*(7), 517–29. doi: 10.1038/nrm1155

Bex, P. J., Metha, A. B., & Makous, W. (1999). Enhanced motion aftereffect for complex motions. *Vision research*, *39*(13), 2229–38.

Binda, P., Cicchini, G. M., Burr, D. C., & Morrone, M. C. (2009). Spatiotemporal Distortions of Visual Perception at the Time of Saccades. *Journal of Neuroscience*, *29*(42), 13147–13157. doi: 10.1523/JNEUROSCI.3723-09.2009

Binda, P., Morrone, M. C., & Bremmer, F. (2012). Saccadic Compression of Symbolic Numerical Magnitude. *PLoS ONE*, *7*(11), e49587. doi: 10.1371/journal.pone.0049587

Binda, P., Morrone, M. C., Ross, J., & Burr, D. C. (2011). Underestimation of perceived number at the time of saccades. *Vision Research*, *51*(1), 34–42. doi: 10.1016/j.visres.2010.09.028

Bizo, L. A., & White, K. G. (1994). The behavioral theory of timing: Reinforcer rate determines pacemaker rate. *Journal of the Experimental Analysis of Behavior*, *61*(1), 1334351. doi: 10.1901/jeab.1994.61-19

Bizo, L. a., & White, K. G. (1994). Pacemaker rate in the behavioral theory of timing. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*(3), 308–321. doi: 10.1037//0097-7403.20.3.308

Bizo, L. a., & White, K. G. (1997). Training with controlled reinforcer density: Implications for models of timing. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*(I), 44–55. doi: 10.1037/0097-7403.23.1.44

Bjoertomt, O., Cowey, A., & Walsh, V. (2002). Spatial neglect in near and far space investigated by repetitive transcranial magnetic stimulation. *Brain*, *125*(9), 2012–

2022.

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. *The Journal of Physiology*, *203*(1), 237–260.1.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport*, *12*(9), 1879–1884. doi: 10.1097/00001756-200107030-00023

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition*, *75*(1), 1–28.

Bottini, R., & Casasanto, D. (2013). Space and time in the child's mind: metaphoric or ATOMic? *Frontiers in Psychology*, *4*, 803. doi: 10.3389/fpsyg.2013.00803

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*(5241), 177–8.

Boynton, G. M. (2004). Adaptation and attentional selection. *Nature neuroscience*, *7*(1), 8–10. doi: 10.1038/nn0104-8

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial vision*, *10*(4), 433–436. doi: 10.1163/156856897X00357

Bregman, A. S., & Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of experimental psychology*, *89*(2), 244–249. doi: 10.1037/h0031163

Bridge, H., Leopold, D. A., & Bourne, J. A. (2015). Adaptive Pulvinar Circuitry Supports Visual Cognition. *Trends in Cognitive Sciences*, *S1364-6613*(15), 00253–3. doi: 10.1016/j.tics.2015.10.003

Bridgeman, B. (2007). Efference copy and its limitations. *Computers in biology and medicine*, *37*(7), 924–929. doi: 10.1016/j.compbiomed.2006.07.001

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722. doi: 10.1016/0042-6989(75)90290-4

Brody, C. D., Romo, R., & Kepecs, A. (2003). Basic mechanisms for graded persistent activity: Discrete attractors, continuous attractors, and dynamic representations. *Current Opinion in Neurobiology*, *13*(2), 204–211. doi: 10.1016/S0959-4388(03)00050-3

Broome, B. M., Jayaraman, V., & Laurent, G. (2006). Encoding and Decoding of Overlapping Odor Sequences. *Neuron*, *51*(4), 467–482. doi: 10.1016/j.neuron.2006.07.018

Brosch, M., & Schreiner, C. E. (2000). Sequence sensitivity of neurons in cat primary auditory cortex. *Cerebral Cortex*, *10*(12), 1155–1167.

Brown, S. W. (1995). Time, change, and motion: The effects of stimulus movement on temporal perception. *Perception & Psychophysics*, *57*(1), 105–116. doi: 10.3758/BF03211853

Bruce, C. J., & Goldberg, M. E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of neurophysiology*, *53*(3), 603–635.

Brune, F., & Lücking, C. (1969). Oculomotorik, Bewegungswahrnehmung und Raumkonstanz der Seedinge. *Der Nervenarzt*, *40*, 413–421.

Bruno, A., Ayhan, I., & Johnston, A. (2010). Retinotopic adaptation-based visual duration compression. *Journal of Vision*, *10*(10), 30. doi: 10.1167/10.10.30

Bueti, D., Bahrami, B., & Walsh, V. (2008). Sensory and Association Cortex in Time Perception. *Journal of Cognitive Neuroscience*, *20*(6), 1054–1062. doi: 10.1162/jocn.2008.20060

Bueti, D., & Macaluso, E. (2010). Auditory temporal expectations modulate activity in visual cortex. *NeuroImage*, *51*(3), 1168–1183. doi: 10.1016/j.neuroimage.2010.03.023

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature reviews. Neuroscience*, *6*(10), 755–765. doi: 10.1038/nrn1764

Buonomano, D. V. (2000). Decoding temporal information: A model based on short-term synaptic plasticity. *Journal of Neuroscience*, *20*(3), 1129–41.

Buonomano, D. V. (2007). The biology of time across different scales. *Nature chemical biology*, *3*(10), 594–7. doi: 10.1038/nchembio1007-594

Buonomano, D. V., Bramen, J., & Khodadadifar, M. (2009). Influence of the interstimulus interval on temporal processing and learning: testing the state-dependent network model. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1865–1873. doi: 10.1098/rstb.2009.0019

Buonomano, D. V., Hickmott, P. W., & Merzenich, M. M. (1997). Context-sensitive synaptic plasticity and temporal-to-spatial transformations in hippocampal slices. *Proceedings of the National Academy of Sciences of the United States of America*(19), 10403–8. doi: 10.1073/pnas.94.19.10403

Buonomano, D. V., & Karmarkar, U. R. (2002). How do we tell time? *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, *8*(1), 42–51. doi: 10.1177/107385840200800109

Buonomano, D. V., & Laje, R. (2010). Population clocks: motor timing with neural dynamics. *Trends in cognitive sciences*, *14*(12), 520–7. doi: 10.1016/j.tics.2010.09.002

Buonomano, D. V., & Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nature reviews. Neuroscience*, *10*(2), 113–25. doi: 10.1038/nrn2558

Buonomano, D. V., & Mauk, M. D. (1994). Neural Network Model of the Cerebellum: Temporal Discrimination and the Timing of Motor Responses. *Neural Computation*, *6*(1), 38–55. doi: 10.1162/neco.1994.6.1.38

Buonomano, D. V., & Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science (New York, N.Y.)*, *267*(5200), 1028–30.

Buonomano, D. V., & Merzenich, M. M. (1998). Net interaction between different forms of short-term synaptic plasticity and slow-IPSPs in the hippocampus and auditory cortex. *Journal of neurophysiology*, *80*(4), 1765–1774.

Burnashev, N., & Rozov, A. (2005). Presynaptic Ca2+ dynamics, Ca2+ buffers and synaptic efficacy. *Cell Calcium*, *37*(5 SPEC. ISS.), 489–495. doi: 10.1016/j.ceca.2005.01.003

Burr, D., Rocca, E. D., & Morrone, M. C. (2013). Contextual effects in interval-duration judgements in vision, audition and touch. *Experimental Brain Research*, *230*(1), 87–98. doi: 10.1007/s00221-013-3632-z

Burr, D., & Ross, J. (2008). A Visual Sense of Number. *Current Biology*, *18*(6), 425–428. doi: 10.1016/j.cub.2008.02.052

Burr, D., & Thompson, P. (2011). Motion psychophysics: 1985-2010. *Vision research*, *51*(13), 1431–56. doi: 10.1016/j.visres.2011.02.008

Burr, D., Tozzi, A., & Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nature neuroscience*, *10*(4), 423–5. doi: 10.1038/nn1874

Burr, D. C., Anobile, G., & Turi, M. (2011). Adaptation affects both high and low (subitized) numbers under conditions of high attentional load. *Seeing and perceiving*, *24*(2), 141–150. doi: 10.1163/187847511X570097

Burr, D. C., Cicchini, G. M., Arrighi, R., & Morrone, M. C. (2011). Spatiotopic selectivity of adaptation-based compression of event duration. *Journal of vision*, *11*(2), 21; author reply 21a. doi: 10.1167/11.2.21

Burr, D. C., Holt, J., Johnstone, J. R., & Ross, J. (1982). Selective depression of motion sensitivity during saccades. *J Physiol*, 1–15. doi: VL - 333

Burr, D. C., & Morrone, M. C. (2010). Vision: Keeping the World Still When the Eyes Move. *Current Biology*, *20*(10), R442–R444. doi: 10.1016/j.cub.2010.03.033

Burr, D. C., & Morrone, M. C. (2011). Spatiotopic coding and remapping in humans. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*(1564), 504–515. doi: 10.1098/rstb.2010.0244

Burr, D. C., & Morrone, M. C. (2012). Constructing stable spatial maps of the world. *Perception*, *41*(September), 1355–1372. doi: 10.1068/p7392

Burr, D. C., Morrone, M. C., & Ross, J. (1994). *Selective suppression of the magnocellular visual pathway during saccadic eye movements.* (Vol. 371) (No. 6497). doi: 10.1038/371511a0

Burr, D. C., Morrone, M. C., & Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans. *Vision research*, *38*(12), 1731–43.

Burr, D. C., & Ross, J. (1982). Contrast sensitivity at high velocities. *Vision Research*, *22*, 479–484. doi: 10.1016/0042-6989(82)90196-1

Burr, D. C., Ross, J., Binda, P., & Morrone, M. C. (2010). Saccades Compress Space, Time, and Number. *Trends in cognitive sciences*, *14*(12), 528–533. doi: 10.1016/B978-0-12-

385948-8.00012-8

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of vision*, *10*, 20. doi: 10.1167/10.6.20

Butterworth, B. (1999). *The Mathematical Brain.* London: Macmillan.

Cai, Z. G., & Connell, L. (2015). Space–time interdependence: Evidence against asymmetric mapping between time and space. *Cognition*, *136*, 268–281.

