Prediction of space and features in action : a psychophysical approach

Rodrigo Balp

Université de Paris

ED 158 3C - Cerveau Cognition et Comportement

Integrative Neuroscience and Cognition Centre

« La prédiction de l'espace et les traits dans l'action : une approche psychophysique »

Prediction of Space and Features in Action: A Psychophysical Approach

Rodrigo BALP

Thèse de doctorat en Neuroscience Cognitives

Présentée et soutenue publiquement le 22 novembre 2019

Dirigée par

Thérèse COLLINS et Florian Waszak

Devant un jury composé de :

Laurent MADELAIN, PU, Université de Lille, rapporteur

Martial Mermillod, PU, Université Grenoble Alpes, rapporteur

Simone Schütz-Bosbach, PU, Université de Munich, examinatrice

Andrea Desantis, CU, ONERA, examinateur

Prediction of Space and Features in Action: A Psychophysical Approach

by Rodrigo BALP

supervisors
Thérèse Collins & Florian Waszak

Summary

Actions are executed by many living beings. They are the way in which we interact with the environment. It is well known that the Central Nervous System associates an action with its sensory effects, creating what is known as an Action-Effect association. The Ideomotor Theory states that these associations are then used to guide and control action, and in a long-term basis they reflect motor learning and can even affect perception. Before an action is executed, the brain predicts (pre-activates) the known effects of the action, and performs a comparison between the predicted and the actual effects. The result of such comparison allows the brain to improve future action, or correct an ongoing one. Forward Models have been proposed to be in charge of such predictions. It is not clear, however, what kind of information Forward Models integrate to be able to predict the consequences of an action. There are two main streams by which visual information is conveyed for visual perception. Broadly, the dorsal pathway is in charge of treating spatial aspects ("where") of an object to be able to recognise it in space; the ventral pathway in turn, is in charge of recognising features about the object ("what"), independently of its location. In a series of experiments, we tested what kind of sensory information is used by Forward Models to make action-effect predictions. With this information in hand we were able to infer the neural origin of received information. By considering eye movements as actions, we assumed the existence of a Forward Model in charge of oculomotor predictions. Supported by current literature, we hypothesised that oculomotor predictions would be based on spatial information as well as in features of the stimulus. In one set of experiments we tested whether visual features of a stimulus influence its spatial perception. This allowed us to determine if one type of information had preferential treatment over the other. In another set of experiments, we tested if spatial information alters a learnt action-effect association. This way we were able to determine if spatial information was treated in the same way as visual features. Taken together, our experiments allowed us to state that spatial information & visual features are treated differently depending on the type of task. We also found that both kinds of information are considered with different priorities by the predictive systems.

Keywords: prediction, space, features, forward model, ideomotor, transsaccadic correspondence, saccadic suppression, action effect

Résumé

Les actions sont exécutées par plusieurs êtres vivants. Elles permettent d'interagir avec notre environnement. Il est bien connu que le Système Nerveux Central associe une action à ses effets sensoriels créant ce que l'on appelle l'Association Action-Effet. La théorie Idéomotrice établie que de telles associations sont alors utilisées pour guider et contrôler l'action ; et au long terme elles aboutissent en apprentissage moteur et elles peuvent même altérer la perception. Juste avant qu'une action soit exécutée, le cerveau prédit (préactivé) ses effets connus et compare les effets préactivés avec les vrais effets ressentis. Le résultat d'une telle comparaison permet soit d'améliorer l'action future, soit de corriger une action en cours d'exécution. Les Modèles Prédictifs (Forward Models) ont été proposés pour mener ces prédictions. En revanche, le type d'information reçue pour arriver à prédire les conséquences sensorielles d'une action n'est pas clair. Ils existent deux principales voies par lesquelles l'information est transmise pour la perception visuelle. La voie dorsale serait, elle, en charge de traiter les aspects spatiales (« où ») d'un objet pour ainsi le reconnaitre dans l'espace. La voie ventrale est en charge de reconnaitre les traits d'un objet (« quoi »), indépendamment de la location de l'objet. Dans une série d'expériences nous avons testé quels types d'informations sensorielles sont utilisées par les Modèles Prédictifs pour faire des prédictions action-effet. Avec cette information, nous étions également capables de déduire l'origine neuronale de l'information reçue. En considérant les mouvements oculaires en tant qu'actions, nous avons assumé l'existence d'un Modèle Prédictif en charge des prédictions oculomotrices. D'après la littérature actuelle, nous avons émis hypothèse que les prédictions du système oculomoteur seraient basées sur l'information spatiale et sur les traits de l'objet. Dans un groupe d'expériences nous avons testé si les traits visuels d'un stimulus influencent sa perception spatiale. Cela nous permettrait de déterminer s'il y a un traitement préférentiel d'une de deux types d'information. Dans un autre groupe d'expériences nous avons testé si l'information spatiale peut altérer une association action-effet déjà apprise. De cette manière, nous étions en mesure de déterminer si l'information spatiale est traitée de la même façon que les traits visuels. Pris dans leur ensemble, nos expériences nous ont permis de constater que l'information spatiale et les traits visuels d'un objet sont traités de différentes manières selon la tâche. Nous avons trouvé aussi que les deux types d'information sont considérées avec des différentes priorités par les systèmes prédictifs.

Mots clefs : prédiction, espace, traits, forward model, ideomotor, correspondance trans-saccadique suppression saccadique, action-effet

Acknowledgments

I especially want to thank Thérèse Collins and Florian Waszak for their guidance, patience, advice, knowledge, help and support. None of this work would have been possible without you two.

Thank you, Pauline, for being a solid life companion and for encouraging me when I was discouraged. Thanks to Ponchita, for constantly reminding me that life is simple and should be enjoyed. Thanks to my family, mom, dad and sister who, although physically far away, were always with me and supported me. Thanks to my friends, the moments together surely were great and needed. Thanks to all the people of the lab who supported me somehow, so thank you Sabine, Aboubaccar, Stéphane and my co-workers Eef, Álvaro, Hamdi, Martina and Klara.

Special thanks to CONACyT for providing financial support during the three years of thesis.

Finally, I would like to thank myself for deciding to pursuit a doctorate. It has been over three years of intellectual and emotional growth.

