

# VU Research Portal

## Updating spatial working memory in a dynamic visual environment

Boon, P.J.

2018

### **document version**

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

### **citation for published version (APA)**

Boon, P. J. (2018). *Updating spatial working memory in a dynamic visual environment*. [PhD-Thesis - Research and graduation internal, Vrije Universiteit Amsterdam].

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

### **E-mail address:**

[vuresearchportal.ub@vu.nl](mailto:vuresearchportal.ub@vu.nl)

# Chapter 1

## Introduction

Have you ever tried riding a bike through the centre of Amsterdam? Then you know how it feels to zigzag among hordes of fellow cyclists, evade scooters that come out of nowhere at high speeds, cars that drive so close that they almost brush against your leg and tourists that tend to step out of a tram right in front of you. And I am not even talking about busses, trams and taxis that obey their own rules. Somehow, we possess this amazing ability to keep track of multiple (moving) objects in the environment, while at the same time moving our eyes, head and bodies to extract relevant information. How is our brain able to accomplish this seemingly impossible task?

Although perceiving and interacting with the visual world around us often seems effortless, this ability is not as straightforward as it might appear. Our brain is not able to process all information in parallel, so we somehow have to select the most relevant information for further processing. The visual system has evolved to optimize information processing by reducing the high resolution processing to a very small part of the retina. Luckily, we developed a sophisticated eye movement system capable to quickly, “just in time”, orient the fovea to the objects of interest. In this way, our eye movements serve as the first filter of visual information.

Interacting with the world around us involves making 3 to 4 of these eye movements per second. The majority of these movements are fast ballistic movements called saccades, which take only 20 to 40 ms to complete. In everyday life we are seldom aware of these rapid movements, since we are busy “looking” and not “moving our eyes”. Therefore, we require a representation of the locations of relevant objects in the surroundings that can act as a basis for directing the eyes. The problem is that after each eye movement objects in a scene fall on a different part of the retina, and are thus represented by a different set of neurons. Despite these continuous changes in sensory input we are able to interact with the world around us with impressive accuracy.

The main aim of this doctoral thesis is to yield a better understanding of how we keep track of relevant information in the visual environment. In the upcoming chapters I will discuss the role of the eye movement system in maintaining spatial information (Chapter 2), and the mechanisms underlying the updating of this information during both self-induced (eye) movements (Chapter 3 and 4) and during movement of objects in the world around us (Chapter 5). Before proceeding to the chapters, I present my perspective on the state-of-the-art of the problem of maintaining relevant information across saccadic eye movements.

## Vision as an active process

### A “picture-in-the-head”?

The rise of computers in the 1950s marked a major change in thinking about human information processing. In contrast to the then prevailing behaviorism (Skinner, 2011), psychologists came to think of cognition as computation. Analogous to a computer, the brain was thought to receive input about the outside world through its ‘sensors’. This information is subsequently translated into a mental representation of the world around us, which can then be processed and manipulated by the mind. Ultimately, such *thinking* can lead to the output in the form of a motor act.

This so called *cognitivist* paradigm (Mandler, 2002; Neisser, 1967) has profoundly influenced the way we think about processing and selecting visual information. It led to a dominant view in vision science that when we see, an internal representation of the outside world is set up somewhere in the brain. This representational idea was epitomized by David Marr (1982), who described vision as a hierarchical computational process that transforms the raw retinal image step by step into a three-dimensional representation of the world.

For a long time this cognitivist account of vision remained undisputed. Perhaps, because the ‘picture-in-the-head’ idea gets so convincingly endorsed by the common sense. Since we have the feeling that we can see everything there is to be seen, it seems reasonable to assume that the brain holds, somewhere, a complete and detailed representation of what happens in front of our eyes. It was only when this image-processing approach was applied in the area of robotics that it became apparent that it might not be suitable to explain real-life behavior. Attempts to implement computer vision led to the realization of the immense computational resources it requires to process a complete image of the environment. Although early robots already displayed some impressive reasoning skills, the problems of perception and action had severely been underestimated. The most famous example of this was the robot Shakey, an enormous steel computer on wheels which was developed in the late 1960s. Although dubbed the ‘first electronic person’, its performance fell far too short from human’s ability; the processing of visual input and subsequent selection of an appropriate action could take up to several minutes (after which the machine would suddenly start moving, with the risk of accidentally breaking a researcher's leg). This phenomenon is known as Moravec’s paradox (Moravec, 1988); tasks that are considered to be very hard, like playing chess or proving a mathematical theorem, require relatively little computation and are easy to solve by a computer. However, activities that are

generally not thought of to require a lot of intelligence, like moving the chess piece or recognizing a chair, require enormous computational resources. This illustrates the sophistication of our sensorimotor apparatus, which is not solely relying on incoming information, but also on predictions based on prior experience with sensory and motor behaviour (Körding & Wolpert, 2004; Slijper, Richter, Over, Smeets, & Frens, 2009). Although the computational approach has led to many valuable insights about for example formal reasoning and memory, we must acknowledge that it does a very poor job in explaining the seemingly effortless sensorimotor skills that we display as we move through daily life.

### **Change blindness**

Despite the obvious computational problems, the idea of a 'picture-in-the-head' has been very pervasive among vision scientists. However, during the last decades a considerable volume of research revealed the true sparseness of our visual representation. Most notably, observers fail to notice large changes in the visual scene if they are timed to coincide with short interruptions of viewing. Film makers have known this for a long time. Most movies we watch contain a lot of continuity errors; sudden changes that occur during an editorial cut or change of camera angle. A plane engine might suddenly disappear (*The usual suspects*) or a broken helicopter windshield suddenly becomes intact again (*Terminator 2*). Unless someone points it out to us we never notice these things. Simons and Levin (1997) experimentally showed that changes in colour, presence, position and identity of objects in a movie are hardly ever detected. Even the main actor can be replaced during an editorial cut without most people noticing. Interestingly, observers also fail to detect changes in a scene when camera position maintains the same, but viewing is briefly interrupted by shortly presenting a blank screen (Rensink, O'Regan, & Clark, 1997), or when changes are introduced during blinks (O'Regan, Deubel, Clark, & Rensink, 2000).

The failure to detect changes across brief interruptions of viewing shows that our experience of a detailed visual representation is in fact a mirage. Although this goes against our common sense, it is not surprising if one takes into account that this "picture-in-the-head" perspective completely neglects two defining features of human vision (and that of most other vertebrates): 1) the mobility of the eye and 2) the inhomogeneity of the visual system. In the time between waking up in the morning and going to bed at night we make on average 230.000 fast, ballistic eye movements. The necessity of these saccades lies in the distribution of the light receptors on the retina. These photoreceptors are not evenly spread. The rods, which are responsible for motion perception and vision in poor light conditions,

are located mainly in the periphery. However, the cones, the cells responsible for detecting color and form, are densely packed together in one region of the retina - the fovea, allowing us to detect color and form in high resolution. On the downside, only stimuli in a small region of space can be processed with a high resolution (approximately two visual degrees). This was elegantly demonstrated in a study by Freeman & Simoncelli (2011) in which people had to judge whether two sequentially presented images were exactly the same. As long as the large scale image statistics remained the same, virtually everything outside of the fixation area can be scrambled and mixed up without people noticing the image had changed.

We are dependent on eye movements to sample detailed information from the environment. However, these eye movements also disrupt the continuity of viewing. Between the onset and landing of a saccade we are virtually blind, so saccades induce 'cuts' in viewing similar to the artificially induced interruptions in a film edit. Research shows that we maintain very little information between successive fixations. In one of the earliest demonstrations of this phenomenon observers read a text that alternated case with each letter (LiKe ThiS FoR eXaMplE). During each saccade, all letters changed their case. Still, subjects had no problem continuing reading. Moreover, they did not even notice that the visual form of the words continuously changed (McConkie & Zola, 1979). In addition to this, also large spatial changes go unnoticed when they occur during a saccade (e.g., Bray, Bansal, & Joiner, 2016; Bridgeman, Hendry, & Stark, 1975; Collins, Rolfs, Deubel, & Cavanagh, 2009; Deubel, Schneider, & Bridgeman, 1996; Li & Matin, 1990). Even if objects are displaced over a distance of one third of the size of an eye movement, observers tend to miss that anything has changed in the visual scene. These findings show that we do not construct any pictures-in-the-head. Instead, perception constitutes of a series of small snapshots from the environment, and little information is maintained across these fixations.

### **An active vision alternative**

Despite the obvious necessity to move the eyes, eye movements are often neglected in cognitive research. The standard view of perceptual processing describes perception and action as independent entities. The idea that the brain first actively constructs conscious visual perception and only later (if at all) prepares an action, was challenged by J.J. Gibson (1979). He claimed that perception cannot be studied without taking later operations into account. Or, as he put it in his influential book; "The vast quantity of experimental research (...) is concerned with snapshot vision, fixed-eye vision, or aperture vision, and it is not

relevant” (Gibson, 1979). It is not relevant, because in natural situations people never watch a scene with their head mounted in a chin rest, passively fixating a dot on a computer screen. Instead, people are in constant interaction with their environment and perception and action are closely intertwined. His ‘ecological approach’ describes perception as “(...) whole body activities devoted to actively extract, isolate, and clarify informative structure in the world”. Instead of the passive collection of information, vision is an active process. On basis of our perception we select the goal of our next action, which in turn determines the content of perception.