Calvert, G. A., & Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris*, *98*(1-3), 191–205. doi: 10.1016/j.jphysparis.2004.03.018

Carlson, B. a. (2009). Temporal-Pattern Recognition by Single Neurons in a Sensory Pathway Devoted to Social Communication Behavior. *Journal of Neuroscience*, *29*(30), 9417–9428. doi: 10.1523/JNEUROSCI.1980-09.2009

Carpenter, R. (1977). *Movements of the eyes.* London: Pion.

Carpenter, R. (1999). A neural mechanism that randomises behaviour. *Journal of Consciousness Studies*, *6*(1), 13–22.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593.

Casasanto, D., Fotakopoulou, O., & Boroditsky, L. (2010). Space and Time in the Child's Mind: Evidence for a CrossDimensional Asymmetry. *Cognitive Science*, *34*(3), 387–405.

Cass, J., & Van der Burg, E. (2014). Remote temporal camouflage: Contextual flicker disrupts perceived visual temporal order. *Vision Research*, *103*, 92–100. doi: 10.1016/j.visres.2014.08.008

Cavanagh, P. (1997). Visual perception. Predicting the present. *Nature*, *386*(6620), 19–21. doi: 10.1038/386019a0

Charpier, S., & Deniau, J. M. (1997). In vivo activity-dependent plasticity at corticostriatal connections: evidence for physiological long-term potentiation. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(13), 7036–7040. doi: 10.1073/pnas.94.13.7036

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, *75*(5), 790–811. doi: 10.3758/s13414-013-0475-4

Chochon, F., Cohen, L., van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of cognitive neuroscience*, *11*(6), 617–630. doi: 10.1162/089892999563689

Chopin, A., & Mamassian, P. (2012). Predictive properties of visual adaptation. *Current Biology*, *22*(7), 622–626. doi: 10.1016/j.cub.2012.02.021

Church, R. M. (1978). The internal clock. In S. Hulse, H. Fowler, & W. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 227–310). Hillsdale, NJ: Erlbaum.

Church, R. M., & Broadbent, H. a. (1990). Alternative representations of time, number, and rate. *Cognition*, *37*(1-2), 55–81. doi: 10.1016/0010-0277(90)90018-F

Church, R. M., & Broadbent, H. A. (1991). A connectionist model of timing. In *Neural network models of conditioning and action.* (pp. 225–240).

Church, R. M., & Meck, W. H. (1984). The numerical attribute of stimuli. *Animal cognition*, 445–464.

Churchland, M. M., Yu, B. M., Sahani, M., & Shenoy, K. V. (2007). Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Current opinion in neurobiology*, *17*(5), 609–18. doi: 10.1016/j.conb.2007.11.001

Cicchini, G. M. (2012). Perception of duration in the parvocellular system. *Frontiers in Integrative Neuroscience*, *6*(March), 1–9. doi: 10.3389/fnint.2012.00014

Cicchini, G. M., Binda, P., Burr, D. C., & Morrone, M. C. (2013). Transient spatiotopic integration across saccadic eye movements mediates visual stability. *Journal of neurophysiology*, *109*(4), 1117–25. doi: 10.1152/jn.00478.2012

Cicchini, G. M., & Morrone, M. C. (2009). Shifts in spatial attention affect the perceived duration of events. *Journal of vision*, *9*(1), 9.1–13. doi: 10.1167/9.1.9

Clark, H. H. (1973). Space, time, semantics and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). New York: Academic Press.

Cohen, L., & Dehaene, S. (1996). Cerebral networks for number processing: Evidence from a case of posterior callosal lesion. *Neurocase*, *2*(3), 155–174.

Colonius, H. (2015). Behavioral Measures of Multisensory Integration: Bounds on Bimodal Detection Probability. *Brain Topography*, *28*(1), 1–4. doi: 10.1007/s10548-014-0421-3

Cook, L. a., Van Valkenburg, D. L., & Badcock, D. R. (2011). Predictability affects the perception of audiovisual synchrony in complex sequences. *Attention, Perception, & Psychophysics*, *73*(7), 2286–2297. doi: 10.3758/s13414-011-0185-8

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc*, *34*(4), 613–617. doi: 10.3758/BF03195489

Costello, P., Jiang, Y., Baartman, B., McGlennen, K., & He, S. (2009). Semantic and subword priming during binocular suppression. *Consciousness and Cognition*, *18*(2), 375–382. doi: 10.1016/j.concog.2009.02.003

Coubart, A., Izard, V., Spelke, E. S., Marie, J., & Streri, A. (2014). Dissociation between small and large numerosities in newborn infants. *Developmental Science*, *17*(1), 11–22. doi: 10.1111/desc.12108

Coull, J. T., Vidal, F., Nazarian, B., & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science*, *303*(5663), 1506–1508.

Cover, T. M. (1965). Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Comput-*

ers, *EC-14*(3), 326–334. doi: 10.1109/PGEC.1965.264137

Creelman, C. D. (1962). Human Discrimination of Auditory Duration. *The Journal of the Acoustical Society of America*, *34*(5), 582. doi: 10.1121/1.1918172

Crespi, S., Biagi, L., D'Avossa, G., Burr, D. C., Tosetti, M., & Morrone, M. C. (2011). Spatiotopic coding of BOLD signal in human visual cortex depends on spatial attention. *PloS one*, *6*(7), e21661. doi: 10.1371/journal.pone.0021661

Critchley, M. (1953). *The parietal lobes.* London: Edward Arnold.

Cunningham, D. W., Billock, V. A., & Tsou, B. H. (2001). Sensorimotor adaptation to violations of temporal contiguity. *Psychological science : a journal of the American Psychological Society / APS*, *12*(6), 532–535. doi: 10.1111/1467-9280.d01-17

Curran, W., & Benton, C. P. (2012). The many directions of time. *Cognition*, *122*(2), 252–257. doi: 10.1016/j.cognition.2011.10.016

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108*(49), 19552–19557. doi: 10.1073/pnas.1113195108

D'Avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D. C., & Morrone, M. C. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nature neuroscience*, *10*(2), 249–55. doi: 10.1038/nn1824

de Boer-Schellekens, L., Eussen, M., & Vroomen, J. (2013). Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. *Frontiers in integrative neuroscience*, *7*(February), 8. doi: 10.3389/fnint.2013.00008

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends in cognitive sciences*, *7*(4), 145–147.

Dehaene, S. (2009). Origins of Mathematical Intuitions. *Annals of the New York Academy of Sciences*, *1156*(1), 232–259. doi: 10.1111/j.1749-6632.2009.04469.x

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371.

Dehaene, S., & Brannon, E. M. (2010). Space, time, and number: a Kantian research program. *Trends in cognitive sciences*, *14*(12), 517–9. doi: 10.1016/j.tics.2010.09.009

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science (New York, N.Y.)*, *320*(5880), 1217–20. doi: 10.1126/science.1156540

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science*, *284*(5416), 970–974.

DeLong, A. J. (1980). Phenomenological space-time: toward an experiential relativity. *Science*, *213*(4508), 681–683. doi: 10.1126/science.7256273

Deneve, S., & Pouget, A. (2004). Bayesian multisensory integration and cross-modal spatial links. *Journal of Physiology-Paris*, *98*(1-3), 249–258. doi: 10.1016/j.jphysparis.2004.03.011

Deubel, H. (2004). Localization of targets across saccades: Role of landmark objects. *Visual Cognition*, *11*(2-3), 173–202. doi: 10.1080/13506280344000284

De-Wit, L., Machilsen, B., & Putzeys, T. (2010). Predictive Coding and the Neural Response to Predictable Stimuli. *Journal of Neuroscience*, *30*(26), 8702–8703. doi: 10.1523/JNEUROSCI.2248-10.2010

Di Luca, M., Knörlein, B., Ernst, M. O., & Harders, M. (2011). Effects of visual-haptic asynchronies and loading-unloading movements on compliance perception. *Brain Research Bulletin*, *85*(5), 245–259. doi: 10.1016/j.brainresbull.2010.02.009

Di Luca, M., Machulla, T.-K., & Ernst, M. O. (2009). Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. *Journal of vision*, *9*(12), 7.1–16. doi: 10.1167/9.12.7

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(July), 1–17. doi: 10.3389/fpsyg.2014.00781

Ditchburn, R. (1955). Eye-movements in Relation to Retinal Action. *Optica Acta: International Journal of Optics*, *1*(February 2015), 171–176. doi: 10.1080/713818684

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, *9*(6), 719–721. doi: 10.1068/p090719

Dobrunz, L. E., & Stevens, C. F. (1999). Response of hippocampal synapses to natural stimulation patterns. *Neuron*, *22*(1), 157–166. doi: 10.1016/S0896-6273(00)80687-X

Dodge, R. (1900). Visual perception during eye movement. *Psychological Review*, *7*(5), 454–465. doi: 10.1037/h0067215

Dodge, R. (1905). The illusion of clear vision during eye movement. *Psychological Bulletin*, *2*(6), 193.

Donohue, S. E., Green, J. J., & Woldorff, M. G. (2015). The effects of attention on the temporal integration of multisensory stimuli. *Frontiers in integrative neuroscience*, *9*(32), 1–14. doi: 10.3389/fnint.2015.00032

Doupe, A. J., & Solis, M. M. (1997). Song- and order-selective neurons develop in the songbird anterior forebrain during vocal learning. *Journal of Neurobiology*, *33*(5), 694–709. doi: 10.1002/(SICI)1097-4695(19971105)33:5¡694::AID-NEU13¿3.0.CO;2-9

Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & psychophysics*, *54*(3), 277–286. doi: 10.3758/BF03205262

Duffy, C. J., & Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *Journal of neurophysiology*, *65*(6), 1329–45.

Duffy, C. J., & Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *Journal of neurophysiology*, *65*(6), 1346–59.

Duhamel, J.-r., Colby, C. L., & Goldberg, M. E. (1992). The Updating of the Representation of Visual. *Science*, *255*(1990), 90–92.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*(5040), 90–2.