Contents

Preface

Actions are executed by many living beings. They are the way in which we interact with the environment. It is well known that the Central Nervous System associates an action with its sensory effects, creating what is known as an Action-Effect association. The Ideomotor Theory states that these associations are used to guide and control action, and in a long-term basis they reflect motor learning and alter perception. Before an action is executed, the brain predicts the effects of the action and pre-activates such estimation in its corresponding sensory areas, then it performs a comparison between the pre-activated and the actual sensory effects. The result of such comparison allows the brain to improve future action, or correct an ongoing one. Forward (and Inverse) Models have been proposed to be in charge of such predictions. It is not clear, however, what kind of information these predictive models integrate to be able to predict the consequences of an action.

In the studies presented in this thesis, we used psychophysics to test the way in which Forward Models make predictions about the location and the features of stimuli. This was done with help of the Ideomotor Theory. While the majority of current studies has mainly investigated how visual features of stimuli are predicted, we tested the way in which spatial predictions occur. We hypothesised that a visual Forward Model would process spatial information as well. We interpret the results based on Forward Models as to facilitate the incorporation of the Ideomotor Theory to oculomotor control.

In a series of experiments, we tested what kind of sensory information is processed by Forward Models to make predictions. We did this by manipulating spatial and feature information about a stimulus. We based our hypotheses on the two main streams by which visual information is conveyed for object recognition. The dorsal pathway, which is in charge of treating spatial aspects ("where") of an object to identify it in space; and the ventral pathway, which is in charge of recognising features about an object ("what") independently of its location. These two kinds of information form the object's identity.

By considering eye movements as actions, we assumed the existence of a visual Forward Model in charge of oculomotor predictions. Supported by current literature, we hypothesised that oculomotor predictions would be based on spatial information as well as in features of the stimulus. With this information in hand we were able to infer the neural origin of information processed for predictions.

In one set of experiments we tested whether visual features of a stimulus influence its spatial perception. This allowed us to determine if one type of information had preferential treatment over

the other, and to what degree. In another set of experiments, we tested whether spatial information can alter a learnt Action-Effect association. This way we were able to establish if spatial information is predicted in the same way as visual features.

Taken together, our experiments allowed us to recognise that spatial information & visual features are treated differently depending on the type of task. We also found that both kinds of information are considered with different priorities by Forward Models.

Roadmap

This thesis contains three main parts. In the first part (I THEORETICAL FRAMEWORK) I will briefly show the theory behind the current work. First, I will show the physiological and neural properties of the motor system, from low-level muscle synapses to high-level cortical brain regions (1 NEURAL BASES OF MOTOR ACTION). Next, I present current theories about how the brain controls actions (2 THEORIES OF ACTION CONTROL). I will point out that the brain can improve its actions and that this improvement can occur due to the brain's ability to learn (2.1 ACTION-EFFECT ASSOCIATIONS). We will see that learning is achieved by keeping an inner model of the outside world which is updated after each action (2.2 IDEOMOTOR THEORY). To update the model, the brain has to evaluate and compare the predicted outcome of the action with the actual outcome once it becomes available to the senses (3 ACTION PREDICTION: FORWARD AND INVERSE MODELS). I will explain that according to current theories of motor action, motor predictions are carried out by a Forward Model which receives necessary information to perform the prediction. The following chapter will deal with the visual system (4 THE VISUAL SYSTEM). Like for motor action, I will outline its general physiological and neural bases (4.1 INFORMATION FLOW THROUGHOUT THE VISUAL SYSTEM), but I will focus on how information is transferred through different networks of the visual pathway. Then I will describe how eye movements take place, in particular saccadic eye movements (4.2 ACTIONS WITHIN THE VISUAL SYSTEM). Notions about the visual system that concern the present work will then be introduced (4.3 SACCADIC CONTROL: SC AND FEF and 4.4 VISUAL OBJECT RECOGNITION). Then, I will expose different approaches in which research has explained visual stability, a long-standing problem in vision (4.6 VISUAL STABILITY AND TRANSSACCADIC CORRESPONDENCE). We will see why visual information does not reach conscious awareness during saccadic eye movements, and how this phenomenon is studied in the laboratory (4.7 SACCADIC SUPPRESSION OF DISPLACEMENT). This will lead us to view saccades as "simple motor actions" which hypothetically follow the same principles as body movements. With this in hand, I will present the general hypothesis of this thesis (5 THE IDEOMOTOR THEORY FOR EYE MOVEMENTS).

The second part (II EXPERIMENTS) presents the experimental work we did to test our hypotheses, as well as data analyses, results and discussion for each study. At the end of the second part (DISCUSSION), I expose some points that deserve discussion, improvements or experiments that can potentially shed some light into how predictions are achieved.

Note on styles

Throughout the document different styles are used to facilitate reading.

REFERENCES, are presented in SMALL CAPS. They will (usually) be accompanied by the page number of the target section. In the electronic version of this document references are hyperlinks to the corresponding section.

{Page number} references will be surrounded by curly brackets in smaller font.

Emphasis in words or phrases use *italics* to highlight a particular notion within the paragraph.

Strong emphasis in words or phrases highlight a particular notion within the section or chapter. ***Textual citations*** use the same typography.

`Typewriter` font is used to refer to results, labels, code, modelling output and in general anything machine- or computer- related.

I THEORETICAL FRAMEWORK

Actions are executed by many living beings. They are, after all, the way in which we interact with the environment. Originally, a mechanism of survival for more primitive life forms, actions have evolved along with the species to become of such complexity that it is difficult to understand how they are formed, planned, executed and achieved. Actions are so anchored in our life, that many of them go unnoticed. However, when we stop to think about how actions come into existence to accomplish a specific goal, we can see how complex they are.

Voluntarily moving a finger is a simple action that needs motor preparation. In such preparation, the brain will determine the timing of such movement, its duration, the location of the executing body part and the force that the muscles will implement. Playing the piano involves a concatenation of such simple actions by coordinating their preparation parameters. The correct execution of all the movements is carried out to attain the executor's goal.

The study of actions has attracted interest for many centuries, and has been studied from different disciplines such as philosophy, theology, physics, psychology, neuroscience, economics, etc. What concerns us in this thesis is the neuroscientific & psychological point of view about actions.

1 NEURAL BASES OF MOTOR ACTION

An action is a physical movement executed by a person or an animal. Strictly speaking, it involves a skeletal muscle contraction or tension. In general, there are four neural subsystems which are responsible for movement control, as shown in FIGURE 1 {p. 2}.