It is not only the selection of information from the visual environment that necessitates eye movements. We have to move our eyes in order to see. Even when we keep our gaze directed at one location we make small jittery eye movements. In the 1950s, different researchers showed that these microsaccades are crucial for perception. When these eye movements are eliminated, for example by placing an image on a suction cup directly mounted on the eyeball (Yarbus, 1967), the visual image quickly fades away (Ditchburn & Ginsborg, 1952; Riggs & Ratliff, 1952; Yarbus, 1967). The visual system simply cannot deal with a completely static retinal image. The steady illumination of retinal receptors results in neural adaptation, which can be counteracted by making small eye movements (Martinez-Conde, Otero-Millan, & Macknik, 2013).

### **Is covert attention really covert?**

Perception is shaped by our eye movements, which serve as the first filter for visual information. However, many theories of visual attention assume eye movements do not play a crucial role in selection of visual information (e.g. Posner, 1980; Wolfe, Cave, & Franzel, 1989). These ideas are based on the fact that it is possible to facilitate processing of information at a location without making an eye movement towards it (Posner, 1980; Posner, Nissen, & Ogden, 1978). This phenomenon, *covert attention*, was first demonstrated using the now classic Posner cueing paradigm. In this task observers fixate at a central fixation point while one out of two boxes on the left and the right is cued. Subsequently, they have to respond to a target stimulus in either the same or the opposite box. Posner showed that when the target appears in the same box as the cue, the speed and accuracy of visual processing are improved relative to a target in the non-cued box. Despite the fact that in most situations these shifts of attention are accompanied by an eye movement (*overt attention*), attentional orienting was assumed to operate independently from the eye movement system (Hoffman & Nelson, 1981; Hunt & Kingstone, 2003; Klein & Pontefract, 1994; Posner, 1980).

The allocation of a covert attention while maintaining fixation has been the preferred way to study visual selection. The fact that the retinal locations of stimuli remain the same throughout an experiment immensely simplifies a lot of psychophysical investigations; eye movements only interfere with information processing, induce interruptions of viewing, and cause stimuli to fall on different parts of the retina all the time. However, as anyone who ever participated in a cognitive psychology experiment can acknowledge, attending somewhere while holding the eyes “still”, feels a bit awkward. The processing benefits associated with covertly attending to a location are minimal compared to the increase in visual acuity that can be accomplished by simply moving the eyes towards that part of the visual scene, so it feels rather unnatural not to utilize the enhanced resolution of foveal vision.<sup>1</sup>

The fact that we are able to attend somewhere while holding the eyes still does not necessarily mean that this process is completely independent from the eye movement system. Many studies have demonstrated that there is a close and obligatory coupling between covert and overt attention. For example, just before an eye movement is executed covert attention shifts towards the endpoint of the upcoming saccade (Born, Ansorge, & Kerzel, 2012; Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Kowler, Anderson, Doshier, & Blaser, 1995; Shepherd, Findlay, & Hockey, 1986). Interestingly, these effects are only observed when an eye movement is actually executed. When eye movements are cancelled at the last moment, no facilitation is found whatsoever (Born, Mottet, & Kerzel, 2014). Apparently, the mere preparation of an eye movement is not sufficient to elicit a shift of attention. The coupling also works the other way around; when an observer is instructed to covertly shift attention, this results in the obligatory preparation of an eye movement towards the attended location (Belopolsky & Theeuwes, 2009a, 2012). The close interdependence between overt and covert attention has raised questions about the underlying mechanisms that control these processes. How closely are saccades and attention linked? Are they part of

---

<sup>1</sup> One might wonder why we can perform covert shifts at all, if they are awkward and unnatural, and wholly unnecessary. They may, however, serve a purpose in some scenarios. For example, it might be beneficial to monitor someone else’s actions without letting them know you are observing them (Laidlaw, Rothwell, & Kingstone, 2016). However, in daily life we hardly ever encounter a situation in which it would be beneficial to scan our environment using our ‘internal spotlight’ instead of our eyes.



one and the same mechanism, or are they two independent but closely coupled processes?

Some researchers propose that covert attention is not an independent operating scanning mechanism, but instead an intrinsic part of action planning. One of the most influential theories on visual attention, the *premotor theory*, posits that covert attention is in fact a consequence of the preparation of an eye movement. It can be seen as a necessary by-product of saccade planning, and we can only measure its effects in isolation when we deliberately hold back the actual motor act (Klein, 1980; Rizzolatti, Riggio, Dascola, & Umiltá, 1987). In line with this, it has been argued that the presaccadic shift of attention supplements the oculomotor act by enhancing the extraction of information about the target of the upcoming saccade. Because saccades are usually sloppy, it can be beneficial to store some information across the eye movement, a mechanism called transsaccadic memory (Currie, McConkie, Carlson-Radvansky, & Irwin, 2000; Irwin & Gordon, 1998). In case the eyes do not land where they were supposed to go, this information can be used to identify the correct object (Hollingworth, 2007; Irwin & Zelinsky, 2002; McConkie & Currie, 1996). In a similar vein, the Visual Attention Model (VAM; Deubel & Schneider, 1996; Schneider, 1995) argues that both covert attention and motor preparation are the consequence of the same selection mechanism. This selection mechanism prioritizes a certain object, which leads to both enhancement of the processing of its features (selection-for-perception) and the preparation of a motor action towards it, such as a saccade or a grasping movement (selection for action). Both theories share the idea that covert attention and action preparation are tightly coupled. However, according to the premotor theory, attention is a consequence of action preparation, while VAM assumes that it is the other way around; saccade programming is a consequence of covert attention instead of its precursor.

In agreement with these ideas, several lines of evidence show that mechanisms responsible for saccade programming play a crucial role in the allocation of covert attention. First, it has been established that covertly attending a location influences the trajectories of saccadic eye movements (see Figure 2.2 for an example). When participants attend to a location and subsequently make an eye movement to another location, the trajectory of this eye movement is slightly altered (Sheliga, Riggio, & Rizzolatti, 1995). Depending on how quickly a saccade is executed, the eyes either curve toward or around this location (McSorley, Haggard, & Walker, 2006). This *saccade curvature* is thought to originate from competition between different saccade targets in the intermediate layers of the

superior colliculus (SC; McPeck, Han, & Keller, 2003; or subsequent adaptation in the brainstem, see Kruijne, Van der Stigchel, & Meeter, 2014). The SC contains a spatial motor map for the generation of eye movements and is the last stage of oculomotor preparation before a motor command is passed on to the saccade generators in the brain stem. The fact that covert attention is represented at this late stage of oculomotor preparation provides strong evidence for a common mechanism for attention and eye movements.

Second, the same brain areas appear to be involved in both attention and eye movement preparation. Research in monkeys has demonstrated that injecting a small electrical current into eye movement related brain structures like the Frontal Eye Fields (FEF) or the SC evokes saccades to the corresponding retinotopic locations. More interesting is what happens when the strength of this stimulation is tuned down to a level that is no longer strong enough to evoke an eye movement. This *subthreshold stimulation* exactly mimics the effects of covert attention. Although the eyes remain at the same location, visual processing is enhanced specifically at the location represented by the stimulation site (FEF; Armstrong, Fitzgerald, & Moore, 2006; Moore & Fallah, 2001, 2004; SC; Cavanaugh & Wurtz, 2004; Müller, Philiastides, & Newsome, 2005). Furthermore, this microstimulation also enhances the sensitivity of neurons in V4 in a way similar to endogenously attending to that location (Armstrong et al., 2006). In humans, similar sensitivity modulations in extrastriate areas have been observed when applying TMS to the FEF (Silvanto, Lavie, & Walsh, 2006). In addition, functional imaging studies show that the same network of frontal and parietal cortical areas is activated during both the preparation of an eye movement and covertly attending to a location (Beauchamp, Petit, Ellmore, Ingeholm, & Haxby, 2001; Corbetta et al., 1998; De Haan, Morgan, & Rorden, 2008; Nobre, Gitelman, Dias, & Mesulam, 2000; Perry & Zeki, 2000).

Although these data are often interpreted as evidence in favor of a functional coupling between eye movement preparation and covert attention, Smith and Schenk (2012) have pointed out that this interpretation might be problematic. Their main argument is that the FEF contains several overlapping but independent neural populations (Sato & Schall, 2003; Thompson, Bichot, & Schall, 1997; Thompson, Biscoe, & Sato, 2005). The FEF neurons responsible for saccadic eye movements are separate from the neurons involved in visual selection. This means that some neurons are involved in the preparation of an eye movement, while others have only visual properties. Microstimulation does not activate single neurons but activates all neural tissue surrounding the electrode tip (Tehovnik,

1996). They argue that based on the current neurophysiological evidence, it is impossible to determine whether the perceptual and neuromodulatory effects of microstimulation are driven by visual or motor signals. The same problem occurs when interpreting the neuroimaging evidence. The coarse spatial resolution of fMRI does not allow us to discern between different neural populations in the same area.