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 149–169. doi: 10.1037/0096-1523.21.1.149

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*, *18*(18), R855–R856. doi: 10.1016/j.cub.2008.07.053

Eagleman, D. M. (2008). Human time perception and its illusions. *Current opinion in neurobiology*, *18*(2), 131–6. doi: 10.1016/j.conb.2008.06.002

Eg, R., & Behne, D. M. (2015). Perceived synchrony for realistic and dynamic audiovisual events. *Frontiers in Psychology*, *6*(June). doi: 10.3389/fpsyg.2015.00736

Eger, E., Henson, R. N., Driver, J., & Dolan, R. J. (2007). Mechanisms of top-down facilitation in perception of visual objects studied by fMRI. *Cerebral Cortex*, *17*(9), 2123–2133. doi: 10.1093/cercor/bhl119

Ehinger, B. V., Konig, P., & Ossandon, J. P. (2015). Predictions of Visual Content across Eye Movements and Their Modulation by Inferred Information. *Journal of Neuroscience*, *35*(19), 7403–7413. doi: 10.1523/JNEUROSCI.5114-14.2015

Eisler, H., Eisler, A. D., & Hellstrom, A. (2008). Psychophysical issues in the study of time perception. In S. Grondin (Ed.), *Psychology of time* (pp. 75–110). Bingley, U.K.: Emerald Group.

Ellingstad, V. S., & Heimstra, N. W. (1969). Velocity-Time Estimation as a Function of Target Speed and Concealment Extent. *Human Factors*, *11*(4), 305–312. doi: 10.1177/001872086901100401

Ellis, H. D. (1973). Proactive effects of interpolated anchors. *Journal of experimental psychology*, *98*(2), 233–238.

Enns, J., & Di Lollo V. (2000). What's new in visual masking? *Trends in cognitive sciences*, *4*(9), 345–352. doi: 10.1016/S1364-6613(00)01520-5

Erdmann, B., & Dodge, R. (1898). *Psychologische Untersuchung über das Lesen auf experimenteller Grundlage.* Halle: Niemeyer.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169. doi: 10.1016/j.tics.2004.02.002

Ernst, M. O., & Luca, M. D. (2011). Multisensory Perception: From Integration to Remapping. In *Sensory cue integration* (pp. 224–250). Oxford University Press. doi: 10.1093/acprof:oso/9780195387247.003.0012

Everling, S., & Munoz, D. P. (2000). Neuronal correlates for preparatory set associated with pro-saccades and anti-saccades in the primate frontal eye field. *Journal of Neuroscience*, *20*(1), 387–400.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. doi: 10.1016/j.tics.2004.05.002

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links Between the Intuitive Sense of Number and Formal Mathematics Ability. *Child Development Perspectives*, *7*(2), 74–79. doi: 10.1111/cdep.12019

Feldman, J. (2015). Bayesian models of perception: a tutorial introduction. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization* (pp. 933–947). Oxford: Oxford University Press.

Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annual review of neuroscience*, *23*, 441–71. doi: 10.1146/annurev.neuro.23.1.441

Fias, W. (1996). The importance of magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical cognition*, *2*(1), 95–110.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons. *Science*, *299*(March), 1898–1902. doi: 10.1126/science.1077349

Fischer, M. H. (2003). Spatial representations in number processing–evidence from a pointing task. *Visual Cognition*, *4*(April 2015), 493–508. doi: 10.1080/13506280244000186

Fornaciai, M., Arrighi, R., & Burr, D. C. (n.d.). Adaptation-induced compression of event time occurs only for translational motion.

Fornaciai, M., Arrighi, R., & Burr, D. C. (2014). Fast Translational Motion, but not Radial, Circular or Biological Motion, Causes Spatially Selective Adaptation of Event Duration. *Procedia - Social and Behavioral Sciences*, *126*, 123–124. doi: 10.1016/j.sbspro.2014.02.339

Fornaciai, M., & Binda, P. (2015). Effect of saccade automaticity on perisaccadic space compression. *Frontiers in Systems Neuroscience*, *9*(127). doi: 10.3389/fnsys.2015.00127

Fornaciai, M., & Di Luca, M. (n.d.). Audio-visual simultaneity in a time to come: predictability improves temporal order judgments.

Fornaciai, M., Markouli, E., & Di Luca, M. (n.d.). The effect of Inter-Stimulus Interval on auditory and visual duration discrimination.

Fornaciai, M., Togoli, I., Burr, D. C., & Arrighi, R. (2015). Motion-induced compression of perceived numerosity. *Perception ECVP 2015 Abstract Supplement*, *44*(S1), 277. doi: 10.1177/0301006615598674

Franconeri, S., Bemis, D., & Alvarez, G. (2009). Number estimation relies on a set of segmented objects. *Cognition*, *113*(1), 1–13. doi: 10.1016/j.cognition.2009.07.002

Freedman, D. J., & Assad, J. A. (2006). Experience-dependent representation of visual categories in parietal cortex. *Nature*, *443*(7107), 85–8. doi: 10.1038/nature05078

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126–132. doi: 10.1037/0278-

7393.10.1.126

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *The Journal of the Acoustical Society of America*, *98*(5), 2524–2531. doi: 10.1121/1.413218

Friston, K. (2005). A theory of cortical responses. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *360*(1456), 815–836. doi: 10.1098/rstb.2005.1622

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *364*(1521), 1211–1221. doi: 10.1098/rstb.2008.0300

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature neuroscience*, *7*(7), 773–778. doi: 10.1167/3.9.34

Gallistel, C. R., & Gelman, R. (1990). The what and how of counting. *Cognition*, *34*(2), 197–199.

Galton, F. (1881). Visualised numerals. *Journal of the Anthropological Institute of Great Britain and Ireland*, 85–102.

Garcia, K. S., & Mauk, M. D. (1998). Pharmacological analysis of cerebellar contributions to the timing and expression of conditioned eyelid responses. *Neuropharmacology*, *37*(4-5), 471–480. doi: 10.1016/S0028-3908(98)00055-0

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of Visual Space in Human Occipital Cortex Are Retinotopic, Not Spatiotopic. *Journal of Neuroscience*, *28*(15), 3988–3999. doi: 10.1523/JNEUROSCI.5476-07.2008

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nature reviews. Neuroscience*, *4*(7), 563–72. doi: 10.1038/nrn1138

Gekas, N., & Seitz, A. R. (2015). Expectations developed over multiple timescales facilitate visual search performance. *Journal of Vision*, *15*(9), 1–22. doi: 10.1167/15.9.10.doi

Genovesio, A., Brasted, P. J., Mitz, A. R., & Wise, S. P. (2005). Prefrontal cortex activity related to abstract response strategies. *Neuron*, *47*(2), 307–20. doi: 10.1016/j.neuron.2005.06.006

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2006). Neuronal activity related to elapsed time in prefrontal cortex. *Journal of neurophysiology*, *95*(January 2006), 3281–3285. doi: 10.1152/jn.01011.2005

Georgeson, M. A. (1985). The effect of spatial adaptation on perceived contrast. *Spatial vision*, *1*(2), 103–12.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279–325. doi: 10.1037/0033-295X.84.3.279

Gibbon, J. (1991). Origins of scalar timing. *Learning and Motivation*, *22*(1-2), 3–38. doi: 10.1016/0023-9690(91)90015-Z

Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical*

Psychology, *36*(2), 283–293. doi: 10.1016/0022-2496(92)90041-5

Gibbs, R. W. (2006). Metaphor interpretation as embodied simulation. *Mind and Language*, *21*(3), 434–458. doi: 10.1111/j.1468-0017.2006.00285.x

Göbel, S. M., Calabria, M., Farnè, A., & Rossetti, Y. (2006). Parietal rTMS distorts the mental number line: simulating 'spatial' neglect in healthy subjects. *Neuropsychologia*, *44*(6), 860–868.

Goel, A., & Buonomano, D. V. (2014). Timing as an intrinsic property of neural networks: evidence from in vivo and in vitro experiments. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *369*(1637), 20120460. doi: 10.1098/rstb.2012.0460

Goodale, M. (2004). An evolving view of duplex vision: separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology*, *14*(2), 203–211. doi: 10.1016/j.conb.2004.03.002

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

Goolkasian, P., & Woodberry, C. (2010). Priming effects with ambiguous figures. *Attention, perception & psychophysics*, *72*(1), 168–178. doi: 10.3758/APP.72.1.168

Gottlieb, J., & Goldberg, M. E. (1999). Activity of neurons in the lateral intraparietal area of the monkey during an antisaccade task. *Nature neuroscience*, *2*(10), 906–12. doi: 10.1038/13209

Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, *391*(6666), 481–484. doi: 10.1038/35135

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics*, *54*(3), 383–394. doi: 10.3758/BF03205274

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attention, perception & psychophysics*, *72*(3), 561–82. doi: 10.3758/APP.72.3.561

Grossberg, S., & Schmajuk, N. (1988). Neural dynamics of adaptive timing and temporal discrimination during associative learning. *Neural Networks*, *1*(1 SUPPL), 98. doi: 10.1016/0893-6080(88)90137-2

Guo, X., Li, X., Ge, X., & Tong, S. (2015). Audiovisual congruency and incongruency effects on auditory intensity discrimination. *Neuroscience Letters*, *584*, 241–246. doi: 10.1016/j.neulet.2014.10.043

Gygi, B., Giordano, B. L., Shafiro, V., Kharkhurin, A., & Zhang, P. X. (2015). Predicting the timing of dynamic events through sound: Bouncing balls. *The Journal of the Acoustical Society of America*, *138*(1), 457–466. doi: 10.1121/1.4923020

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature neuroscience*, *5*(4), 382–385. doi: 10.1038/nn827

Halfwerk, W., Page, R. a., Taylor, R. C., Wilson, P. S., & Ryan, M. J. (2014). Crossmodal comparisons of signal components allow for relative-distance assessment. *Current Biology*, *24*(15), 1751–1755. doi: 10.1016/j.cub.2014.05.068

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research*, *186*(4), 517–524. doi: 10.1007/s00221-007-1253-0

Harrington, D. L., & Haaland, K. Y. (1999). Neural underpinnings of temporal processing: a review of focal lesion, pharmacological, and functional imaging research. *Reviews in the neurosciences*, *10*(2), 91–116. doi: 10.1515/REVNEURO.1999.10.2.91

Hartcher-O'Brien, J., Di Luca, M., & Ernst, M. O. (2014). The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PloS one*, *9*(3), e89339. doi: 10.1371/journal.pone.0089339

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science*, *341*(September), 1123–1126. doi: 10.1126/science.1239052

Hauser, M. D., Carey, S., & Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proceedings. Biological sciences / The Royal Society*, *267*(1445), 829–833. doi: 10.1098/rspb.2000.1078