1.1 Lower Motor Neurons and Local Circuit Neurons

Located in the grey matter of the spinal cord and the tegmentum of the brainstem, the first subsystem contains *lower motor neurons* and *local circuit neurons*. Lower motor neurons send output signals out of the brainstem & spinal cord into skeletal muscles of the body. Each lower motor neuron contacts muscle fibres within a single muscle. The ensemble of lower motor neurons innervating a single muscle are called the muscle's *motor neuron pool*. Lower motor neurons can be further subdivided into two groups: α-motor neurons and smaller ϒ-motor neurons. α-motor neurons are responsible for the generation of force signals into the muscle fibres needed for movement and

posture. A single α-motor neuron is associated with a set of unique muscle fibres, which together are called *motor unit*. They constitute the smallest force that can be executed by the muscle. The recruitment of different motor units is the way in which the nervous system produces specific movements for different contexts with a particular force. For example, eye movements are a rotation of the eyes in the orbits, and require precise small forces to locate the eye at a certain position. Extra-ocular motor units are very small and have a very high proportion of muscle fibres capable of applying specific forces at very high velocities. Force is regulated by increasing or decreasing the amount of motor units active at a time. Notably for the aim of this thesis, ϒ-motor neurons' function is to regulate the sensory input of an action. Local circuit neurons receive input from descending brain regions and represent almost all synaptic input for lower motor neurons.

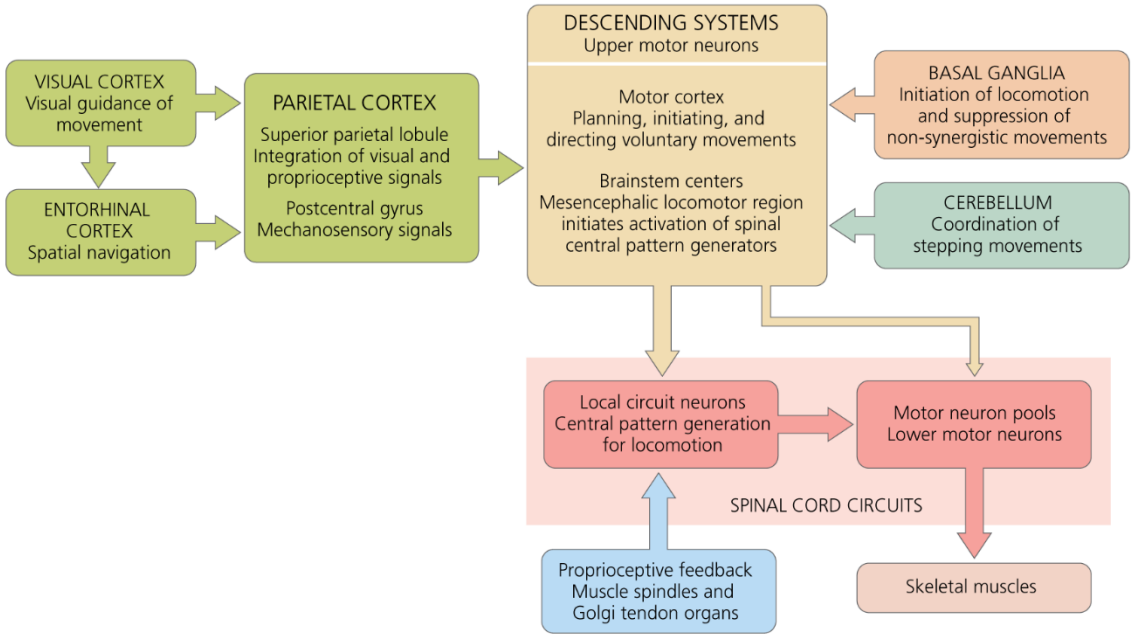

Figure 1 The four systems in charge of movement control. Descending systems include the motor cortex which has efferent connections to local circuit neurons. Altered by basal ganglia and the cerebellum, upper motor neurons located in the motor cortex can also signal lower motor neurons which selectively activate motor units. Parallel to the descending pathway of motor execution there is a path of proprioceptive feedback with terminal synapses in sensory cortices and the cerebellum itself. In green, higher functional control is guided by the visual, parietal and entorhinal cortices. Together, these systems conform a loop that guides action. After Drew and Marigold, 2015 and Purves, 2007.

1.2 Upper motor neurons & the Cerebral Cortex

The second subsystem consists of *upper motor neurons*, located either in the brainstem or in the cerebral cortex. Upper motor neurons synapse with local circuit neurons and lower motor neurons.

Upper motor neurons located in the frontal lobe are located in the primary motor cortex (Brodmann's area 4) and several sites in the pre-motor cortex (Brodmann area 6) and the supplementary motor area (SMA). They are responsible for muscle tone regulation, and for the orientation of the eyes, the head and the body in general. Pathways that begin in the cortex have been found to be essential in the planning, initiation and directing sequences of voluntary movements that involve the head, the trunk and the limbs (Drew and Marigold, 2015). Located in the Frontal Eye Field (FEF, Brodmann's area 8) lay important cortical areas which are essential for the control of eye movements.

1.3 Cortical Motor Areas

John H. Jackson found that the motor cortex contains a "spatial map" of the muscles of the body. In the beginning of the 20th century more evidence about the motor maps in the cortex was confirmed by Charles Sherrington, who published the mapping of the motor cortex in apes (Sherrington, 1906). Using microelectrodes stimulation, Edward Evarts recorded the electrical activity of individual motor neurons in monkeys (Evarts, 1986). He proposed that the primary motor cortex contributes to the initial phase of recruitment of lower motor neurons involved in the generation of finely controlled movements. His findings, among many others, suggested that the activity of upper motor neurons in the cortex controls movements, rather than individual joint or muscles.

Next to the primary motor cortex (M1, Brodmann's area 4) lie the Pre-Motor Cortex (PMC) laterally and the Supplementary Motor Area (SMA) rostrally (together they constitute Brodmann's area 6; see FIGURE 2), which also contribute to motor function. Over 30% of the axons in the corticospinal tract come from neurons in the pre-motor cortex. It uses information from other cortical regions to select movements according to a goal and a context of action. These three regions contribute differently to movement.