A crucial question is whether saccade generation and covert attention can operate independently. To address this question, a number of studies have focused on patients suffering from different forms of ophthalmoplegia (Craighero, Carta, & Fadiga, 2001; Gabay, Henik, & Gradstein, 2010; Rafal, Posner, Friedman, Inhoff, & Bernstein, 1988; Smith, Rorden, & Jackson, 2004). This disorder is characterized by a weakness or paralysis of one or more extraocular muscles, which leads to the inability to execute eye movements to certain locations in the visual field. Patients were asked to do the classic cueing task introduced by Posner (1980). This task usually shows that cueing a peripheral location leads to enhanced processing of targets shown at that location. However, patients with ophthalmoplegia showed no cueing effects for the affected locations, suggesting that they were unable to shift covert attention (Craighero et al., 2001; Gabay et al., 2010; Rafal et al., 1988; Smith et al., 2004).

Craighero and colleagues (Craighero, Nascimben, & Fadiga, 2004) developed a method to impair the ability to execute eye movements in healthy participants. In their *eye abduction* paradigm, participants were asked to fixate a location on a screen that was placed to the right of their sagittal plane, forcing them to rotate the eye into the temporal hemifield. Placing the eye at this extreme position prevented them from making eye movements further into the temporal hemifield. While this manipulation did not affect visual acuity, it was no longer possible to execute an eye movement to one side of the visual field. Similar to the ophthalmologic patients, cueing effects were abolished at locations to which no eye movement could be executed, an effect that was replicated in a variety of cueing and visual search tasks (Michalczyk, Paszulewicz, Bielas, & Wolski, 2018; Smith, Ball, & Ellison, 2014; Smith, Ball, Ellison, & Schenk, 2010; Smith, Schenk, & Rorden, 2012). The inability to make an eye movement to a location led to a corresponding deficit in covert attention. A fully functioning eye movement system appears to be necessary in order to shift attention.

Although these findings are consistent with the premotor theory, it is unknown whether the manipulation really disrupted oculomotor preparation. The degree of attentional allocation is inferred based on the cueing effect (the difference

between the reaction time on valid versus invalid cued trials), which has been shown to be influenced by both preceding trials and strategic top-down modulation (Macaluso & Doricchi, 2013). To establish a direct link between covert attention and eye movement preparation, it is crucial to determine whether the attentional deficits observed in eye abduction tasks stemmed from the inability to prepare an action towards it. In a recent study (Chapter 2) we used saccadic curvature to examine whether eye abduction eliminates the target-distractor competition in the oculomotor system. The inability to move the eyes to a location was shown to affect the amount of activity visual distractors evoke in the oculomotor system. Activity was reduced for distractors located beyond the oculomotor range, but was not abolished completely. This can be explained by the fact that other forms of motor planning, such as head movements, are also represented in the oculomotor system. Although in most experiments head movements are artificially restricted, normal behavior usually involves combined head eye-movements. There is a vast body of evidence demonstrating that the SC not only codes for eye movements but is also involved in the control of other forms of motor behavior such as head and arm movements (Freedman, Stanford, & Sparks, 1996; Guitton, Crommelinck, & Roucoux, 1980; Roucoux, Guitton, & Crommelinck, 1980; Stuphorn, Bauswein, & Hoffmann, 2000). For example, electrical stimulation of the monkey SC elicits combined eye head movements (Freedman et al., 1996; Segraves & Goldberg, 1992). Moreover, these forms of orienting behavior seem to be highly interchangeable, as was revealed by experiments with patients that can no longer make eye movements due to acute ophthalmoplegia. Directly after the onset of the condition eye movements were replaced by head movements showing all the characteristics of normal saccadic eye movements (Gaymard et al., 2000). In addition to the link between attention and eye movements, there also appears to be a strong relationship between attention and head movements preparation. Covert orienting leads to elevated electromyographic (EMG) activity from the neck muscles (Corneil, Munoz, Chapman, Admans, & Cushing, 2008), while the planning of a head movement results in attention shifting in the same direction just before the movement commences (Cicchini, Valsecchi, & De'Sperati, 2008), similar to the presaccadic shifts of attention demonstrated earlier (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Kowler et al., 1995). Instead of driving isolated eye movements, the oculomotor system might be better described as a gaze motor system in which eye and head movements are exchangeable.

The premotor theory posits that covert attention is an emergent property of movement planning, which is not per se limited to the eye movement system.

Locations can be represented in any effector system, depending on the task or goal of the observer (Rizzolatti & Craighero, 1998; Rizzolatti, Riggio, & Sheliga, 1994). In line with this, several parietal regions are involved in the transformation of sensory input into maps that code for action intention of different effector systems. For example, the anterior intraparietal area (AIP) responds in a similar way as LIP, but instead of coding potential eye movement goals it appears to be involved in the preparation grasping actions (Murata, Gallese, Luppino, Kaseda, & Sakata, 2000). The ventral intraparietal area (VIP; Duhamel, Colby, Goldberg, 1998) and medial intraparietal area (MIP; Ferraina et al., 1997) display similar properties. Just like LIP, these areas project onto corresponding premotor areas in the frontal cortex (Luppino, Murata, Govoni, & Matelli, 1999; Matelli, Govoni, Galletti, Kutz, & Luppino, 1998), thereby possibly playing a role in the transformation of visual information into action goals.

Although the oculomotor system is definitely not unique, it is likely to have a privileged role in the representation of relevant locations and the allocation of covert attention. First, its retinotopic nature allows for a straightforward translation from perceptual information into action codes. Second, in everyday life most actions are preceded by an eye movement. Try grabbing your cup of coffee standing next to you, and you will notice that even before your hands start moving your eyes have probably already fixated it (Land & Hayhoe, 2001). In line with this, it has been suggested that our eye movements work as *deictic pointers* (Ballard, Hayhoe, Pook, & Rao, 1997). By indexing the object that has to be acted upon, this strategy frees up resources that would otherwise be necessary to hold things in mind. In this way, perception and motor control can be vastly simplified.

To summarize, visual selection is an active process. Our eye movements supply us with a mechanism that allows efficient filtering of the most relevant information from the surroundings. On basis of this information we select goals of the next eye movement. The studies discussed above are part of a growing body of evidence suggesting covert attention to peripheral locations might act to supplement, not substitute for, actual movements of the eyes. In line with the premotor theory, attention might be an unavoidable consequence of movement planning, using the same neural substrates as eye movement planning, but with the actual motor response withheld.

### **An 'enactive' view of perception**

The previous section discussed the role of the (eye) movement system as an integral part of perception and attention. The idea that action is not just the end product of cognitive operations, but instead an intrinsic part of perception itself

was most explicitly developed in robotics, but has also had its impact on cognitive psychology and neuroscience. Inspired by the ecological approach of Gibson (1979), the 90s saw a number of theories emphasizing the key role of actions in understanding the world around us. Although they differ in details, they all share the idea that perceptual processes and action are closely intertwined. For example, Varela and colleagues (Varela, Rosch, & Thompson, 1992) argued that the function of perception is not to construct a veridical representation of the world, but to determine possible actions.

This 'enactive' view of cognition is also reflected in the sensorimotor contingency theory put forward by O'Regan and Noë (O'Regan, & Noë, 2001). This influential theory again emphasizes that there is no need for perception to construct a picture-in-the-head. It stresses that the outside world simply is out there, serving as its own, external representation. Within this view, "vision is a mode of exploration of the world that is mediated by knowledge of what we call sensorimotor contingencies." In other words, cognitive capacities are developed by learning the regularities of how an action, for example, an eye movement, influences the way we perceive things. Our knowledge about an object's shape, for example, is formed by our experience of how our eye movements change the image of this object on the retina.

A related idea is that of the ideomotor theory, which can be traced back all the way to the nineteenth century (Harleß, 1861; James, 1890; Lotze, 1852). Just like the sensorimotor contingency theory, it posits that people learn to associate movements with their outcomes by actively exploring their environment. However, in contrast to pure enactivism, this theory does assume some form of internal representation by claiming that stimuli and responses share the same representational resources or structures (Hommel, Musseler, Aschersleben, & Prinz, 2001; Prinz, 1997). In other words, actions are represented in terms of their perceptual consequences and vice versa.

In summary, although the representationalist idea of a picture-in-the head remains persistent across the vision sciences, the previous decades have marked a gradual transition into a more action-oriented paradigm. The allocation of covert attention can be regarded as an intrinsic part of oculomotor selection, and perception and cognition can be best understood as part of our ongoing interaction with the world around us.

### The relationship between spatial working memory and the oculomotor system

The previous section discussed the prominent role of the eye movement system in vision. We do not seem to construct a detailed representation of the environment; change blindness demonstrated that most information is discarded (at least from awareness) upon every new fixation. Instead we use eye movements to select the most relevant 'snapshots' from the external representation which is the world around us. To be able to navigate through our dynamic environment we do need to keep track of the locations of the most relevant objects. We need to determine where to move the eyes next, which locations remain relevant while moving our eyes elsewhere, and refrain from inspecting the same location over and over again. In this section, I will discuss the role of the oculomotor system in storage of spatial information in working memory.