He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment: implications for configural processing. *Psychonomic bulletin & review*, *16*(3), 509–517. doi: 10.3758/PBR.16.3.509

He, L., Zhou, K., Zhou, T., He, S., & Chen, L. (2015). Topology-defined units in numerosity perception. *Proceedings of the National Academy of Sciences*, *112*(41), E5647–E5655. doi: 10.1073/pnas.1512408112

Hein, E., Rolke, B., & Ulrich, R. (2006). Visual attention and temporal discrimination: Differential effects of automatic and voluntary cueing. *Visual Cognition*, *13*(1), 29–50. doi: 10.1080/13506280500143524

Hellström, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, *97*(1), 35–61. doi: 10.1037/0033-2909.97.1.35

Hellström, A. (2003). Comparison is not just subtraction: effects of time- and space-order on subjective stimulus difference. *Perception & psychophysics*, *65*(7), 1161–1177. doi: 10.3758/BF03194842

Hellström, Å., & Cederström, C. (2014). Intramodal and crossmodal pairing and anchoring in comparisons of successive stimuli. *Attention, Perception, & Psychophysics*, *76*(4), 1197–1211. doi: 10.3758/s13414-014-0635-1

Hellström, A., & Rammsayer, T. H. (2004). Effects of time-order, interstimulus interval, and feedback in duration discrimination of noise bursts in the 50- and 1000-ms ranges. *Acta psychologica*, *116*(August 2015), 1–20. doi: 10.1016/j.actpsy.2003.11.003

Henson, R. N. a., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*(3), 263–270. doi: 10.1016/S0028-3932(02)00159-8

Heron, J., Hanson, J. V. M., & Whitaker, D. (2009). Effect before Cause: Supramodal Recalibration of Sensorimotor Timing. *PLoS ONE*, *4*(11), e7681. doi: 10.1371/journal.pone.0007681

Heron, J., Whitaker, D., & McGraw, P. V. (2004). Sensory uncertainty governs the extent of audio-visual interaction. *Vision Research*, *44*(25), 2875–2884. doi: 10.1016/j.visres.2004.07.001

Hinton, S. C., & Meck, W. H. (2004). Frontal–striatal circuitry activated by human peak-interval timing in the supra-seconds range. *Cognitive Brain Research*, *21*(2), 171–182. doi: 10.1016/j.cogbrainres.2004.08.005

Hirsh, I. J., & Sherrick Jr., C. E. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology*, *62*(5), 423–432. doi: 10.1037/h0045283

Hoerl, C., & McCormack, T. (2011). Time in cognitive development. In C. Callender (Ed.), *Oxford handbook of the philosophy of time* (pp. 439–459). Oxford University Press. doi: 10.1093/oxfordhb/9780199298204.003.0015

Hogendoorn, H., Verstraten, F. a. J., & Johnston, A. (2010). Spatially localized time shifts of the perceptual stream. *Frontiers in Psychology*, *1*(NOV), 1–8. doi: 10.3389/fpsyg.2010.00181

Holson, R. R., Bowyer, J. F., Clausing, P., & Gough, B. (1996). Methamphetamine-stimulated striatal dopamine release declines rapidly over time following microdialysis probe insertion. *Brain Research*, *739*(1-2), 301–307. doi: S0006-8993(96)00837-2 [pii]

Holt, E. B. (1903). Eye-movement and central anaesthesia. *The Psychological Review: Monograph Supplements*, *4*(1), 1–45.

Hommel, B., & Zmigrod, S. (2013). Feature Integration across Multimodal Perception and Action: A Review. *Multisensory Research*, *26*(1-2), 143–157. doi: 10.1163/22134808-00002390

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, *148*, 574–591. doi: 10.1113/jphysiol.2009.174151

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154.2.

Ivry, R. B., Keele, S. W., & Diener, H. C. (1988). Dissociation of the lateral and medial cerebellum in movement timing and movement execution. *Experimental Brain Research*, *73*, 167–180. doi: 10.1007/BF00279670

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in cognitive sciences*, *12*(7), 273–80. doi: 10.1016/j.tics.2008.04.002

Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current opinion in neurobiology*, *14*(2), 225–232. doi: 10.1016/j.conb.2004.03.013

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition*, *106*(3), 1221–1247.

Jaekl, P., Seidlitz, J., Harris, L. R., & Tadin, D. (2015). Audiovisual Delay as a Novel Cue to Visual Distance. *Plos One*, *10*(10), e0141125. doi: 10.1371/journal.pone.0141125

Jaekl, P., Soto-Faraco, S., & Harris, L. R. (2012). Perceived size change induced by audiovisual temporal delays. *Experimental brain research*, *216*(3), 457–62. doi: 10.1007/s00221-011-2948-9

Jamieson, D. G., & Petrusic, W. M. (1975). The dependence of time-order error direction on stimulus range. *Canadian Journal of Psychology*, *29*(3), 175–182. doi: 10.1037/h0082023

Jevons, W. S. (1871). *Theory of political economy* (1911st ed.). London: Macmillan.

Johnston, A. (2010). Modulation of time perception by visual adaptation. In A. C. Nobre & J. T. Coull (Eds.), *Attention and time* (pp. 187–200). Oxford University Press.

Johnston, A., Arnold, D. H., & Nishida, S. (2006). Spatially localized distortions of event time. *Current biology : CB*, *16*(5), 472–9. doi: 10.1016/j.cub.2006.01.032

Johnston, A., Bruno, A., Watanabe, J., Quansah, B., Patel, N., Dakin, S., & Nishida, S. (2008). Visually-based temporal distortion in dyslexia. *Vision research*, *48*(17), 1852–8. doi: 10.1016/j.visres.2008.04.029

Johnston, D., & Wu, S. (1995). *Foundations of Cellular Neurophysiology.* Cambridge, MA: MIT Press.

Johnston, K., & Everling, S. (2008). Neurophysiology and neuroanatomy of reflexive and voluntary saccades in non-human primates. *Brain and Cognition*, *68*(3), 271–283. doi: 10.1016/j.bandc.2008.08.017

Johnstone, J. R., & Riggs, L. A. (1979). Upper-velocity threshold for detection of movement. *Optics letters*, *4*(10), 309–310.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological review*, *96*(3), 459–491. doi: 10.1037/0033-295X.96.3.459

Kaliuzhna, M., Prsa, M., Gale, S., Lee, S. J., & Blanke, O. (2015). Learning to integrate contradictory multisensory self-motion cue pairings. *Journal of vision*, *15*(1), 15.1.10. doi: 10.1167/15.1.10

Kanai, R., Paffen, C. L. E., Hogendoorn, H., & Verstraten, F. A. J. (2006). Time dilation in dynamic visual display. *Journal of vision*, *6*(12), 1421–30. doi: 10.1167/6.12.8

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron*, *53*(3), 427–38. doi: 10.1016/j.neuron.2007.01.006

Kaufman, E. L., & Lord, M. W. (1949). The discrimination of visual number. *The American journal of psychology*, *62*(4), 498–525.

Kaufman, L., & Kaufman, J. H. (2000). Explaining the moon illusion. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(1), 500–5. doi: 10.1073/pnas.97.1.500

Kayser, C., & Shams, L. (2015). Multisensory Causal Inference in the Brain. *PLOS Biology*, *13*(2), e1002075. doi: 10.1371/journal.pbio.1002075

Kerlin, J. R., & Shapiro, K. L. (2015). Multisensory Integration: How Sound Alters Sight. *Current Biology*, *25*(2), R76–R77. doi: 10.1016/j.cub.2014.12.016

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological review*, *95*(2), 274–295. doi: 10.1037/0033-295X.95.2.274

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147. doi: 10.1016/S1364-6613(00)01452-2

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception ECVP 2007 Abstract Supplement*, 14. doi: 10.1068/v070821

Knapen, T., Rolfs, M., & Cavanagh, P. (2009). The reference frame of the motion aftereffect is retinotopic. *Journal of vision*, *9*(5), 16.1–7. doi: 10.1167/9.5.16

Knill, D., & Richards, W. (Eds.). (1996). *Perception as Bayesian inference.* Cambrige: Cambridge University Press.

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*, *134*(3), 372–384. doi: 10.1016/j.actpsy.2010.03.010

Kohlrausch, A., van Eijk, R., Juola, J. F., Brandt, I., & van de Par, S. (2013). Apparent causality affects perceived simultaneity. *Attention, perception & psychophysics*, *75*(7), 1366–73. doi: 10.3758/s13414-013-0531-0

Kohn, A. (2007). Visual adaptation: physiology, mechanisms, and functional benefits. *Journal of neurophysiology*, *97*(5), 3155–3164. doi: 10.1152/jn.00086.2007

Kopinska, A., & Harris, L. R. (2004). Simultaneity constancy. *Perception*, *33*(9), 1049–1060. doi: 10.1068/p5169

Kostarakos, K., & Hedwig, B. (2012). Calling Song Recognition in Female Crickets: Temporal Tuning of Identified Brain Neurons Matches Behavior. *Journal of Neuroscience*, *32*(28), 9601–9612. doi: 10.1523/JNEUROSCI.1170-12.2012

Krauzlis, R. J., Lovejoy, L. P., & Zénon, A. (2013). Superior Colliculus and Visual Spatial Attention. *Annual Review of Neuroscience*, *36*(1), 165–182. doi: 10.1146/annurev-neuro-062012-170249

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research*, *39*(16), 2669–2679. doi: 10.1016/S0042-6989(98)00287-9

Kruschke, J. K., & Fragassi, M. M. (1996). The perception of causality: Feature binding in interacting objects. *Proceedings of the 18th Annual Conference of the Cognitive Science Society*, 441–446.

Kwon, J., Ogawa, K.-i., & Kim, J. (2014). Temporal Processing for Visual Motion Information in Audiovisual Integration. *ICME International Conference on Complex Medical*

Engineering, 67–72.