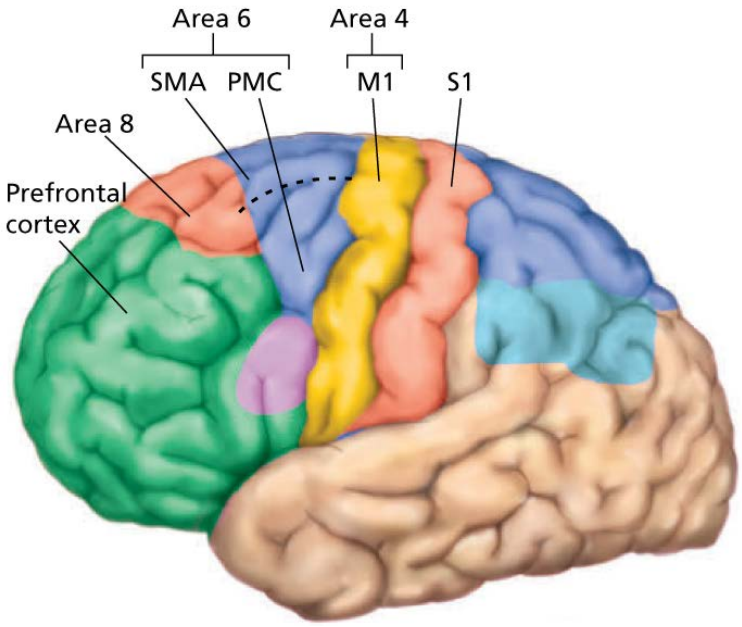

Figure 2 Lateral view of the brain showing the premotor cortex and the primary motor cortex from Purves, 2007.

The main difference between the pre-motor cortex and primary motor cortex is the strength of their connections to lower motor neurons (see FIGURE 1 {p. 2}), the pre-motor cortex containing only a fraction of the primary motor cortex's connections. Another difference is the way in which space is processed: movements encoded by the primary motor cortex comprise personal space, while movements encoded by pre-motor cortex encompass external space (Passingham, 1993). The SMA is in charge of internally-generated movements, as opposed to movements generated by sensory events (Roland et al., 1980). Other functions of the SMA include the stabilisation of the body's posture (Penfield and Welch, 1951), bimanual movement control (Serrien et al., 2002) and the coordination in time of movements' sequences (Lee and Quessy, 2003; Shima and Tanji, 1998).

Although neurons in the pre-motor cortex can show similar functionality to those of M1 or the SMA, like firing in accordance to a movement's timing, other neurons appear to encode the *intention* to perform a particular movement and seem to be particularly involved in the selection of movements based on external events (sensory based), unlike the SMA. The pre-motor cortex also contains a subset of neurons that respond when a movement is executed and when the same movement is observed being performed by another individual (di Pellegrino et al., 1992). These neurons are called *mirror motor neurons*. It has also been found that mirror motor neurons respond less when the same actions are mimicked without an action goal (Rizzolatti et al., 1996).

1.4 The Cerebellum & the Basal Ganglia

The third and fourth subsystems control movement indirectly. They do this by regulating synaptic activity of upper motor neurons located in the cerebral cortex and the brainstem. The third subsystem consists of the cerebellum. The cerebellum's function is primordial for detecting and attenuating the difference (motor error) between an intended movement and the movement that was actually executed. Interestingly, the cerebellum not only intermediates these errors in real-time, but also in a long-term basis that later becomes a form of motor learning.

The fourth subsystem is the basal ganglia, located deep in the forebrain. Their main function is to inhibit upper motor neurons from unnecessary or unwanted movement. Such inhibition allows this subsystem in the regulation of transitions from one pattern of voluntary movement to another. Parkinson's and Huntington's diseases have their roots in problems associated with basal ganglia disorders.

The four subsystems are in charge of movement control whether it is voluntary or not. Movements are therefore the result of a complex temporal and spatial set of brain patterns that ultimately send tension or contraction signals to muscles.

2 THEORIES OF ACTION CONTROL

Motor behaviour is executed to achieve specific goals based on an agent's context, beliefs and desires. It can be voluntary, but can also be reflexive and have no conscious preparation. Actions give us the ability to produce an effect in the environment that surrounds us. After an action has been executed, the surrounding environment experiences some kind of impact. This impact is usually called *effect*, but more precisely it refers to the perceptually available consequences of the action. The effects of an action are perceived by our senses, and this allows us to determine if the goal was successfully accomplished. Perception is strongly coupled with action because it allows us to evaluate the correctness of our actions by classifying their effects. Take for example lighting a match. Our goal would be to have fire, the action would be scratching the match against the side of the matchbox and the desired effect would be fire on the match. After the action is executed, we will see, feel and listen to the consequences of the action; after collecting this sensory input we would be able to determine if our goal was correctly achieved or not. Nevertheless, the action has to be carefully prepared. From grabbing and opening the matchbox, to taking out a match, to holding it with the fingers and sliding it along the side with a specific force on the fingers for it to slide fast without breaking. Although the match lighting may be seen as the action, it actually comprises a series of actions that succeed each other. Examples like this can be found every day without us being aware that they happen at all. Right now, as you read these words, the visual system of your brain is carefully making eye movements to place the fovea on each word, yet another example of an unconscious action. From now on, when referring to action, I specifically mean *voluntary* action, to make the distinction with involuntary or reflexive actions. A very important aspect of human action is the fact that in order to achieve an agent's desired effect, the brain must have *previous knowledge* about a specific action's effects (Kühn et al., 2010). In other words, an internal representation of action and effect relationships. These Action-Effect associations are necessary for goal-driven behaviour. Not having such a mapping would imply that actions are executed "blindly" and with absolute uncertainty about goal fulfilment. A lack of incoming perceptual information would make learning impossible. It is therefore of main scientific interest to study motor action experimentally by analysing how Action-Effect associations are created.

2.1 Action-Effect Associations

The concept of association between action and perception has its roots in the first half of the 19th century, when Herbart, Lotze, Harless, Laycock, Carpenter and James, among others, began to theorise how action control operated. Carpenter, for example, realised that «*the expectation of a*

result is sufficient to determine [...] the muscular movements by which it is produced»[1]. Laycock had a similar revelation when he noticed that hydrophobic patients were not only physically afraid of water, but would also be afraid of the suggested idea of water. He concluded that intention for future actions consciously arises before the act itself (Laycock, 1860). Herbart's ideas on action control went further, coming closer to the current knowledge about actions. He considered the role of *learning* in the development of motor codes. He proposed that actions were initiated by the to-be-produced sensory effects by means of an *automatic association* of actions and their sensory effects when executed, and a purposeful use of the associations to initiate actions and activate their effects (Herbart, 1826). William James termed *ideomotor action* the co-occurrence of sensory input and action, and suggested that the direct relation between ideas and movements was a critical factor to induce voluntary movements. This view assumed that an action was represented in terms of its sensory effects (James, 1890).