Working memory is the mechanism that helps us maintain and manipulate visual information in the absence of sensory input. Working memory was traditionally regarded as a separate, higher order cognitive function, consisting of several specialized buffers for the storage and manipulation of information (Baddeley & Hitch, 1974). Neurophysiological recordings in monkeys demonstrated that neurons in the prefrontal cortex (PFC) show delay-period activity during working memory tasks (Funahashi, Bruce, & Goldman-Rakic, 1990; Funahashi, Bruce, & Goldman-Rakic, 1989; Goldman-Rakic, 1991), which led to the characterization of this area as the critical neural substrate for maintaining information (Goldman-Rakic, 1995). This interpretation was further backed by neuropsychological evidence showing that lesions of the PFC resulted in disrupted working memory performance (e.g. Gross, 1963; Jacobsen, 1936; Milner & Petrides, 1984; Pierrot-Deseilligny, Rivaud, Gaymard, & Agid, 1991; Pierrot-Deseilligny, Müri, Nyffeler, & Milea, 2005). However, recent years saw a gradual turn towards the idea that working memory might arise through the coordinated recruitment of sensory and motor processing systems (Belopolsky, 2015; Christophel, Klink, Spitzer, Roelfsema, & Haynes, 2017; Postle, 2006; Theeuwes, Olivers, & Chizk, 2005; Zimmer, 2008). Instead of consisting of specialized storage buffers, working memory might be better described as persistent activity in any of the brain areas involved in translation of sensory information into the most suitable response. Depending on the precise requirements of a task, information might be represented at any stage of this sensorimotor continuum (Christophel et al., 2017; Zimmer, 2008). Detailed sensory information might be represented in the sensory areas, while the prefrontal cortex may play a role in the maintenance of more

abstract information (Freedman, Riesenhuber, Poggio, & Miller, 2001). The most straightforward way to represent spatial information is using a prospective motor code. In other words, we memorize the location of an object by maintaining the motor plan that will eventually be used to act upon it. The retinotopic layout of the oculomotor areas makes it ideally suited for the maintenance of spatial information (Theeuwes, Belopolsky, & Olivers, 2009). In line with the premotor theory of attention discussed above, endogenous maintenance of spatial information might rely on the same areas responsible for the planning of eye movements.

Different lines of evidence have linked motor preparation to spatial working memory. First, there is the observation that the execution of voluntary eye movements disrupts the ability to maintain a set of locations in memory (Baddeley, 1986; Postle, Idzikowski, Sala, Logie, & Baddeley, 2006; Smyth & Scholey, 1994). Although such disruptive effects have also been observed during movements of the arms (Lawrence, Myerson, Oonk, & Abrams, 2001; Logie & Marchetti, 1991) or fingers (Farmer, Berman, & Fletcher, 1986; Salway & Logie, 1995), the reduction in performance is significantly larger during movements of the eyes (Pearson & Sahraie, 2003). Further evidence for a role of the eye movement system in the maintenance spatial information was provided in a study by Tremblay and colleagues (S. Tremblay, Saint-Aubin, & Jalbert, 2006), who studied eye movement patterns during the maintenance of a series of locations. They concluded that eye movements are used as a form of rehearsal, helping an observer to maintain spatial locations in memory.

Second, there is neurophysiological evidence linking neural activity in the oculomotor system to spatial working memory. A paradigm that is often used to investigate spatial working memory functions in both primates and humans involves memory-guided saccades. In these tasks, a participant has to look at a fixation point while remembering the location of a flashed target somewhere in the periphery. After a certain delay, a 'go signal' indicates that an eye movement to the remembered target has to be executed. A series of neurophysiological studies in monkeys demonstrated that neurons in regions known to be involved in preparation of eye movements, such as LIP (Duhamel, Colby, & Goldberg, 1992) and FEF (Sommer & Wurtz, 2001; Umeno & Goldberg, 2001), show sustained activity when maintaining the location of a delayed eye movement. Moreover, this sustained activity was also observed when the task did not involve a delayed eye movement, but merely remembering the location and identity of a stimulus in a change detection task (Armstrong, Chang, & Moore, 2009).



In humans, several studies have demonstrated that the accuracy of memory guided saccades is impaired in patients with lesions affecting the PPC or the FEF (Mackey, Devinsky, Doyle, Golfinos, & Curtis, 2016; Pierrot-Deseilligny, Rivaud, Gaymard, & Agid, 1991; Ploner, Rivaud-Péchéoux, Gaymard, Agid, & Pierrot-Deseilligny, 1999; Rivaud, Müri, Gaymard, Vermersch, & Pierrot-Deseilligny, 1994). In addition, imaging studies show that memorizing the location of a stimulus induces sustained BOLD activity in frontal and posterior parietal regions (Corbetta et al., 1998; Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998; Curtis, Rao, & D'Esposito, 2004; Kastner et al., 2007; Todd & Marois, 2004), which decays as soon as a location is no longer relevant (Schluppeck, Curtis, Glimcher, & Heeger, 2006; Srimal & Curtis, 2008). A number of studies have directly linked the strength of this activity to working memory performance. Notably, the strength of delay period activity predicts how well spatial information is memorized; people that show stronger sustained activity perform better on a spatial memory task (Curtis et al., 2004). In addition, disturbing this sustained response by applying transcranial magnetic stimulation to the FEF leads to deficits in memory retrieval (Campana, Cowey, Casco, Oudsen, & Walsh, 2007).

### **The representation of priority in the oculomotor system**

Due to the limited spatial resolution of neuroimaging techniques, it is hard to determine whether delay-period activity is spatially specific. However, the development of multivariate decoding techniques now enables the identification of content-specific activity in humans. Convincing evidence for the existence of a common mechanism for visual attention, working memory and eye movements was provided in an elegant study by Jerde and Curtis (Jerde & Curtis, 2013). They used these techniques to compare neural representations of spatial information in three different tasks: maintaining a location in working memory, covertly attending to a location, and maintaining a saccade plan. In all three tasks, a location could subsequently be predicted on basis of the pattern of BOLD activity. Most interestingly, in both the FEF and the IPS (the human homologue of monkey LIP) classifiers trained to predict a location in one task cross-predicted locations in the other spatial tasks. For example, a classifier trained to discriminate the locus of covert attention could also be used to reliably predict a location maintained in working memory or the target of an upcoming eye movement. Retinotopic oculomotor maps appear to be agnostic with regard to the specific task shaping its activity pattern.

The patterns of neural activity in the oculomotor-related areas appear to be indistinguishable across spatial tasks. An influential idea is that brain areas

involved in oculomotor control, like the IPS, FEF, and the SC, serve as priority maps (Awh, Belopolsky, & Theeuwes, 2012; Belopolsky, 2015; Bisley & Goldberg, 2010; Fecteau & Munoz, 2006; Ipata, Gee, Bisley, & Goldberg, 2009; Serences & Yantis, 2006; Zelinsky & Bisley, 2015). These maps contain dynamic representations of space, in which the salience of objects in the environment and the goals of the observer are thought to be amalgamated into one priority landscape, in which activity tags the locations in the environment that are behaviourally the most relevant. These locations compete for selection, with the winner being selected as the goal of the next action.

Different oculomotor areas have previously been described as such priority maps (Fecteau & Munoz, 2006). Both area LIP (Bisley & Goldberg, 2010) and the FEF (Kastner et al., 2007; Thompson et al., 2005), have been reported to contain maps coding for the most likely action goals. In addition to these cortical areas, there is also evidence for the representation of prioritized space at the subcortical level. The motor-related intermediate layers of the SC (SCi), which receive input from frontal and parietal areas (White & Munoz, 2011), were recently shown to contain a dynamic representation of potential action goals (White et al., 2017).

In an elegant eye movement study by Theeuwes and colleagues (Theeuwes et al., 2005, see also Belopolsky & Theeuwes, 2011; Belopolsky & Van der Stigchel, 2013; Boon, Theeuwes, & Belopolsky, 2014) participants were asked to memorize the location of a stimulus and subsequently make an eye movements to a different location. Analysis of the trajectories of these eye movements revealed that eyes slightly curve away from the memorized location, in a similar way as when ignoring a salient distractor. This curvature is likely originating from the competition between different potential eye movement goals within the SCi (McPeck et al., 2003). The fact that both attended and memorized locations compete with saccade goals indicates that there is a strong overlap between the mechanisms involved in spatial attention, working memory, and the eye movement system.

Even though the selection of our next eye movement goal is a winner-take-all mechanism (we can only make one eye movement at a time), areas functioning as priority maps are probably not limited to representing a single activity peak. Instead, they seem to constitute a dynamic priority landscape, representing multiple important locations at the same time (Godijn & Theeuwes, 2002b; Sprague, Ester, & Serences, 2014). These can be the target of a future eye movement, a covertly attended stimulus, or simply the locations of a memorized object. Moreover, locations that have already been visited have to be inhibited in some way (Itti & Koch, 2001), a mechanism called 'inhibition of return' (Godijn &

Theeuwes, 2002a; Posner & Cohen, 1984). If not, we would constantly revisit the most salient locations over and over again. Priority maps have been shown to contain such a mechanism, reducing the activity of neurons representing already visited targets (Mirpour, Arcizet, Ong, & Bisley, 2009).