Kwon, J., Ogawa, K.-i., & Miyake, Y. (2014). The effect of visual apparent motion on audiovisual simultaneity. *PloS one*, *9*(10), e110224. doi: 10.1371/journal.pone.0110224

Lappe, M., Awater, H., & Krekelberg, B. (2000). Postsaccadic visual references generate presaccadic compression of space. *Nature*, *403*(6772), 892–5. doi: 10.1038/35002588

Latimer, K., Curran, W., & Benton, C. P. (2014). Direction-contingent duration compression is primarily retinotopic. *Vision Research*, *105*, 47–52. doi: 10.1016/j.visres.2014.09.004

Lebedev, M. a., O'Doherty, J. E., & Nicolelis, M. a. L. (2008). Decoding of temporal intervals from cortical ensemble activity. *Journal of neurophysiology*, *99*(1), 166–186. doi: 10.1152/jn.00734.2007

Legenstein, R., & Maass, W. (2007). Edge of chaos and prediction of computational performance for neural circuit models. *Neural Networks*, *20*(3), 323–34. doi: 10.1016/j.neunet.2007.04.017

Leon, M. I., & Shadlen, M. N. (2003). Representation of Time by Neurons in the Posterior Parietal Cortex of the Macaque. *Neuron*, *38*(2), 317–327. doi: 10.1016/S0896-6273(03)00185-5

Lester, R. a., Clements, J. D., Westbrook, G. L., & Jahr, C. E. (1990). Channel kinetics determine the time course of NMDA receptor-mediated synaptic currents. *Nature*, *346*(6284), 565–567. doi: 10.1038/346565a0

Leung, S. (2002). *Causation & causality: two mirrors of the cause.* London: Empiricus Books.

Levi, D. M. (2008). Crowding-An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*(5), 635–654. doi: 10.1016/j.visres.2007.12.009

Levitin, D. J., MacLean, K., Mathews, M., Chu, L., & Jensen, E. (2000). The perception of cross-modal simultaneity. *International Journal of Computing Anticipatory Systems*, 6–9.

Lewis, P. A., & Miall, R. C. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia*, *41*(12), 1583–92.

Lewis, P. a., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Current opinion in neurobiology*, *13*(2), 250–5. doi: 10.1016/S0959-4388(03)00036-9

Lewis, P. A., & Miall, R. C. (2006a). Remembering the time: a continuous clock. *Trends in cognitive sciences*, *10*(9), 401–406.

Lewis, P. A., & Miall, R. C. (2006b). A right hemispheric prefrontal system for cognitive time measurement. *Behavioural processes*, *71*(2-3), 226–34. doi: 10.1016/j.beproc.2005.12.009

Libertus, M. E., Feigenson, L., & Halberda, J. (2013a). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences*, *25*, 126–133.

doi: 10.1016/j.lindif.2013.02.001

Libertus, M. E., Feigenson, L., & Halberda, J. (2013b). Numerical approximation abilities correlate with and predict informal but not formal mathematics abilities. *Journal of Experimental Child Psychology*, *116*(4), 829–838. doi: 10.1016/j.jecp.2013.08.003

Liverence, B. M., & Scholl, B. J. (2012). Discrete events as units of perceived time. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 549–554. doi: 10.1037/a0027228

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*(11), 3416–3468.

Love, S. A., Petrini, K., Cheng, A., & Pollick, F. E. (2013). A Psychophysical Investigation of Differences between Synchrony and Temporal Order Judgments. *PLoS ONE*, *8*(1), e54798. doi: 10.1371/journal.pone.0054798

Lupiáñez, J., Milán, E. G., Tornay, F. J., Madrid, E., & Tudela, P. (1997). Does IOR occur in discrimination tasks? Yes, it does, but later. *Perception & Psychophysics*, *59*(8), 1241–1254. doi: 10.3758/BF03214211

Lyon, D. C., Nassi, J. J., & Callaway, E. M. (2010). A Disynaptic Relay from Superior Colliculus to Dorsal Stream Visual Cortex in Macaque Monkey. *Neuron*, *65*(2), 270–279. doi: 10.1016/j.neuron.2010.01.003

Macaluso, E., Frith, C. D., Driver, J., & Kingdom, U. (2002). Crossmodal spatial influences of touch on extrastriate visual areas take current gaze direction into account. *Neuron*, *34*(0896-6273), 647–658. doi: 10.1016/S0896-6273(02)00678-5

Macar, F., Lejeune, H., Bonnet, M., Ferrara, A., Pouthas, V., Vidal, F., & Maquet, P. (2002). Activation of the supplementary motor area and of attentional networks during temporal processing. *Experimental Brain Research*, *142*(4), 475–485.

Machulla, T.-K., Di Luca, M., Froehlich, E., & Ernst, M. O. (2012). Multisensory simultaneity recalibration: storage of the aftereffect in the absence of counterevidence. *Experimental Brain Research*, *217*(1), 89–97. doi: 10.1007/s00221-011-2976-5

Maij, F., Brenner, E., & Smeets, J. B. J. (2011). Peri-saccadic mislocalization is not influenced by the predictability of the saccade target location. *Vision Research*, *51*(1), 154–159. doi: 10.1016/j.visres.2010.10.024

Malapani, C., Rakitin, B., Levy, R., Meck, W. H., Deweer, B., Dubois, B., & Gibbon, J. (1998). Coupled temporal memories in Parkinson's disease: a dopamine-related dysfunction. *Journal of cognitive neuroscience*, *10*(3), 316–331. doi: 10.1162/089892998562762

Mamassian, P., Landy, M. S., & Maloney, L. T. (2002). Bayesian modelling of visual perception. In R. Rao, B. Olshausen, & M. Lewicki (Eds.), *Probabilistic models of the brain; perception and neural function.* (pp. 13–36). Cambridge, MA: MIT Press.

Maquet, P., Lejeune, H., Pouthas, V., Bonnet, M., Casini, L., Macar, F., … Degueldre, C.

(1996). Brain activation induced by estimation of duration: a PET study. *Neuroimage*, *3*(2), 119–126.

Marder, C. P., & Buonomano, D. V. (2003). Differential effects of short- and long-term potentiation on cell firing in the CA1 region of the hippocampus. *Journal of Neuroscience*, *23*(1), 112–121.

Margoliash, D. (1983). Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *3*(5), 1039–1057.

Maricq, A. V., & Church, R. M. (1983). The differential effects of haloperidol and methamphetamine on time estimation in the rat. *Psychopharmacology*, *79*(1), 10–15. doi: 10.1007/BF00433008

Marinovic, W., & Arnold, D. H. (2012). Separable temporal metrics for time perception and anticipatory actions. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1730), 854–859. doi: 10.1098/rspb.2011.1598

Markram, H., Pikus, D., Gupta, A., & Tsodyks, M. (1998). Potential for multiple mechanisms, phenomena and algorithms for synaptic plasticity at single synapses. *Neuropharmacology*, *37*(4-5), 489–500. doi: 10.1016/S0028-3908(98)00049-5

Matell, M. S., King, G. R., & Meck, W. H. (2004). Differential modulation of clock speed by the administration of intermittent versus continuous cocaine. *Behavioral neuroscience*, *118*(1), 150–6. doi: 10.1037/0735-7044.118.1.150

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *BioEssays*, *22*(1), 94–103. doi: 10.1002/(SICI)1521-1878(200001)22:1¡94::AID-BIES14¿3.0.CO;2-E

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Brain research. Cognitive brain research*, *21*(2), 139–70. doi: 10.1016/j.cogbrainres.2004.06.012

Matteson, H. H. (1970). Effects of surround luminance on perceptual latency. *Journal of the Optical Society of America*, *60*(8), 1125–1131. doi: 10.1364/JOSA.61.001169

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual review of neuroscience*, *27*, 307–40. doi: 10.1146/annurev.neuro.27.070203.144247

Mauk, M. D., & Donegan, N. H. (1997). A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. *Learning & memory*, *4*(1), 130–158. doi: 10.1101/lm.4.1.130

Maunsell, J. H., Ghose, G. M., Assad, J. a., McAdams, C. J., Boudreau, C. E., & Noerager, B. D. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Visual neuroscience*, *16*(1), 1–14. doi: 10.1017/S0952523899156177

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. *Journal of experimental psychology. Human perception and performance*, *29*(6), 1102–1125. doi: 10.1037/0096-1523.29.6.1102

McDowell, J. E., Dyckman, K. a., Austin, B. P., & Clementz, B. a. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades: Evidence from studies of humans. *Brain and Cognition*, *68*(3), 255–270. doi: 10.1016/j.bandc.2008.08.016

McGraw, P. V., Whitaker, D., Skillen, J., & Chung, S. T. L. (2002). Motion adaptation distorts perceived visual position. *Current Biology*, *12*(23), 2042–2047. doi: 10.1016/S0960-9822(02)01354-4

Meck, W. H. (1983). Selective adjustment of the speed of internal clock and memory processes. *Journal of experimental psychology. Animal behavior processes*, *9*(2), 171–201. doi: 10.1037/0097-7403.9.2.171

Meck, W. H. (1984). Attentional bias between modalities: effect on the internal clock, memory, and decision stages used in animal time discrimination. *Annals of the New York Academy of Sciences*, *423*, 528–41.

Meck, W. H. (1986). Affinity for the dopamine D2 receptor predicts neuroleptic potency in decreasing the speed of an internal clock. *Pharmacology, Biochemistry, and Behavior*, *25*(6), 1185–1189. doi: 10.1016/0091-3057(86)90109-7

Meck, W. H. (2002). Choline Uptake in the Frontal Cortex Is Proportional to the Absolute Error of a Temporal Memory Translation Constant in Mature and Aged Rats. *Learning and Motivation*, *33*(1), 88–104. doi: 10.1006/lmot.2001.1101

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of experimental psychology. Animal behavior processes*, *9*(3), 320–334. doi: 10.1037/0097-7403.9.3.320

Meck, W. H., & Malapani, C. (2004). Neuroimaging of interval timing. *Brain research. Cognitive brain research*, *21*(2), 133–7. doi: 10.1016/j.cogbrainres.2004.07.010

Meck, W. H., Penney, T. B., & Pouthas, V. (2008). Cortico-striatal representation of time in animals and humans. *Current opinion in neurobiology*, *18*(2), 145–52. doi: 10.1016/j.conb.2008.08.002

Medina, J. F., Garcia, K. S., Nores, W. L., Taylor, N. M., & Mauk, M. D. (2000). Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. *Journal of Neuroscience*, *20*(14), 5516–5525. doi: 20/14/5516 [pii]

Melcher, D. (2005). Spatiotopic transfer of visual-form adaptation across saccadic eye movements. *Current biology*, *15*(19), 1745–8. doi: 10.1016/j.cub.2005.08.044

Melcher, D., & Morrone, M. C. (2003). Spatiotopic temporal integration of visual motion across saccadic eye movements. *Nature neuroscience*, *6*(8), 877–881. doi: 10.1038/nn1098

Merritt, D. J., Casasanto, D., & Brannon, E. M. (2010). Do monkeys think in metaphors? Representations of space and time in monkeys and humans. *Cognition*, *117*(2), 191–202. doi: 10.1016/j.cognition.2010.08.011

Meyerhoff, H. S., Vanes, L. D., & Huff, M. (2015). Spatiotemporal Predictability Alters Perceived Duration of Visual Events : Memento Effect Revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 613–622. doi: 10.1037/xhp0000048

Miall, C. (1989). The Storage of Time Intervals Using Oscillating Neurons. *Neural Computation*, *1*(3), 359–371. doi: 10.1162/neco.1989.1.3.359

Miall, R., & Wolpert, D. (1996). Forward Models for Physiological Motor Control. *Neural Networks*, *9*(8), 1265–1279. doi: 10.1016/S0893-6080(96)00035-4

Michotte, A. (1963). *The perception of causality* (Vol. 6) (No. 8).