The concept of learning introduced by Herbart and James allowed Anthony Greenwald (1970) years later to restate the theory in more verifiable terms. He proposed that mechanisms of action control imperatively require sensory feedback, and that the idea of a response and the contiguity of events are sufficient for instrumental conditioning. Greenwald's ideas about the ideomotor mechanism initially proposed by James involved three types of "events": Stimulus, Response and Effect (S-R-E). The learning of an action would involve unique sequences of S-R-E triplets. Repeated exposure to any triplet would make the agent automatically associate the relationships between them. These associations would be learnt by the agent so that later effects could be anticipated. In the ideomotor mechanism, a motor command is coded with the intrinsic feedback that it aims to generate (Shin et al., 2010).

To prove the existence of these automatic associations, Greenwald proposed an experimental method which involved an acquisition phase (also called training or practice phase) and a test phase. During the acquisition phase participants were repeatedly exposed to a unique combination of S-R-E in which a stimulus S involved a specific response R from the subject. The response R was linked to a specific sensory effect E. In the test phase, subjects were tested on their responses by modifying the previously experienced S-R-E triplets. When subjects encountered an S-R-E triplet they had not seen before, they displayed longer response time to the stimulus S. Already experienced S-R-E triplets displayed faster and more accurate responses to the stimulus S. Thus, the acquisition phase was used to create the association among Stimulus, Response & Effects; the test phase served to get a measurement of the priming effect of E. Nowadays the acquisition-test paradigm is still used to

[1] On Carpenter, 1876, p. 287

investigate Action-Effects associations. The automatic learning of contingent relationships between actions and their effects has been termed Action-Effect Paradigm, and there is vast evidence supporting both the paradigm and the automatic acquisition of A-E associations.

The basic concept of Action-Effect anticipation is that when subjects are exposed to sequences of stimuli containing a specific order or a set of rules behind them, they tend to improve their performance and respond faster as compared to randomly arranged stimuli, even when they are not aware of these rules. In a pioneer study conducted by Morin and Grant (1955) they presented an 8-choice task to subjects. There was a row of response buttons and parallel to it a row of stimulus lights. When one of the response buttons was pressed, a light (the effect) was lit. They used different Stimulus–Response and Response–Effect mappings across subjects, and varied their compatibility and difficulty: some would match perfectly (each button turns on the light right above it), others were random, and yet others were inverse mirror mapping (leftmost button turns on rightmost light, and so on). After several practice sessions, the feedback lights were no longer used, and response times (RT) increased significantly. This was true especially for mappings without structure. The increase in RTs showed that subjects relied on the feedback during the practice phase.

Elsner and Hommel also investigated the learning of Action–Effect associations (Elsner and Hommel, 2001). In an acquisition phase, subjects performed two different key-presses which in turn produced two different tones. In the test phase, the same tones were presented as target stimuli for a speeded-choice response task. RTs were faster in response to the keypress-tone association learnt at the acquisition phase as compared to the alternative action. With these results, Elsner and Hommel showed that subjects acquired bidirectional associations between the motor code of the action and the perceptual code of the auditory effect.

Wilfried Kunde showed that the temporal influence of Response-Effect compatibility persists even when responses are cued in advance (Kunde, 2003). He claimed that this result suggests that at least part of the compatibility originated from response-generation processes, which occur later than the response selection stage. Janczyk and colleagues showed that free-choice and forced-choice actions (i.e.: goal-driven & stimulus-driven actions) are bound similarly to their effects and demonstrated that Action-Effect associations occur for any type of action (Janczyk et al., 2012).

The ability to know that a given action will in fact create some *context specific* effect comes from the capability of pre-evaluating the sensory effect of one's actions (Elsner and Hommel, 2001; Hommel et al., 2001). Such ability would have occurred initially by the execution of random movements which would be associated with their environmental effects, and in a second step, the activation of the effects in memory would lead to the (p)reactivation of the movements that lead to this effect (Harleß,

1861; Lotze, 1852). By executing and repeating the movement, its relevant sensations are learnt and associated to the motor command that originated the movement.

These findings corroborate effect-based theories of action control which assume that the selection, initiation and execution of movements is mediated by anticipation of their sensory effects.

2.2 Ideomotor Theory

The Ideomotor Theory roots in the question of how an "idea" produces its intended action, and claims that performing an action results in a *bidirectional* association between the action's motor code and the sensory effects the action produces (Shin et al., 2010). The term "bidirectional" refers to a top-down and bottom-up neural activation, the former consisting of the motor action and the latter of its perceived effects. As Greenwald suggested, an action is automatically triggered by the anticipation of its sensory effects (Greenwald, 1970). Therefore, because of the bidirectionality, any activation of the effect "image" will trigger the corresponding action. Once the bidirectional associations are acquired, they can be used to select an action by anticipating —or simulating— their perceptual consequences (Elsner and Hommel, 2001; Herwig et al., 2007; Nattkemper and Prinz, 1997). In this way, the re-activation of such sensations will produce the same movement, or a tendency towards it. Nowadays, it is well established that the effects of an action play an important role in the selection, preparation, and execution of motor actions.

The Ideomotor Theory attempts to explain how perception arises. It points out that perception results from the interaction between the sensory input and an internal model of the world generating predictions about the ongoing sensory input (Schröger et al., 2015). One of the novelties introduced by this theory is the bidirectionality that allows action and perception to be conceived jointly, whereas previous approaches devised them separately. Yet, two things remained unexplained: a) How the internal model generates predictions and b) the nature of the interaction between low-level information (the sensory input) and high-level information (the predictions generated by the internal model and motor planning).