Further evidence for the role of eye movement preparation in spatial memory was provided by a study using the eye abduction paradigm. The inability to execute an eye movement to a memorized location led to a small decrease in the ability to report its exact location (Pearson, Ball, & Smith, 2014). These results are in line with the similar decrease in oculomotor competition observed when a visual distractor was presented outside of the oculomotor range (Boon, Theeuwes, & Belopolsky, 2017). The anatomical constraints of the eye movement system influence the ability to represent spatial information in the brain.

To conclude, the oculomotor system plays a crucial role not only in saccade preparation and the allocation of covert attention, but also in maintaining spatial information in working memory. It appears to represent the *priority* of objects in the environment, irrespective of the content determining their prioritized status. In line with the premotor theory of attention discussed in the previous section, memorizing these relevant locations might rely on the maintenance of a prospective motor code. In other words, spatial working memory might be nothing else than the sustained prioritization of a location as a potential (eye) movement goal.

### The updating of spatial memory across saccades

The eye movement system plays a crucial role in maintenance of spatial information. However, each eye movement also dramatically changes visual input, which poses a great challenge for keeping track of relevant information. Whether it is the planning of a future eye movement, memorizing a location of an object, or simply refraining from revisiting already inspected objects - any of these tasks would not be possible if this information would be computed anew with every fixation. In the remainder of this chapter I will discuss the mechanisms involved in updating endogenously maintained information across saccades and object movements. I will deliberate on how both extra-retinal and retinal signals can be used to update visual and memorized information, and argue that timing is a crucial factor in determining the relative contribution of these sources of information.

### Extra-retinal signals

Visual input is processed in a retinotopic (eye-centered) reference frame (Talbot & Marshall, 1941), which implies that with each eye movement the relevant objects are represented by a completely different set of neurons. Despite this continuously changing retinal input we experience the world as stable. This discrepancy has preoccupied the minds of researchers for several centuries. Why do we not notice anything of the dramatic shifts in visual input that are induced by eye movements?

One group of theories, dating back to Helmholtz (Helmholtz, 1867; Hering, 1868; Holst & Mittelstaedt, 1950) emphasized the role of so-called extra-retinal signals. These non-retinal signals allow the brain to differentiate between changes in sensory input that arise from the environment from those that arise as a consequence of self-generated actions. Extra-retinal information can be subdivided into afferent (proprioceptive) signals and efferent (motor) signals (Matin, 1972). The afferent signals provide feedback from the eye muscles regarding the rotation of the eye in the socket. Because it takes much more time before the brain can utilize the afferent signals, most research has focused on the role of efferent signals. When an eye movement is launched, a copy of the motor program (also known as “the efferent copy” or “corollary discharge”) is retained, which can later be used for correction or cancellation of the displacement of the retinal image induced by the eye movement. This would enable the brain to anticipate the consequences of eye movements and bridge the gap between successive fixations.

The cancellation theory was inspired by two fascinating observations. First, Purkinje discovered that gently pushing a finger on the side of the eyeball with the other eye closed results in an illusory tilting of the visual scene. The pushing slightly alters the retinal image, comparable to the shift of the retinal image induced by an eye movement. However, the fact that a motor command is absent supposedly results in the inability to compensate for this movement. More than a century later, both Von Holst (von Holst & Mittelstaedt, 1950) and Sperry (Sperry, 1950) came up with another method to decouple the motor command from the actual movement of the eye. Although they used different animal models (insects and fish respectively), both found a similar effect when they surgically inverted the eyeball of these animals. As long as it moved straight forward no abnormal behaviour was observed. However, as soon as the animal decided to make a turn to the right or the left this resulted in a rapid spinning in that direction. Independent of each other, both researchers concluded that instead of

compensating for the retinal displacement, the motor command now acted to exaggerate the effects of self-induced movement.

Cancellation theories are based on the idea that the brain contains a representation of visual surroundings which shifts with each eye movement. However, from an 'active vision' perspective, there is no need to update anything at all. As was already noticed by Gibson (Gibson, 1966), the only representation we need is already there, which is the outside world itself. This representation is already stable, so the problem of visual stability is in fact a non-issue. We can select all relevant bits of visual information from this 'external' representation by simply moving our eyes towards it. While the active vision perspective denies the necessity for a mechanism for updating rich internal representations, completely discarding any mechanism compensating for eye movements does neglect certain aspects of spatial cognition. As already demonstrated by Yarbus (1967), the eyes do not move around randomly, hoping to accidentally land on an interesting object. Instead, we efficiently select the most relevant locations to align the fovea. Crucially, these potential motor goals have to be updated across each eye movement. Instead of updating a complete picture of the surroundings extraretinal signals might serve to update only the spatial layout of objects relevant for future behaviour (Bays & Husain, 2007; Cavanagh, Hunt, Afraz, & Rolfs, 2010; Wurtz, 2008).

Convincing evidence in favour of a spatial compensation mechanism driven by extraretinal signals was provided by Mays and Sparks (1980). They used a task in which monkeys were trained to make an eye movement to a briefly flashed target somewhere in the visual periphery. However, just before the monkey moved its eyes they electrically stimulated the superior colliculus. This microstimulation induced an eye movement to another location. Despite the interference caused by this involuntary saccade, the animal was able to make an accurate eye movement to the remembered target directly afterwards. Even though the saccade target was no longer visible, its location was automatically updated to compensate for the displacement induced by the artificially evoked eye movement. This compensation mechanism is unlikely to be driven by afferent eye position signals; although similar results are obtained when electrically inducing eye movements in cortical oculomotor areas like the FEF (Schiller & Sandell, 1983; Tehovnik & Sommer 1996), this compensation mechanism does not work when saccades are evoked in motor structures downstream from the SC (Mays & Sparks, 1981).

Although apparently driven by extra-retinal signals, the neural mechanisms behind this this compensation mechanism were initially unclear. However, in their

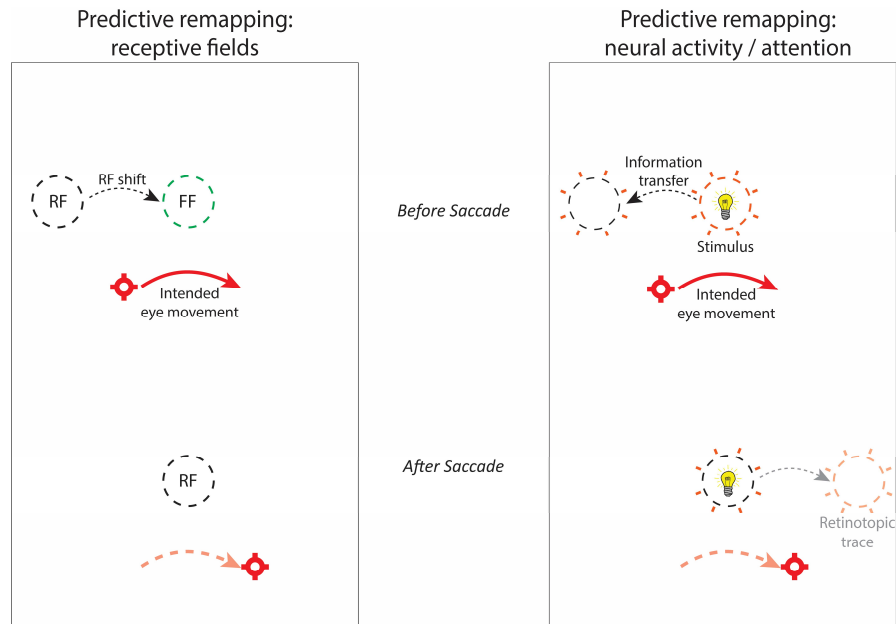
landmark study, Duhamel and colleagues (Duhamel et al., 1992) showed that already before an eye movement some neurons in area LIP of rhesus macaques begin to fire in response to stimuli that will occupy their receptive field after the saccade is completed. Consistent with cancellation theories described previously, these parietal cells seem to use a copy of the motor command to predict the sensory consequences of eye movements (see Figure 1.1, left panel). This “predictive remapping” of neural activity has subsequently also been discovered in two other brain regions associated with eye movement control: the FEF (Umeno & Goldberg, 1997, 2001) and the SC (Dunn, Hall, & Colby, 2010; M. F. Walker, Fitzgibbon, & Goldberg, 1995).

In addition to motor areas, remapping of visual activity has been observed in the early visual areas (Nakamura & Colby, 2002). However, the earlier in the visual pathway, the smaller the proportion of neurons displaying remapping properties becomes, and the later remapping happens relative to the saccade onset. This is in line with the general idea that the efference copy driving this transfer of neural activity is generated in the SC. Several experiments show that these signals are likely to be projected to the FEF, LIP, and eventually visual cortex through the colliculo-thalamic pathway. For example, blocking this pathway by inactivation of the mediodorsal nucleus of the thalamus, stops most FEF cells from showing any remapping effects (Sommer & Wurtz, 2006). In humans, thalamic and frontoparietal lesions have been shown to lead to similar deficits in the updating of spatial content. When these patients are asked to make a sequence of two saccades, they are unable to update the location of the second saccade after the first one, which resulted in a systematic inaccuracy of the second eye movements (Bellebaum, Daum, Koch, Schwarz, & Hoffmann, 2005; Ostendorf, Liebermann, & Ploner, 2010).