Miller, X. R. L., Pluta, S. R., Stein, B. E., & Rowland, B. A. (2015). Relative Unisensory Strength and Timing Predict Their Multisensory Product. *Journal of Neuroscience*, *35*(13), 5213–5220. doi: 10.1523/JNEUROSCI.4771-14.2015

Milne, J. L., Chapman, C. S., Gallivan, J. P., Wood, D. K., Culham, J. C., & Goodale, M. a. (2013). Connecting the dots: Object connectedness deceives perception but not movement planning. *Psychological Science*, *24*(8), 1456–65. doi: 10.1177/0956797612473485

Miniussi, C. (1999). Orienting attention in time: Modulation of brain potentials. *Brain*, *122*(8), 1507–1518. doi: 10.1093/brain/122.8.1507

Mitchell, C. T., & Davis, R. (1987). The perception of time in scale model environments. *Perception*, *16*(1), 5–16.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: Examining temporal ventriloquism. *Cognitive Brain Research*, *17*(1), 154–163. doi: 10.1016/S0926-6410(03)00089-2

Morel, a., Garraghty, P. E., & Kaas, J. H. (1993). Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *The Journal of comparative neurology*, *335*(3), 437–459. doi: 10.1002/cne.903350312

Morgan, M. J., Raphael, S., Tibber, M. S., & Dakin, S. C. (2014). A texture-processing model of the 'visual sense of number'. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1790), 20141137–20141137. doi: 10.1098/rspb.2014.1137

Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, *376*(6540), 507–509. doi: 10.1038/376507a0

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature neuroscience*, *8*(7), 950–4. doi: 10.1038/nn1488

Morrone, M. C., Ross, J., & Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *Journal of Neuroscience*, *17*(20), 7941–53.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature neuroscience*, *3*(12), 1322–8. doi: 10.1038/81860

Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. *Journal of experimental psychology. Human perception and performance*, *15*(2), 315–30. doi: 10.1037/0096-1523.15.2.315

Murakami, I. (2001). A flash-lag effect in random motion. *Vision Research*, *41*(24), 3101–3119. doi: 10.1016/S0042-6989(01)00193-6

Nakamura, K., & Colby, C. L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(6), 4026–4031. doi: 10.1073/pnas.052379899

Newberry, N. R., & Nicoll, R. A. (1984). A bicuculline-resistant inhibitory post-synaptic potential in rat hippocampal pyramidal cells in vitro. *The Journal of Physiology*, *348*(1), 239–254. doi: 10.1113/jphysiol.1984.sp015107

Nicol, J. R., Watter, S., Gray, K., & Shore, D. I. (2009). Object-based perception mediates the effect of exogenous attention on temporal resolution. *Visual Cognition*, *17*(4), 555–573. doi: 10.1080/13506280802113860

Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proceedings of the National Academy of Sciences*, *109*(29), 11860–11865. doi: 10.1073/pnas.1204580109

Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and Spatial Enumeration Processes in the Primate Parietal Cortex. *Science*, *313*(5792), 1431–1435. doi: 10.1126/science.1130308

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science (New York, N.Y.)*, *297*(5587), 1708–1711. doi: 10.1126/science.1072493

Nieder, A., & Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(19), 7457–7462. doi: 10.1073/pnas.0402239101

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*(6487), 256–257. doi: 10.1038/370256b0

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences*, *6*(9), 387–393. doi: 10.1016/S1364-6613(02)01963-0

Nijhawan, R. (2008). Visual prediction: psychophysics and neurophysiology of compensation for time delays. *The Behavioral and brain sciences*, *31*(2), 179–198; discussion 198–239. doi: 10.1017/S0140525X08003804

Nishida, S., & Johnston, A. (1999). Influence of motion signals on the perceived position of spatial pattern. *Nature*, *397*(6720), 610–612. doi: 10.1038/17600

Ong, W. S., Hooshvar, N., Zhang, M., & Bisley, J. W. (2009). Psychophysical evidence for spatiotopic processing in area MT in a short-term memory for motion task. *Journal* *of neurophysiology*, *102*(4), 2435–40. doi: 10.1152/jn.00684.2009

Ostendorf, F., Fischer, C., Finke, C., & Ploner, C. J. (2007). Perisaccadic Compression Correlates with Saccadic Peak Velocity: Differential Association of Eye Movement Dynamics with Perceptual Mislocalization Patterns. *Journal of Neuroscience*, *27*(28), 7559–7563. doi: 10.1523/JNEUROSCI.2074-07.2007

Ostendorf, F., Fischer, C., Gaymard, B., & Ploner, C. J. (2006). Perisaccadic mislocalization without saccadic eye movements. *Neuroscience*, *137*(3), 737–745. doi: 10.1016/j.neuroscience.2005.09.032

Panichi, M., Burr, D., Morrone, M. C., & Baldassi, S. (2012). Spatiotemporal dynamics of perisaccadic remapping in humans revealed by classification images. *Journal of Vision*, *12*(4), 11–11. doi: 10.1167/12.4.11

Pariyadath, V., & Eagleman, D. (2007). The effect of predictability on subjective duration. *PloS one*, *2*(11), e1264. doi: 10.1371/journal.pone.0001264

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature neuroscience*, *4*(7), 739–744. doi: 10.1038/89532

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial vision*, *10*(4), 437–442. doi: 10.1163/156856897X00366

Pelli, D. G. (2008). Crowding: a cortical constraint on object recognition. *Current Opinion in Neurobiology*, *18*(4), 445–451. doi: 10.1016/j.conb.2008.09.008

Pepperberg, I. M. (2006). Cognitive and communicative abilities of Grey parrots. *Applied Animal Behaviour Science*, *100*(1-2), 77–86. doi: 10.1016/j.applanim.2006.04.005

Perrett, S. P., Ruiz, B. P., & Mauk, M. D. (1993). Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses. *Journal of Neuroscience*, *13*(4), 1708–1718.

Peterken, C., Brown, B., & Bowman, K. (1991). Predicting the future position of a moving target. *Perception*, *20*(1), 5–16. doi: 10.1068/p200005

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547–555. doi: 10.1016/j.neuron.2004.10.014

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A Magnitude Code Common to Numerosities and Number Symbols in Human Intraparietal Cortex. *Neuron*, *53*(2), 293–305. doi: 10.1016/j.neuron.2006.11.022

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*(6), 983–993.

Pinto, Y., van Gaal, S., de Lange, F. P., Lamme, V. a. F., & Seth, A. K. (2015). Expectations accelerate entry of visual stimuli into awareness. *Journal of Vision*, *15*(8), 13. doi: 10.1167/15.8.13

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25. doi: 10.1080/00335558008248231

Prablanc, C., & Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *Journal of neurophysiology*, *67*(2), 455–69.

Rabinovich, M., Huerta, R., & Laurent, G. (2008). Transient dynamics for neural processing. *Science*, *321*(July), 48–50. doi: 10.1126/science.1155564

Rammsayer, T. H. (1993). On dopaminergic modulation of temporal information processing. *Biological psychology*, *36*(3), 209–22. doi: 10.1016/0301-0511(93)90018-4

Rammsayer, T. H. (2006). Effects of pharmacologically induced changes in NMDA receptor activity on human timing and sensorimotor performance. *Brain Research*, *1073-1074*, 407–416. doi: 10.1016/j.brainres.2005.12.019

Rammsayer, T. H. (2008). Neuropharmacological approaches to human timing. In S. Grondin (Ed.), *Psychology of time* (pp. 295–320). Bingley, U.K.: Emerald Group.

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: cognitive and perceptual factors. *Perception & psychophysics*, *50*(6), 565–74. doi: 10.3758/BF03207541

Rammsayer, T. H., & Verner, M. (2014). The effect of nontemporal stimulus size on perceived duration as assessed by the method of reproduction. *Journal of Vision*, *14*(5), 1–10. doi: 10.1167/14.5.17.doi

Rammsayer, T. H., & Verner, M. (2015). Larger visual stimuli are perceived to last longer from time to time: The internal clock is not affected by nontemporal visual stimulus size. *Journal of Vision*, *15*(3), 1–11. doi: 10.1167/15.3.5.doi

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

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature neuroscience*, *4*(3), 317–23. doi: 10.1038/85191

Reale, R. a., & Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat. *The Journal of comparative neurology*, *192*(2), 265–91. doi: 10.1002/cne.901920207

Reutimann, J. (2004). Climbing Neuronal Activity as an Event-Based Cortical Representation of Time. *Journal of Neuroscience*, *24*(13), 3295–3303. doi: 10.1523/JNEUROSCI.4098-03.2004

Reyes, a., & Sakmann, B. (1999). Developmental switch in the short-term modification of unitary EPSPs evoked in layer 2/3 and layer 5 pyramidal neurons of rat neocortex. *Journal of Neuroscience*, *19*(10), 3827–3835.