2.2.1 Common Coding Theory

The Common Coding Theory complements the Ideomotor Theory by stating that perception and action use the same code, or internal representation (Prinz, 1990, 1984). Until the arrival of the Common Code Theory, perception and action were thought to use separate "languages", which

according to Prinz, implied «*that afferent and efferent codes are incompatible at all levels*»[2]. Prinz's rationale was based on the idea that the set of instructions presented to a participant *before* the experiment, is the *true* beginning of the task, meaning that the rules that will guide action and lead perception begin their formation before the experiment commences. Instructions presented by the experimenter guide participants to perform an action given an experimental condition. In this case, *afferent processes* refer to the information in the stimulus that will guide an action given a condition, and *efferent process control* to what the subject does under such conditions. The separate-coding approach implied that there was a specific unit in charge of *translating* from stimuli to response, and another unit in charge of *choosing* a response given stimuli. The question which Prinz highlighted was, how can choice of one code among several efferent codes be accomplished on the basis of afferent information? This problem was solved by viewing actions and percepts as characterised by the same attributes, the only difference being that percepts refer to ongoing, actor-independent events, while actions refer to to-be-generated, actor-dependent events. Being represented using the same code facilitates the treatment of both streams of information. The common representation would be possible because of goal-directedness of action, because «*common codes would only serve to produce epiphenomenal copies of the ongoing processing*»[3]. Recent studies have suggested that a common representation would help a "mirror system" simulate movements we see in others. This would be achieved by mapping the observed action into one's motor code for the same action (Rizzolatti et al., 2001).

2.2.2 Intentional Binding & Sensory Attenuation

Action-Effect anticipation brings into view two experimental effects: Intentional Binding and Sensory Attenuation. Intentional Binding studies have found that a voluntary action is perceptually *bound* in time to the following sensory effects. This implies a sense of agency, in such a way that the awareness of a voluntary movement brings together in time its learnt sensory effects (Engbert et al., 2008). Intentional Binding is based on two main findings. The first is that the mental representation of an action activates its sensory effects. The second one, exposed that the strength of agency operates over a limited time-window, in such a way that as the time-gap between an action and its effects increases, subjects have the impression they did not cause the sensory effect (Wegner and Wheatley,

[2] On Prinz, 1990; p. 169
[3] On Prinz, 1990; p. 172

1999). Given that the sense of agency does not concern the present work, we will limit Intentional Binding to the latency between an action and its perceived effects.

In 1983 Benjamin Libet and colleagues carried out an experiment that completely changed the view on voluntary action, and even the philosophical concept of free-will (Libet et al., 1983). Libet recorded cerebral activity (EEG) before and during a fully voluntary hand movement (completely endogenously generated), and related EEG activity to participants' subjective experience of *wanting to move the fingers/hand*. While being recorded, subjects sat in front of a display that showed a clock they should watch at all time. They were instructed to wait for a full revolution of the clock, after which they should flex the fingers or the wrist of the right hand whenever they felt the "urge" to do so. After the voluntary action was executed, subjects reported the recalled timing at which they had felt the urge to move. Libet found that the onset of the readiness potential activity preceded the reported time of conscious intention to move by 350 ms, on average. Using a similar approach as Libet in his famous 1983 paper, Haggard and colleagues measured the perceived time of actions and their effects to study conscious action (Haggard et al., 2002). They found that subjects perceived voluntary actions occurring later, while their effects were perceived earlier than they had actually happened. They induced involuntary movements using Transcranial Magnetic Stimulation (TMS) and found that such time-binding effect was not present. They concluded that the Central Nervous System (CNS) executes a mechanism that produces Intentional Binding, to make the subject aware of its actions and their consequences. A few years later Engbert et al. generalised previous results to visual, auditory and somatic effects of action, and stated that the binding effect —i.e.: agency— was produced by efferent motor commands (Engbert et al., 2008). These results show important evidence of how action control occurs by bringing into "the picture" awareness about one's actions.

The Sensory Attenuation phenomenon makes a main distinction between the perceived effects of an endogenously-generated and an exogenously-generated action, and states that exogenous events triggered by an agent tend to attenuate the perception of the effects. An excellent example of Sensory Attenuation is that is difficult to tickle oneself (Blakemore et al., 2001). Many studies have shown that electrophysiological responses to a predicted stimulus results in an attenuated response, whereas responses to an unknown stimulus increases. Schafer & Marcus (1973), for example, showed that self-stimulation (that is, a self-administered auditory stimulus, as opposed to a machine induced one) resulted in smaller auditory N1 component amplitudes. Sensory Attenuation has been extensively studied in voluntary action control, and has led to the conclusion that when an action's effect is predicted, the phenomenological experience of the sensory event is altered (Blakemore et al., 2000).

Cardoso-Leite and colleagues used Signal Detection Theory (SDT) in their experiments to evaluate sensitivity (d prime or d') and criterion (c) of a task in which subjects had to detect the orientation of a Gabor patch which was present in 50% of the trials (Cardoso-Leite et al., 2010). They found that d' for the detection of the patch decreased by 10% when the patches were triggered by the previously learn action (congruent trails). The observed Sensory Attenuation was explained by the pre-activation account, which hypothesises that the execution (or the preparation) of an action pre-activates sensory networks which contain the expected perceptual consequence. These networks increase their mean level of activity to some "pedestal" level once the learning phase is completed. Thus, the found decrease in sensitivity for congruent trials was due to the small signal difference with respect to the pedestal level. In other words, it was easier to detect a signal closer to the baseline level than to detect a signal at the pedestal level, because it included the pre-activation signal. This way, sensitivity in congruent trials was reduced not because the mean of the signal distribution had decreased, but because the mean of the noise distribution had increased (Cardoso-Leite et al., 2010). The following figure illustrates the pre-activation hypothesis.