Although remapping is difficult to study using fMRI, several studies do indicate that saccade related areas in the parietal and frontal cortex have an important function in the updating of spatial content across saccades. For example, Heide and colleagues (Heide et al., 2001) showed that the execution of saccade sequences to remembered targets activates a network of areas including the FEF and the intraparietal sulcus (IPS). Others demonstrated actual cross-hemispheric transfer of activity in parietal areas (Medendorp, Goltz, Vilis, & Crawford, 2003; Merriam, Genovese, & Colby, 2003). In one of these studies participants were instructed to make a horizontal eye movement from one side of the screen to the other. Just before this eye movement was executed, a visual stimulus was briefly presented at a central location, eliciting activity in contralateral parietal areas. After the eye

movement brought the location of the stimulus into the opposite visual hemifield, this neural activity was being remapped to the other hemisphere (Merriam et al., 2003). Similar results have been obtained in a study which involved execution of a sequence of two saccades. Just like visual information, the second saccade goal was remapped across hemispheres when the first saccade brought it into the opposing hemifield (Medendorp et al., 2003). Using a comparable task in combination with EEG, Bellebaum and colleagues (Bellebaum & Daum, 2006; Bellebaum, Hoffmann, & Daum, 2005) found that ERP components originating from the parietal cortex were also remapped after an intervening saccade. In line with this, both damage to the parietal lobe (Sapir, Hayes, Henik, Danziger, & Rafal, 2004) and applying TMS to this area (Collins & Jacquet, 2017; Rushworth & Taylor, 2006; van Koningsbruggen, Gabay, Sapir, Henik, & Rafal, 2009) can lead to deficits in the ability to update locations across saccades.

A number of recent studies have characterized this remapping mechanism at a behavioral level, demonstrating fast, or even anticipatory shifts of attentional facilitation around the time of a saccade (Jonikaitis, Szinte, Rolfs, & Cavanagh, 2013; Rolfs, Jonikaitis, Deubel, & Cavanagh, 2011; Szinte, Carrasco, Cavanagh, & Rolfs, 2015). In one study (Jonikaitis et al., 2013), an irrelevant, but salient cue was flashed while participants were planning a saccade. It is known that such a cue captures attention (Posner & Cohen, 1984), resulting in a short-lived processing benefit at its location. Instead of measuring neural activity, a visual probe was used to determine how these processing benefits were updated across an eye movement. The results showed that attention was maintained at the spatiotopic location of the cue before and after the saccade, despite a change in the retinal location of the cue induced by this saccade. Facilitation at the retinotopic location of the cue, however, decayed very rapidly after the eye landed. Importantly, right before an impending saccade, attention was already allocated to the future retinotopic location of the cue, demonstrating presaccadic remapping of visual attention (Figure 1.1, right panel).



**Figure 1.1. Example of remapping.** Left panel: Presaccadic shift of receptive fields. Just before an eye movement is executed a neuron's receptive field (RF) shifts in the direction of the upcoming saccade, to the location of its future receptive field (FF). Right panel: Remapping of neural activity or covert attention. A stimulus presented just before an eye movement elicits a corresponding neural response or shift of covert attention. This is predictively remapped in the opposite direction of the impending saccade. After the eye movement this activity/ attention will be aligned with the world-centered location of the stimulus.

Recently, Jonikaitis & Belopolsky (Jonikaitis & Belopolsky, 2014) demonstrated that such spatiotopic representations also emerge rapidly in the oculomotor system of human observers. They used saccadic curvature to examine whether the competition between target and distractors across eye movements occurs in retinotopic or spatiotopic coordinates. In this study, participants performed a sequence consisting of a horizontal and a vertical saccade. Oculomotor competition was induced by briefly presenting a task-irrelevant distractor at different times during the sequence. Despite the intervening saccade, the second saccade curved away from the spatiotopic location of the distractor that was presented before the first saccade. Furthermore, saccade curvature away increased with an increase in salience of the distractor that was presented before the first saccade. The results clearly showed that not only the information about distractor's spatial location, but also the information about its relative salience was transferred across saccades. The oculomotor system appears to contain the mechanisms for a rapid and automatic updating of spatial information.



### Retinal signals

Even though there is quite some empirical evidence for the existence of an anticipatory compensation mechanism driven by extraretinal signals, several lines of evidence suggest that these signals are often sacrificed to preserve perception of a stable visual world. Most prominently, compensation theories fail to explain the finding that observers tend to miss object displacements if they occur during a saccade (e.g., Bray et al., 2016; Bridgeman et al., 1975; Collins et al., 2009; Deubel et al., 1996; Li & Matin, 1990). When making an eye movement to a target stimulus, this target can be displaced over a distance up to one third of saccade amplitude without being noticed (Bridgeman et al., 1975). Although these ‘perisaccadic’ object movements induce a large discrepancy between the predicted and actual postsaccadic location of an object, this mismatch is often ignored. Information about the location of the target is not lost though. When the target is not continuously present before and after the saccade but reappears at a displaced location after a 50 ms blank interval after the saccade, displacements are detected with a high accuracy (Deubel, Bridgeman, & Schneider, 1998; Deubel et al., 1996; Deubel, Schneider, & Bridgeman, 2002). This “target blanking” paradigm shows that there is a highly accurate prediction about the target location based on the extraretinal signal, but that this information is nevertheless often discarded.

One explanation for these findings is that the visual system relies on the null hypothesis that the visual world is stable across saccades. According to this idea there is no need for a precise prediction of the postsaccadic location of an object. As long as there is a rough match between pre- and postsaccadic stimulus location and identity, objects can often easily be relocalized on basis of remembered identity information (Currie et al., 2000; Irwin & Gordon, 1998; Irwin & Zelinsky, 2002; McConkie & Currie, 1996). For example, when an eye movement accidentally does not land on its target, memorized features of this target are used to drive fast corrective saccades (Hollingworth, Richard, & Luck, 2008). Moreover, several studies demonstrated that an object’s position is encoded according to its spatial relationship to other stimuli in the environment. The detection of object displacement relies to a large extent on changes in position relative to other objects (Deubel, 2004; Deubel, Koch, & Bridgeman, 2010; Germeys, Graef, Panis, Eccelpoel, & Verfaillie, 2004; Verfaillie, 1997; Verfaillie & De Graef, 2000). Objects that are continuously present across eye movements act as landmarks. As described by the sensorimotor contingency idea postulated by O’Regan and Noë (O’Regan & Noë, 2001), we have certain expectancies about the way eye movements change our perception of an object. These expectancies are based on

the sensorimotor regularities learned through experience. One such a rule is that objects do not tend to change location during a saccadic eye movement. This might be the reason why if an object is displaced during a saccade, the brain attributes the mismatch with expected input to some form of oculomotor error, instead of object displacement. In such a case a second, stationary but briefly blanked object will be perceived as moving into the opposite direction (Currie et al., 2000; Deubel, 2004; Deubel et al., 2010; Ostendorf & Dolan, 2015). The displaced object will serve as a landmark for the blanked object, and the change in spatial relational information is always attributed to the object that was not continuously present before and after the saccade.

The question is what source of information the brain uses to maintain spatial constancy across saccades. Both extraretinal and retinal mechanisms provide solutions that are far from perfect. Although extraretinal signals allow fast, even predictive, updating of stimulus locations, our eye movements are often not very accurate. As was already observed by MacKay (1973), this would introduce errors between the predicted and actual post-saccadic location of an object. In addition, extraretinal signals appear to underestimate the eccentricity of eye position. Bridgeman and Stark (Bridgeman & Stark, 1991) cleverly used the eye poking illusion discovered by Purkinje to measure the relative contribution of different extraretinal signals in judging the position of a point of light in an otherwise dark scene. Participants were instructed look at this point with one eye covered. Subsequently, they either poked either their occluded or their viewing eye. Both have a different effect on the proprioceptive signal and efference copy, which allowed them to estimate the contribution of both signals. They showed that both efference copy and proprioception of the eye muscles are used in calibrating the position of the eye. However, even if added up, the gain of the extraretinal compensation was only 7% of the total size of eye displacements.