Roelofs, C., & Zeeman, W. (1951). Influence of different sequences of optical stimuli on the estimation of duration of a given interval of time. *Acta Psychologica*, *8*, 89–128. doi: 10.1016/0001-6918(51)90007-8

Roitman, J., & Brannon, E. (2003). Nonverbal Representations of Time and Number in Animals and Human Infants. In *Animals* (pp. 143–182). doi: 10.1201/9780203009574.ch6

Roitman, J. D., Brannon, E. M., & Platt, M. L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biology*, *5*(8), 1672–1682. doi: 10.1371/journal.pbio.0050208

Roitman, J. D., Brannon, E. M., & Platt, M. L. (2012). Representation of numerosity in posterior parietal cortex. *Frontiers in Integrative Neuroscience*, *6*(May), 1–9. doi: 10.3389/fnint.2012.00025

Rolfs, M., & Vitu, F. (2007). On the limited role of target onset in the gap task: support for the motor-preparation hypothesis. *Journal of vision*, *7*(10), 7.1–20. doi: 10.1167/7.10.7

Rolke, B., Dinkelbach, A., Hein, E., & Ulrich, R. (2007). Does attention impair temporal discrimination? Examining non-attentional accounts. *Psychological Research*, *72*(1), 49–60. doi: 10.1007/s00426-006-0092-0

Rolke, B., & Hofmann, P. (2007). Temporal uncertainty degrades perceptual processing. *Psychonomic bulletin & review*, *14*(3), 522–526. doi: 10.3758/BF03194101

Roseboom, W., Nishida, S., & Arnold, D. H. (2009). The sliding window of audio-visual simultaneity. *Journal of Vision*, *9*(12), 4–4. doi: 10.1167/9.12.4

Rosenbaum, D. A. (1975). Perception and extrapolation of velocity and acceleration. *Journal of experimental psychology. Human perception and performance*, *1*(4), 395–403. doi: 10.1037/0096-1523.1.4.395

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature*, *386*(6625), 598–601. doi: 10.1038/386598a0

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in neurosciences*, *24*(2), 113–121. doi: 10.1016/S0166-2236(00)01685-4

Rossetti, Y., Jacquin-Courtois, S., Rode, G., Ota, H., Michel, C., & Boisson, D. (2004). Does action make the link between number and space representation? Visuo-manual adaptation improves number bisection in unilateral neglect. *Psychological science*, *15*(6), 426–30. doi: 10.1111/j.0956-7976.2004.00696.x

Roth, M. J., Synofzik, M., & Lindner, A. (2013). The cerebellum optimizes perceptual predictions about external sensory events. *Current Biology*, *23*(10), 930–935. doi: 10.1016/j.cub.2013.04.027

Roy, J. E., Riesenhuber, M., Poggio, T., & Miller, E. K. (2010). Prefrontal Cortex Activity during Flexible Categorization. *Journal of Neuroscience*, *30*(25), 8519–8528. doi: 10.1523/JNEUROSCI.4837-09.2010

Rushworth, M. F. S., Krams, M., & Passingham, R. E. (2001). The attentional role of the left parietal cortex: the distinct lateralization and localization of motor attention in the human brain. *Journal of Cognitive Neuroscience*, *13*(5), 698–710.

Ruskin, D. N., Bergstrom, D. a., Kaneoke, Y., Patel, B. N., Twery, M. J., & Walters, J. R.

(1999). Multisecond oscillations in firing rate in the basal ganglia: robust modulation by dopamine receptor activation and anesthesia. *Journal of neurophysiology*, *81*(5), 2046–55. doi: 0022-3077

Sadagopan, S., & Wang, X. (2009). Nonlinear Spectrotemporal Interactions Underlying Selectivity for Complex Sounds in Auditory Cortex. *Journal of Neuroscience*, *29*(36), 11192–11202. doi: 10.1523/JNEUROSCI.1286-09.2009

Salvioni, P., Murray, M. M., Kalmbach, L., & Bueti, D. (2013). How the Visual Brain Encodes and Keeps Track of Time. *Journal of Neuroscience*, *33*(30), 12423–12429. doi: 10.1523/JNEUROSCI.5146-12.2013

Sawamura, H., Shima, K., & Tanji, J. (2002). Numerical representation for action in the parietal cortex of the monkey. *Nature*, *415*(February), 918–922. doi: 10.1038/415918a

Schiff, W., & Oldak, R. (1990). Accuracy of judging time to arrival: effects of modality, trajectory, and gender. *Journal of experimental psychology. Human perception and performance*, *16*(2), 303–316. doi: 10.1037/0096-1523.16.2.303

Schiller, P., & Sandell, J. (1983). Interactions between visually and electrically elicited saccades before and after superior colliculus and frontal eye field ablations in the rhesus monkey. *Experimental Brain Research*, *49*(3), 381–392. doi: 10.1007/BF00238780

Schnier, F., & Lappe, M. (2012). Mislocalization of stationary and flashed bars after saccadic inward and outward adaptation of reactive saccades. *Journal of neurophysiology*, *107*(11), 3062–3070. doi: 10.1152/jn.00877.2011

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309. doi: 10.1016/S1364-6613(00)01506-0

Schultz, W. (2000). Multiple reward signals in the brain. *Nature reviews. Neuroscience*, *1*(3), 199–207. doi: 10.1038/35044563

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*, 1593–1599. doi: 10.1126/science.275.5306.1593

Seibold, V. C., Fiedler, A., & Rolke, B. (2011). Temporal attention shortens perceptual latency: A temporal prior entry effect. *Psychophysiology*, *48*(5), 708–717. doi: 10.1111/j.1469-8986.2010.01135.x

Seilheimer, R. L., Rosenberg, A., & Angelaki, D. E. (2014). Models and processes of multisensory cue combination. *Current Opinion in Neurobiology*, *25*, 38–46. doi: 10.1016/j.conb.2013.11.008

Sekuler, R., Sekuler, a. B., & Lau, R. (1997). Sound alters visual motion perception. *Nature*, *385*(6614), 308. doi: 10.1038/385308a0

Seriès, P., & Seitz, A. R. (2013). Learning what to expect (in visual perception). *Frontiers in human neuroscience*, *7*(October), 668. doi: 10.3389/fnhum.2013.00668

Shimojo, S. (2001). Sensory modalities are not separate modalities: plasticity and interactions. *Current Opinion in Neurobiology*, *11*(4), 505–509. doi: 10.1016/S0959-4388(00)00241-5

Simon, O., Mangin, J.-F., Cohen, L., Le Bihan, D., & Dehaene, S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron*, *33*(3), 475–487.

Sincich, L. C., Park, K. F., Wohlgemuth, M. J., & Horton, J. C. (2004). Bypassing V1: a direct geniculate input to area MT. *Nature neuroscience*, *7*(10), 1123–8. doi: 10.1038/nn1318

Smith, A. T., Wall, M. B., Williams, A. L., & Singh, K. D. (2006). Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience*, *23*(2), 561–569. doi: 10.1111/j.1460-9568.2005.04526.x

Snowden, R. J. (1998). Shifts in perceived position following adaptation to visual motion. *Current biology*, *8*(24), 1343–1345. doi: 10.1016/S0960-9822(07)00567-2

Snyder, J. S., & Alain, C. (2007). Toward a neurophysiological theory of auditory stream segregation. *Psychological bulletin*, *133*(5), 780–799. doi: 10.1037/0033-2909.133.5.780

Sokolov, A. N., Ehrenstein, W. H., Pavlova, M. a., & Cavonius, C. R. (1997). Motion extrapolation and velocity transposition. *Perception*, *26*(7), 875–889. doi: 10.1068/p260875

Solomon, S. G., & Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. *Current biology : CB*, *24*(20), R1012–22. doi: 10.1016/j.cub.2014.09.001

Sommer, M. a., & Wurtz, R. H. (2002). A pathway in primate brain for internal monitoring of movements. *Science (New York, N.Y.)*, *296*(5572), 1480–1482. doi: 10.1126/science.1069590

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*(7117), 374–377. doi: 10.1038/nature05279

Sommer, M. A., & Wurtz, R. H. (2008). Visual perception and corollary discharge. *Perception*, *37*(3), 408–18. doi: 10.1068/p5873

Sparks, D. L., & Mays, L. E. (1983). Spatial localization of saccade targets. I. Compensation for stimulation-induced perturbations in eye position. *Journal of neurophysiology*, *49*(1), 45–63.

Sparks, D. L., Mays, L. E., & Porter, J. D. (1987). Eye movements induced by pontine stimulation: interaction with visually triggered saccades. *Journal of neurophysiology*, *58*(2), 300–18.

Spencer, R. M. C., Karmarkar, U., & Ivry, R. B. (2009). Evaluating dedicated and intrinsic models of temporal encoding by varying context. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *364*(1525), 1853–1863. doi: 10.1098/rstb.2009.0024

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of comparative and physiological psychology*, *43*, 482–489. doi: 10.1037/h0055479

Stavy, R., & Tirosh, D. (2000). *How students (mis-) understand science and mathematics:*

Intuitive rules. Teachers College Press.

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of cognitive neuroscience*, *19*(12), 1964–1973. doi: 10.1162/jocn.2007.91213

Stevens, J. K., Emerson, R. C., Gerstein, G. L., Kallos, T., Neufeld, G. R., Nichols, C. W., & Rosenquist, A. C. (1976). Paralysis of the awake human: visual perceptions. *Vision research*, *16*(1), 93–98.

Stevens, S. S. (1971). *Psychophysics.* John Wiley.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. doi: 10.1016/j.tics.2009.06.003

Synofzik, M., Lindner, A., & Thier, P. (2008). The Cerebellum Updates Predictions about the Visual Consequences of One's Behavior. *Current Biology*, *18*(11), 814–818. doi: 10.1016/j.cub.2008.04.071

Tanaka, Y., & Shimojo, S. (1996). Location vs feature: Reaction time reveals dissociation between two visual functions. *Vision Research*, *36*(14), 2125–2140. doi: 10.1016/0042-6989(95)00272-3

Terry, K. M., Valdes, L. a., & Neill, W. T. (1994). Does "inhibition of return" occur in discrimination tasks? *Perception & psychophysics*, *55*(3), 279–286. doi: 10.3758/BF03207599

Thomaschke, R., & Dreisbach, G. (2013). Temporal predictability facilitates action, not perception. *Psychological science*, *24*(7), 1335–40. doi: 10.1177/0956797612469411

Thompson, P. (1981). Velocity after-effects: The effects of adaptation to moving stimuli on the perception of subsequently seen moving stimuli. *Vision Research*, *21*(3), 337–345. doi: 10.1016/0042-6989(81)90161-9

Tibber, M. S., Greenwood, J. a., & Dakin, S. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *Journal of Vision*, *12*(6), 1–19. doi: 10.1167/12.6.8

Tolias, A. S., Moore, T., Smirnakis, S. M., Tehovnik, E. J., Siapas, A. G., & Schiller, P. H. (2001). Eye movements modulate visual receptive fields of V4 neurons. *Neuron*, *29*(3), 757–767. doi: 10.1016/S0896-6273(01)00250-1

Treisman, M. (1963). Temporal discrimination and the indifference interval. Implications for a model of the "internal clock". *Psychological monographs*, *77*(13), 1–31.

Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & psychophysics*, *66*(7), 1171–1189. doi: 10.3758/BF03196844

Turi, M., & Burr, D. (2012). Spatiotopic perceptual maps in humans: evidence from motion adaptation. *Proceedings of the Royal Society B: Biological Sciences*, *279*(April), 3091–3097. doi: 10.1098/rspb.2012.0637

Ulrich, R., Nitschke, J., & Rammsayer, T. H. (2006). Perceived duration of expected and unexpected stimuli. *Psychological research*, *70*(2), 77–87. doi: 10.1007/s00426-004-0195-4

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of neurophysiology*, *78*(3), 1373–83.

Umeno, M. M., & Goldberg, M. E. (2001). Spatial processing in the monkey frontal eye field. II. Memory responses. *Journal of neurophysiology*, *86*(5), 2344–2352.

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid Recalibration to Audiovisual Asynchrony. *Journal of Neuroscience*, *33*(37), 14633–14637. doi: 10.1523/JNEUROSCI.1182-13.2013

Van der Burg, E., Cass, J., & Alais, D. (2014). Window of audio-visual simultaneity is unaffected by spatio-temporal visual clutter. *Scientific Reports*, *4*, 1–7. doi: 10.1038/srep05098

van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. *Perception & psychophysics*, *70*(6), 955–968. doi: 10.3758/PP.70.6.955

van Eijk, R. L. J., Kohlrausch, A., Juola, J. F., & van de Par, S. (2010). Temporal order judgment criteria are affected by synchrony judgment sensitivity. *Attention, perception & psychophysics*, *72*(8), 2227–2235. doi: 10.3758/BF03196697

van Rossum, M. C. (2001). A novel spike distance. *Neural computation*, *13*(4), 751–763. doi: 10.1162/089976601300014321

Vatakis, A., & Spence, C. (2008). Evaluating the influence of the 'unity assumption' on the temporal perception of realistic audiovisual stimuli. *Acta Psychologica*, *127*(1), 12–23. doi: 10.1016/j.actpsy.2006.12.002

Viswanathan, P., & Nieder, A. (2013). Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(27), 11187–92. doi: 10.1073/pnas.1308141110

Volkmann, F. C. (1962). Vision during voluntary saccadic eye movements. *Journal of the Optical Society of America*, *52*, 571–578.

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip - Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Die Naturwissenschaften*, *37*(20), 464–476. doi: 10.1007/BF00622503

Vroomen, J., & de Gelder, B. (2004). Temporal ventriloquism: sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(3), 513–518. doi: 10.1037/0096-1523.30.3.513

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Attention, perception & psychophysics*, *72*(4), 871–884. doi: 10.3758/APP.72.4.871

Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, *22*(1), 32–35. doi: 10.1016/j.cogbrainres.2004.07.003

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. doi: 10.1037/a0029333

Wagemans, J., van Lier, R., & Scholl, B. J. (2006). Introduction to Michotte's heritage in perception and cognition research. *Acta Psychologica*, *123*(1-2), 1–19. doi: 10.1016/j.actpsy.2006.06.003

Walker, M. F., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *Journal of neurophysiology*, *73*(5), 1988–2003.

Walker, R., Walker, D. G., Husain, M., & Kennard, C. (2000). Control of voluntary and reflexive saccades. *Experimental brain research*, *130*(4), 540–4.

Wallace, M. T., & Stevenson, R. a. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*, *64*, 105–123. doi: 10.1016/j.neuropsychologia.2014.08.005

Wallach, H., & Lewis, C. (1966). The effect of abnormal displacement of the retinal image during eye movements. *Perception & Psychophysics*, *1*(1), 25–29. doi: 10.3758/BF03207816

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488. doi: 10.1016/j.tics.2003.09.002

Walsh, V., & Pasual-Leone, A. (2003). *Transcranial Magnetic Stimulation: A Neurochronometrics of Mind.* MIT Press.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. (1969). Auditory sequence: confusion of patterns other than speech or music. *Science (New York, N.Y.)*, *164*(879), 586–587. doi: 10.1126/science.164.3879.586

Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & psychophysics*, *33*(2), 113–120. doi: 10.3758/BF03202828

Wearden, J. H. (2003). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.), *Time and mind ii* (pp. 21–39). Gottingen: Hogrefe & Huber.

Weber, E. H. (1851). *Annotationes anatomicae et physiologicae.* Leipsig: CF Koehler.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological bulletin*, *88*(3), 638–667. doi: 10.1037/0033-2909.88.3.638

Wenderoth, P., & Wiese, M. (2008). Retinotopic encoding of the direction aftereffect. *Vision Research*, *48*(19), 1949–1954. doi: 10.1016/j.visres.2008.06.013

Whalen, J., Gelman, R., & Gallistel, C. (1999). Nonverbal Counting in Humans: The Psychophysics of Number Representation. *Psychological Science*, *10*, 130–137. doi: 10.1111/1467-9280.00120

Wickens, J., Begg, A., & Arbuthnott, G. (1996). Dopamine reverses the depression of rat corticostriatal synapses which normally follows high-frequency stimulation of cortex In vitro. *Neuroscience*, *70*(1), 1–5. doi: 10.1016/0306-4522(95)00436-M

Wittenberg, M., Bremmer, F., & Wachtler, T. (2008). Perceptual evidence for saccadic updating of color stimuli. *Journal of vision*, *8*(14), 9.1–9. doi: 10.1167/8.14.9

Wolpert, D. M., & Flanagan, J. (2001). Motor prediction. *Current Biology*, *11*(18), R729–R732. doi: 10.1016/S0960-9822(01)00432-8

Woodworth, R. (1906). Vision and localization during eye movements. *Psychological bulletin*, *3*.

Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of vision*, *8*(3), 24.1–11. doi: 10.1167/8.3.24

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, *48*(20), 2070–2089. doi: 10.1016/j.visres.2008.03.021

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-NOOPmonth-old infants. *Cognition*, *74*(1), B1–B11. doi: 10.1016/S0010-0277(99)00066-9

Xuan, B., Chen, X.-C., He, S., & Zhang, D.-R. (2009). Numerical magnitude modulates temporal comparison: An ERP study. *Brain Research*, *1269*, 135–142. doi: 10.1016/j.brainres.2009.03.016

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of vision*, *7*(10), 2.1–5. doi: 10.1167/7.10.2

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *Journal of experimental psychology. Human perception and performance*, *10*(5), 601–621. doi: 10.1037/0096-1523.10.5.601

Yarrow, K., Haggard, P., Heal, R., Brown, P., & Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. *Nature*, *414*(6861), 302–305. doi: 10.1038/35104551

Yeshurun, Y. (2004). Isoluminant stimuli and red background attenuate the effects of transient spatial attention on temporal resolution. *Vision Research*, *44*(12), 1375–1387. doi: 10.1016/j.visres.2003.12.016

Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. *Psychological Science*, *14*(3), 225–231. doi: 10.1111/1467-9280.02436

Yeshurun, Y., & Sabo, G. (2012). Differential effects of transient attention on inferred parvocellular and magnocellular processing. *Vision Research*, *74*, 21–29. doi: 10.1016/j.visres.2012.06.006

Zimmermann, E., Born, S., Fink, G. R., & Cavanagh, P. (2014). Masking produces compression of space and time in the absence of eye movements. *Journal of neurophysiology*, *112*(12), 3066–3076. doi: 10.1152/jn.00156.2014

Zimmermann, E., & Lappe, M. (2009). Mislocalization of Flashed and Stationary Visual Stimuli after Adaptation of Reactive and Scanning Saccades. *Journal of Neuroscience*, *29*(35), 11055–11064. doi: 10.1523/JNEUROSCI.1604-09.2009

Zimmermann, E., & Lappe, M. (2010). Motor signals in visual localization. *Journal of vision*, *10*, 1–11. doi: 10.1167/10.6.2.Introduction

Zimmermann, E., & Lappe, M. (2011). Eye Position Effects in Oculomotor Plasticity and Visual Localization. *Journal of Neuroscience*, *31*(20), 7341–7348. doi: 10.1523/JNEUROSCI.6112-10.2011

Zimmermann, E., Morrone, M. C., & Burr, D. C. (2014). The visual component to saccadic compression. *Journal of Vision*, *14*(12), 13–. doi: 10.1167/14.12.13.doi

Zimmermann, E., Morrone, M. C., Fink, G. R., & Burr, D. (2013). Spatiotopic neural representations develop slowly across saccades. *Current Biology*, *23*(5), R193–R194. doi: 10.1016/j.cub.2013.01.065

Zirnsak, M., Steinmetz, N. a., Noudoost, B., Xu, K. Z., & Moore, T. (2014). Visual space is compressed in prefrontal cortex before eye movements. *Nature*, *507*(7493), 504–7. doi: 10.1038/nature13149

Zorzi, M., Bonato, M., Treccani, B., Scalambrin, G., Marenzi, R., & Priftis, K. (2012). Neglect impairs explicit processing of the mental number line. *Frontiers in human neuroscience*, *6*, 125. doi: 10.3389/fnhum.2012.00125

Zorzi, M., Priftis, K., Meneghello, F., Marenzi, R., & Umiltà, C. (2006). The spatial representation of numerical and non-numerical sequences: evidence from neglect. *Neuropsychologia*, *44*(7), 1061–1067.

Zuber, B. L., & Stark, L. (1966). Saccadic suppression: elevation of visual threshold associated with saccadic eye movements. *Experimental neurology*, *16*(1), 65–79.

Zucker, R. S., & Regehr, W. G. (2002). Short-term synaptic plasticity. *Annual Review of Physiology*, *64*, 355–405. doi: DOI 10.1146/annurev.physiol.64.092501.114547