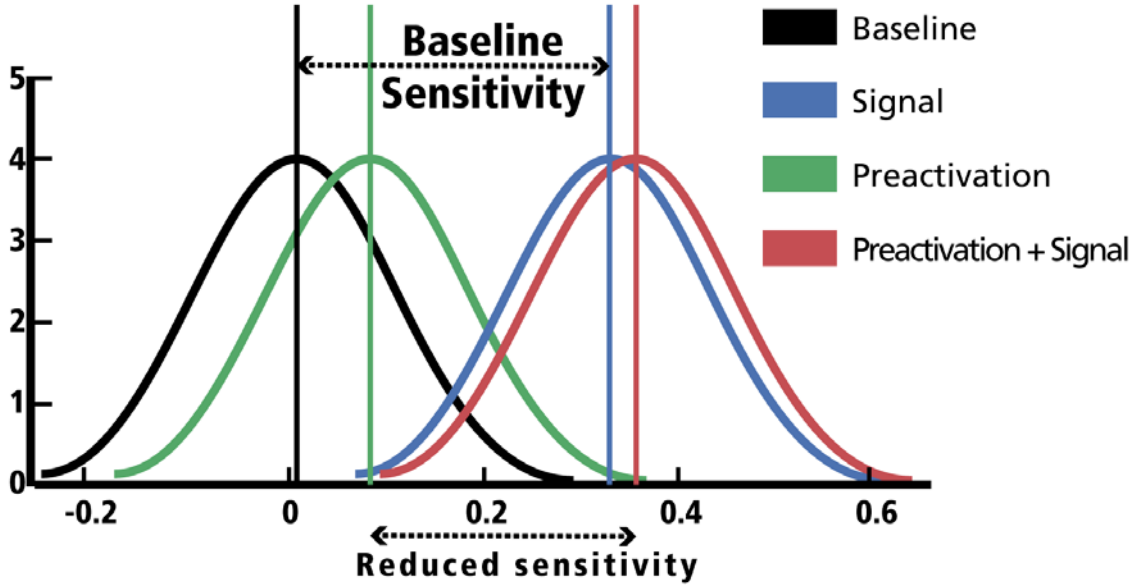

Figure 3 Schematic representation of the pre-activation hypothesis. Performing an action results in the internal pre-activation of the sensory representation of the action's expected perceptual effects. The internal action effect anticipation increases the mean level of activity in the network representing the expected effect to some pedestal level. After Roussel et al., 2013.

3 Action Prediction: Forward and Inverse Models

The theories previously described implied the existence of a system in charge of comparing the expected sensory effects with the actual effects, or as put before, "in charge of generating predictions of continuous sensory input". The concept of prediction arises side by side with the fact that action is guided not only by perception, but by *prior* experience. In this sense prediction is the result of a mechanism that takes into account previous knowledge about the effects of an action, and generates the most plausible outcome of such action. Such prediction will be tightly linked to the timing between the action and its perceived effects, as has been demonstrated by the Intentional Binding effect studies.

Before going into the theory of predictive models, we must first present some basic background on the relationships between cause and effects. This will be needed in the following sections, because internal models, as used in this context, are black-boxes that receive some input —generally seen as the cause— and send some output —regarded as the effect. Basically, action control has been built upon cause and effects relationships: a movement is generated (cause) which induces an environmental change (effect) perceived by the senses (cause) to be finally compared with the expected environmental change (effect/cause). Models of prediction[4] developed concurrently to Action Control theories and include many domains of study, such as cybernetics, robotics and artificial intelligence, among others.

3.1 Predictive Models

The notion of prediction has given rise to computational models and theories that explain how prediction systems work. Such theories aimed to explain adaptive behaviour and were later adopted by Action Control theorists to explain how an internal model could manage predictions to guide action. It has been suggested that the CNS simulates behaviour, and that such simulation helps in planning, learning and control (Miall and Wolpert, 1996). As we saw in previous sections, there are various brain regions involved in planning, learning and control; thus, such brain areas play a very important role in how predictions are administered.

Prediction has been studied based on the flow of information between brain networks. Internal models have been broadly categorised into two different classes depending on their function. The

[4] I will freely interchange predictive model with model of prediction.

first category consists of Forward Models (FM) which are in charge of the causal representation of the motor system (Jordan and Rumelhart, 1992). The FM's task is to predict the sensory consequences of an action based on a copy of the motor command. In theory, the Forward Model receives as input the motor command to be executed and the current state of the system. Such input could, for example, include information about how muscles & joints must be contracted to attain a desired position (the motor command) and their current state. The FM captures the state transition given a motor command and estimates the next state. This estimation can be further "cascaded" to another FM in charge of estimating the expected sensory feedback, or as depicted in FIGURE 3 {15}, a single FM could do all the work. Taken altogether, given the actual state of the system and the motor command to execute, these FMs could predict the sensory effects of such motor command. In principle, forward models could be used by multiple motor systems, and with all possibility there might be several forward models in different brain regions. As stated above in THE CEREBELLUM & THE BASAL GANGLIA {p. 5}, the cerebellum has been observed to be primordial for detecting and attenuating the motor error between an intended movement and the movement actually executed, and has been proposed as a candidate for containing forward models (Miall and Wolpert, 1995). This has been corroborated by studies in which patients with cerebellum damage show poor performance on tracking tasks in which they are deprived of visual feedback of their hand's position (Haggard et al., 1995; Ishikawa et al., 2016; Stein, 2009). This suggests that when no forward model is present, hand coordination, or even ocular-manual coordination, are not possible (Vercher et al., 1996; Vercher and Gauthier, 1988). Forward Models have also been proposed to be involved in oculomotor control, supposedly located in the brain stem; however, given that oculomotor control takes place all over the brain, it is very probable that other forward models in charge of eye movements be at other regions of the brain (Ishikawa et al., 2016).

The second category of internal models entails Inverse Models (IM), which are in charge of *inverting* the causal flow of the motor system (Atkeson, 1989). Inverse Models take as input the system's current state and the desired state transition, and generate as output the causal events (motor commands) that with most certainty would produce the new/desired state. The uses of FMs and IMs include several neural aspects like cancelling sensory reafference, (distal supervised) learning, state estimation & prediction, internal feedback, overcome of time delay, mental practice, etc. (Miall and Wolpert, 1996).

Other studies that used Event Related Potentials (ERPs) have proposed that attenuation of P2 or N1 components is a marker of a correctly predicted sensory feedback by the FM. The repetition of a stimulus results in a decreased brain response due to sensory adaptation (as seen in INTENTIONAL

BINDING & SENSORY ATTENUATION {p. 10}). Harrison and Ziessler used ERPs to show that effect anticipation affects perceptual, cognitive, and motor phases of response preparation (Harrison and Ziessler, 2016).