Retinal information, on the other hand, is considerably more reliable. Unfortunately, it has one big disadvantage; the acquisition of postsaccadic retinal input requires a considerable amount of time (Goodale, Króliczak, & Westwood, 2005). A recent study from our lab showed that it takes approximately 180 ms before new spatial information is integrated (van Leeuwen & Belopolsky, Submitted). Given the speed and frequency of our eye movements it is unlikely that we rely solely on this information. As usually both retinal and extraretinal signals are available, an optimal system should combine both sources to make the most reliable estimation of the postsaccadic location of an object. Different studies have shown that the relative contribution of extraretinal and retinal signals

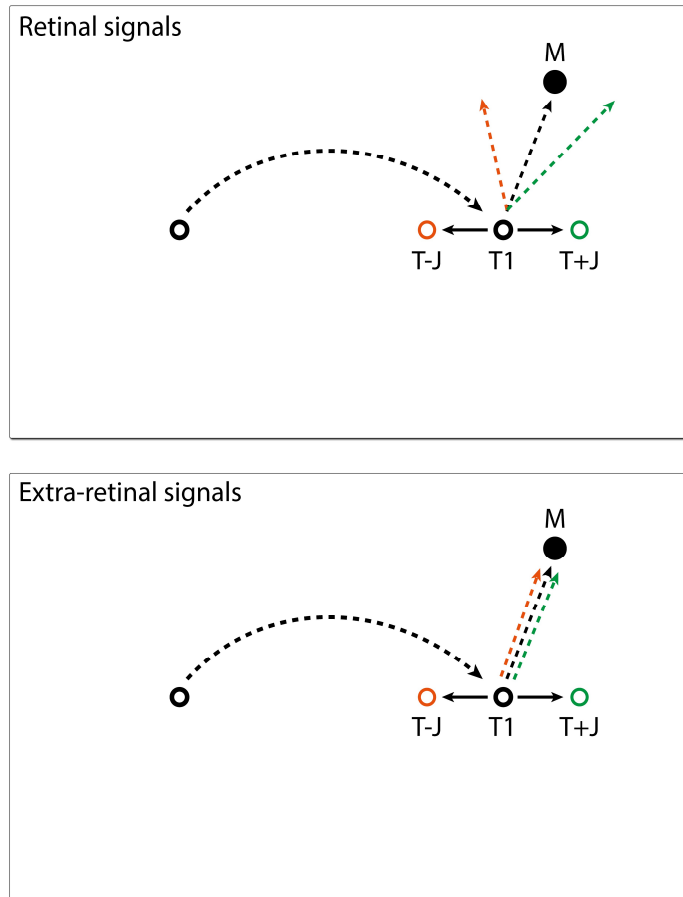
depends on the reliability of both signals. Where people normally rely heavily on landmark objects when they are available, a recent study showed that much more weight is put on extraretinal signals when these landmark objects are not completely stable, but vibrating (Byrne & Crawford, 2010). Niemeier and colleagues (Niemeier, Crawford, & Tweed, 2003) showed that the relative contribution of extraretinal signals correlates with the precision of oculomotor control; people that display more oculomotor error rely more heavily on retinal cues. Apparently the visual system optimally integrates sensory and motor information, resulting in the most reliable estimate of the postsaccadic locations of objects.

### **Updating of endogenously maintained information across saccades**

Most of the research described above involved updating of visual information. In some cases participants had to compare stimuli across saccades and indicate whether an object changed location or not. In other studies, a transient stimulus was used to elicit a neural response or corresponding shift of attention, which was subsequently updated across an eye movement. For the described remapping effects to occur it is crucial that this stimulus is presented in a short time window preceding the eye movement. However, in real world situations we often have to keep track of objects for multiple seconds, during which dozens of eye movements can be executed to other regions of interest. The situation becomes more complex if an object of interest is no longer visible, as a result of being, for example, occluded by another object. This means that the *internal, memorized* position of the object has to be updated. The question is whether sustained activity is updated in a similar fashion as exogenously triggered activity.

The implicit assumption that the world remains stable across saccades seems to influence spatial judgement. This assumption relies on object correspondence, in other words, the fact that the same object is visible before and after an eye movement. Note that when maintaining information in memory, direct object correspondence does not play a role, simply because a memorized object is no longer present. This might lead to different processes of updating spatial information in memory. Since the assumption of stability is never violated, updating might rely completely on extra-retinal signals, driving automatic remapping of the neural activity representing a remembered stimulus. Alternatively, the remembered location might be anchored to other objects in the visual scene, such as the saccade target. Chapter 3 of this thesis tackles precisely this question. In this chapter, the mechanism for updating the location of a remembered stimulus during eye movements was investigated. Participants were

memorizing the location of a stimulus and subsequently had to make an eye movement to a saccade target. Crucially, this target could be displaced during this eye movement. As described above, such displacements are typically not noticed by participants (Bridgeman et al., 1975; Deubel et al., 1996). It was assumed that if updating visual-spatial memory relies exclusively on retinal signals, the error in localizing the memorized location would be equal to the target displacement. As the only reference object (the saccade target) is perceived as stable across the saccade, the assumption of stability is not violated, and an 'assumption theory' (e.g. Deubel et al., 1996; McConkie & Currie, 1996) would predict the displacement to be attributed to imperfect oculomotor coordination (i.e. the oculomotor error). Alternatively, if memory relies exclusively on extraretinal signals, then the saccade target displacement should have no effect on localizing the memorized location after a saccade (see Figure 1.2). This approach allows the quantification of the relative contribution of extraretinal oculomotor signals and postsaccadic retinal signals in updating memorized locations across saccades.



**Figure 1.2. Experimental predictions in the memory task from Chapter 3.** Participants memorized a location (M) and made a saccade to a target (T). This target could remain at the same location, or be displaced in the direction of the saccade or the opposite direction (T+J, T-J respectively). Top panel: Localization of a memorized stimulus relies entirely on retinal signals (sensory feedback), the percentage of correction will be 100%. Bottom panel: Localization relies entirely on extra-retinal signals (efference copy), the percentage of correction will be 0%.

The results showed that displacement of the saccade target caused a small but systematic bias in the direction of this displacement, ranging between 20 and 40 % of total target displacement. This is similar to the effects found in an earlier study using a comparable paradigm (Munuera, Morel, Duhamel, & Deneve, 2009). The updating of working memory across saccades relies neither completely on extraretinal nor on retinal information. The relative contribution of retinal signals is not fixed but depends on both the time available to integrate this information as well as the distance between the saccade target and the remembered location. The more time elapses between end of the first saccade and the start of the

localization saccade, the more interference there is from the post-saccadic retinal input.

The combination of multiple sources of information in calibrating the postsaccadic location of a stimulus is in line with other studies demonstrating an optimal integration strategy when having to locate objects across eye movements (Atsma, Majj, Koppen, Irwin, & Medendorp, 2016; Munuera et al., 2009; Niemeier et al., 2003; Poletti, Burr, & Rucci, 2013; Vaziri, Diedrichsen, & Shadmehr, 2006; Ziesche & Hamker, 2011). Most of these studies emphasized the relative reliability of both sources of information, while a more crucial factor might be the timing between an intervening saccade and a subsequent orienting action. This can be explained by the fact that it takes time to process and integrate retinal information once the eyes have landed (Goodale et al., 2005; van Leeuwen & Belopolsky, Submitted). This is in line with recent findings of Atsma and colleagues (Atsma et al., 2016), showing that postsaccadic viewing time is an important factor when integrating pre- and postsaccadic spatial information. In their study, participants had to make an eye movement and subsequently indicate the initial, presaccadic, location of either the fixation object, the saccade target, or another object. However, similar to the experiments described in Chapter 3 of this thesis, these objects could be visible at a displaced location for a brief period after the saccade. As expected, localization was biased towards this post-saccadic, displaced location. What is more interesting is that the amount of time the objects was visible upon saccade landing had a major impact on the size of this localization bias. The longer the object was visible upon saccade landing, the larger its influence was on the perceived location. In line with the results described in Chapter 3, it appears to take some time to integrate post-saccadic input. In other words, localization depends mainly on the availability, not the reliability, of information.

#### **The time course of updating spatial working memory**

The results described in Chapter 3 show that we quickly have access to a precise representation of relevant objects after an intervening eye movement, supposedly based on a copy of the oculomotor command. Only later on the retinal information is taken into account. The question is how long it takes to form such a spatiotopic representation. Current models of spatial working memory updating during eye movements suggest that, in contrast to the rapid remapping of visual signals, the transformation of a memory representation into spatiotopic coordinates is a time-consuming and effortful process. Given a close link between attentional and working memory representations (Awh & Jonides, 1998; Belopolsky, Kramer and Godijn, 2008, but see Belopolsky & Theeuwes, 2009), attentional facilitation is

often used as an indication of updating of spatial working memory representations across saccades. In these studies, participants have to respond to a probe presented at different time intervals after an intervening saccade. When this probe is presented at the retinotopic location of the remembered object, there was attentional facilitation almost immediately after the saccade. For probes presented at the spatiotopic location, facilitation was only observed after some time had passed (Golomb, 2010; Golomb, Chun, & Mazer, 2008; Lisi, Cavanagh, & Zorzi, 2015). These findings led to the hypothesis that the native coordinate system of memory representations is retinotopic; it was proposed that, contrary to the automatic updating of visually evoked activity, memorized locations naturally move along with each eye movement and a special effort is necessary to gradually transform these retinotopic representations into the world-centered (spatiotopic) coordinates (Golomb et al., 2008; Golomb, Marino, Chun, & Mazer, 2011).

Several studies demonstrated that reporting the retinotopic, eye-centered location of a remembered stimulus is even easier than remembering its world-centered location (Golomb & Kanwisher, 2012; Shafer-Skelton & Golomb, 2017). Although these results are quite convincing, this seems rather counterintuitive. The world-centred reference frame, by definition, is the only relevant one when interacting with the world around us. In line with the sensorimotor contingency theory (O'Regan, & Noë, 2001), our brain relies on certain implicit assumption about the way eye movements change perception. One of these assumptions is that an object never occupies the same retinal location before and after a saccadic eye movement. Furthermore, memorized information is represented in the same brain areas that display remapping properties. It would be rather inefficient if updating of this activity involves a different, voluntary mechanism. Neurophysiological studies show that not only visually evoked activity, but also memorized stimuli are being remapped across saccades. Some FEF neurons fire before an eye movement brings a the location of a previously shown stimulus into their receptive field (Umeno & Goldberg, 2001). In addition, remembered saccade targets appear to be rapidly updated across both voluntary (Boon et al., 2016) and artificially induced saccades (Mays & Sparks 1980; Schiller & Sandell, 1983; Tehovnik & Sommer 1996).

Does updating of endogenously maintained information in the oculomotor system involve the same mechanism as updating of exogenous attention, or is this a time-consuming and effortful process? In chapter 4 of this thesis we investigated the time-course of updating memorized locations in the oculomotor system (Boon,

Zeni, Theeuwes, & Belopolsky, in press). To examine this, we modelled our paradigm after Jonikaitis and Belopolsky (Jonikaitis & Belopolsky, 2014). Participants made a double-step saccade, but instead of presenting a salient distractor, they were instructed to keep a location in memory. Previous studies have shown that this evokes persistent activity in the oculomotor map, causing the eyes to curve away from the remembered location (Belopolsky & Theeuwes, 2011; Boon et al., 2014; Theeuwes et al., 2005). The trajectory of the second saccade was used to assess whether the memorized location was represented in retinotopic or spatiotopic coordinates after the first eye movement. Crucially, we took advantage of the variability in the time interval between the two saccades to determine when the spatiotopic representation emerges. If memorized locations are rapidly updated, we expected the second saccade to curve away from the spatiotopic location even when the intersaccadic interval is short. However, if the formation of a spatiotopic representation is a slow and effortful process, saccades should only curve away from the spatiotopic location after the longer intersaccadic intervals.

The results described in Chapter 4 showed that locations of behaviorally relevant objects are rapidly updated across saccades. Within 130 ms after an intervening saccade, at the eyes curved away from the spatiotopic location of a remembered item. Although there likely was also residual retinotopic activity, this was masked by the dominant spatiotopic representation. The results contradict the view that postulates a gradual deliberate shift from retinotopic to spatiotopic memory representations. The rapid emergence of spatiotopic memory representations observed here is comparable to the updating of exogenous attentional signals (Jonikaitis et al., 2013; Rolfs et al., 2011; Szinte et al., 2015). In these tasks nothing had to be memorized, but instead an irrelevant but salient cue was used to capture attention. When this cue was flashed shortly before saccade execution, attention resided at its spatiotopic location directly afterwards. Furthermore, studies of saccadic IOR have reported similar rapidly emerging spatiotopic representations (He, Ding, & Wang, 2015; Hilchey, Klein, Satel, & Wang, 2012; Pertzov, Zohary, & Avidan, 2010, but see Mathôt & Theeuwes, 2010). The study that is most comparable to ours used saccade curvature to show that spatiotopic representations also emerge rapidly and automatically in the oculomotor system (Jonikaitis & Belopolsky, 2014). Saccades curved away from the spatiotopic location of an attended location after a similar interval as we show here for endogenously maintained locations, indicating that a single mechanism might be involved in the updating of relevant locations, independent of whether these locations are prioritized in an endogenous or exogenous manner.



The research described in Chapter 4 also showed that the time-course of updating spatial information was related to the ability of a participant to memorize a location. Updating appeared to be faster when memory performance was high. Apparently, some people could update a memorized location within very short intersaccadic intervals while others were not able to do so. This effect was especially pronounced in the version of the task in which participants were shown two saccade targets simultaneously, which allowed the preprogramming of a saccade sequence. Such rapid sequences allow for, but also necessitate the updating of the location of the second saccade target. Besides actively maintaining the remembered location, people also had to maintain and update the location of the second saccade target, which might have affected the speed by which they were updated into spatiotopic coordinates. However, also when the second saccade target was not visible until the first saccade had been executed, a similar (although weaker) trend was found. It appears that for some people, the mere fact of having to execute extremely fast saccade sequences, be it preprogrammed or not, is incompatible with simultaneous updating of a memorized location. Alternatively, these fast saccade sequences might be so taxing for them that they drop the memorized location altogether.

To summarize, spatial information is updated rapidly across intervening eye movements. Extraretinal information provides us with a precise prediction of the postsaccadic locations of relevant object, allowing them to be available for subsequent action shortly following saccade landing. The results of Chapter 3 show that it takes some time before postsaccadic retinal input influences localization. Whereas earlier studies claimed that both sources of information are optimally integrated based on their relative reliability, the main factor determining their contribution might instead be the timing of a subsequent localization response. The fact that updating based on retinal input takes more time is illustrated in the experiments in Chapter 5 of this thesis. In this chapter, the updating of a location on a moving object was studied. Although object movement induces changes in retinal input, they are not accompanied by efference copy. Instead, the brain has to rely on visual information for updating. The results described in Chapter 5 show that even though this mechanism is rather efficient, it takes considerably longer than the updating shown during eye movements.

We do not maintain much visual information across saccades. We do not need to, because information is already available in the world around. The main function of remapping of neural activity might be to facilitate the rapid updating of

potential action goals. Oculomotor areas contain topographic representations of space which could act as priority maps, in which different relevant locations are coded as potential action goals. Furthermore, the oculomotor system contains the mechanism for updating these 'pointers' by remapping them across saccades (Cavanagh et al., 2010). This mechanism allows us to make rapid sequences of successive actions, without having to process retinal information and recalibrate the spatial layout upon every new fixation. Nevertheless, if there is enough time available, this information does play a significant role. Although chapter 3 shows that its contribution levels of around 30%, this is based on experiments in visually sparse settings. This contribution might be considerably higher in naturalistic environments.

### Summary

On basis of the research discussed a number of conclusions can be drawn. First, motor structures play a crucial role in the maintenance of spatial information. The ability to maintain spatial information over short periods of time relies to a large extent on the same mechanisms used to plan and execute actions. Spatial working memory might be nothing more than the maintenance of a prospective motor code in the absence of visual stimulation. In other words, we memorize the location of an object by maintaining the motor plan that will eventually be used to act upon it.

Second, for the maintenance of spatial continuity across saccades we rely on both extraretinal and retinal information. A number of studies have demonstrated that the relative contribution of these sources of positional information is modified depending on the relative reliability and precision of both signals. The work in this thesis is in line with this idea. However, the most important factor determining their relative contribution appears to be the availability of information. Extraretinal signals are predictive, which enables the rapid updating of action goals. Compensation mechanisms driven by these signals have mainly been identified within oculomotor regions, which has led to the hypothesis that maintenance of accurate spatial representations across saccades is especially important for actions, but is less critical for perception (Bays & Husain, 2007). It is not necessary to have a precise prediction of the post-saccadic location of a continuously present object because it can easily be relocated upon saccade landing. However, in order to interact with objects and to avoid obstacles, the updating of potential movement goals is crucial. Moreover, given the frequency of our eye movements, this updating has to be fast. Even when doing a simple daily task such as brewing a cup of coffee we make multiple eye movements per second.

Gradually updating our movement goals after each of these eye movements would render such task impossible.

Although retinal information is more precise it takes considerable time to integrate new input upon saccade landing. Only if behavior does not necessitate a fast response do these signals influence subsequent action. This is illustrated by the work described in Chapter 5, which involves updating a memorized location on a moving object. In this case, extraretinal signals are not available, and a perceiver has to rely on visual input in updating the location. This resulted in a more gradual updating of oculomotor competition.

Third, the updating of endogenously maintained information is likely to rely on the same mechanism that facilitates the updating of covert attention across saccades. This remapping mechanism was initially thought to underlie the maintenance of visual stability across saccades. Given the sparse nature of conscious visual perception, this problem might in fact be non-existent. Instead, remapping might drive the updating of spatial information. It helps in the transsaccadic maintenance of a set of prioritized locations. These locations, which are agnostic to the actual visual content they indicate, have been dubbed 'attentional pointers' (Cavanagh et al., 2010). However, they might be better characterized as potential motor goals.

Fourth, saccade curvature provides a great tool for measuring the dynamics of spatial representations. Where most researchers used perceptual discrimination tasks to investigate the updating of memorized content, measuring saccade trajectories allowed us to directly and unobtrusively tap into the oculomotor system. This might also to some extent explain the discrepancies in timing between ours and earlier studies. The efference copy is likely to be generated in the SC, one of the latest stages along the sensorimotor continuum, and subsequently fed back to cortical areas through the thalamus. Neurophysiological studies suggest that the earlier in the visual pathway, the later relative to saccade onset activity is remapped, so it is not unlikely that there is some delay between the updating of oculomotor competition within the SC and the emergence of perceptual enhancement at the corresponding location.

In sum, spatial information is likely to be represented in the effector systems that facilitate a suitable response towards the object they indicate, most prominently the eye movement system. This information is continuously updated as we move through our dynamic environment and move our eyes to locations of interest.

Luckily, the oculomotor system already contains the mechanisms for efficient updating. Remapping of neural activity allows action plans to be available shortly after intervening eye movements. Retinal information also plays a role in keeping track of the objects around us, but only if there is sufficient time to integrate it.