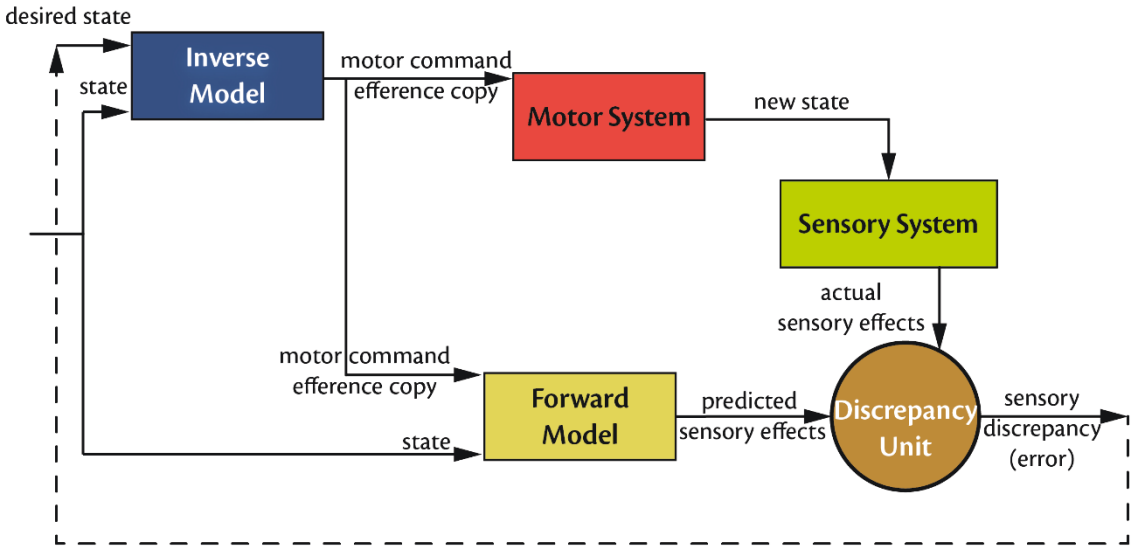

Figure 3 A "Predictive System" showing the interaction between a Forward Model, an Inverse Model and the Motor System, to finally give output to the discrepancy between predicted and actual effects. The Inverse model receives as input the actual state of the system (for example, the state of a finger) and the desired state of the system, as an output it yields the motor command that can achieve such state given the actual one. The motor system later receives the Inverse Model's output as input to execute the necessary joint or muscles changes; after execution, some state is achieved. The achieved state is then used as feedback for the Inverse Model as a training signal. Such training signal will in further occasions, refine the motor command given the initial desired and actual states. The forward model simulates the state change outputting the expected sensory effects of the action. Once the sensory system has sent the actual sensory feedback, the actual and the expected sensory consequences are compared to guide learning or to correct the action. Based on Miall and Wolpert, 1996.

FIGURE 3 shows an example of how the interaction between a single FM and an IM could guide and achieve motor learning. The figure is a combination of previously proposed Forward and Inverse (Miall and Wolpert, 1996, 1995). The Inverse Model receives as inputs the desired state and the current state of the system, and outputs (a candidate of) the motor command that would create the transition state. The candidate motor command and the efference copy are fed as input to the motor system and to the forward model, respectively. The Forward Model receives as input the motor command, the efference copy and the current state of the system, and outputs the expected sensory effects that would be produced in the current state with the received motor command. The efference copy is a duplication of a movement's motor code generated by the motor system. The motor system

implements the movement using this motor command, which transitions the system into a new state. Note that in FIGURE 3 time moves from left to right, meaning that the forward model receives input before the motor system executes the movement. The sensory system also passes to the new state and outputs some real sensory consequences of the action. The actual sensory feedback and the predicted/estimated one are fed into a comparison unit that outputs the sensory discrepancy or error, that is, the difference between the actual and the expected sensory effects. Once this happens, the current state-transition loop ends. The sensory error is fed back into the Inverse Model in order to begin a new state-transition loop. This way, the next action can be corrected, and in the long-term, learnt.

3.2 Predictive Coding

As seen in previous sections, it was suggested that perception is the result of the interaction between sensory input and an internal model of the world that generates predictions about the ongoing sensory input (Schröger et al., 2015). Until the arrival of Predictive Coding, there were few theories that explained how this "internal model" accomplished *predictions*, or the way in which predictions were operated (Atick, 2011; Barlow, 1961a, 1961b; Barlow and Levick, 1976; Srinivasan et al., 1982).

Perception was originally viewed as a mechanism that resulted from "unconscious inferences" (Helmholtz, 1867). Even though perception is nowadays explained in broader terms, inference remains the best way in which prediction can be explained and implemented. Under the inferential view, the internal model of the world contains inferred causes of sensorial input (Schröger et al., 2015). The brain's inferred causes are then used to generate predictions about sensory input. Such predictions are compared with actual sensory input, and the difference is encoded as "*prediction error*" (Rao and Ballard, 1999); prediction errors can then be used to update the internal model of the world.

Predictive coding therefore conceives visual perception as an iterative inference process that matches[5] top-down *predictions* with bottom-up *evidence* (Mumford, 1992; Friston, 2005; Egner et al., 2010). This dynamic matching process sends information hierarchically through the visual pathway by executing two tasks by two different type of modules: 1) *representational modules* that encode the probability of stimulus as an "expectation" and then supply lower-level areas with predictions regarding expected inputs; and 2) *error modules* that encode discrepancy between predictions and

[5] Or "differentiates".

sensorial bottom-up evidence as "surprise" and then transfer such information to higher-levels in which representations are adapted in order to eradicate errors (Friston, 2005). More recent studies further support predictive coding, contrasting it to the traditional view of the visuo-cortical pathway in which neurons are simply feature-based decoders (Hubel and Wiesel, 1962; Riesenhuber and Poggio, 2000).

Currently, perception is seen as the result of Bayesian inferences in which bottom-up sensory information and top-down predictions are fed-forward and backwards, and integrated hierarchically across the cortices (Lee and Mumford, 2003). This dynamic system of feedback and feedforward recurrent loops aims to minimise prediction errors. Given that at each level of the cortical hierarchy only the prediction error is passed onto the higher levels, only mismatching information is fed forward, hence reducing computational resources.

Studies have found experimental evidence supporting predictive coding's basic principles. One of such studies found that V1 neural activity was attenuated when a stimulus was predictable (Alink et al., 2010); another found that V1 and other regions displayed strong activations when the stimulus was not predicted, which was interpreted as a way of encoding the surprise level of the mismatched prediction (den Ouden et al., 2009). These two findings suggest that both expectancy and surprise —the basic information types of predictive coding (Friston, 2005)— determine neural activity in the occipital and temporal lobes (Egner et al., 2010). Finally, studies of repetition suppression—the fact that stimulus-evoked neural activity is attenuated on stimulus repetition— have found that it reflects a decrease of top-down perceptual prediction error when an expected stimulus is processed, contrary to unexpected stimuli (Summerfield et al., 2008). These findings propose that repetition suppression is mediated by expectation either visually (Summerfield et al., 2011) or auditory (Todorovic et al., 2011).

Even though there exists vast evidence of visual surprise signals, it is controversial if such signals are indeed prediction error computations, or if they reflect an attentional highlighting of unexpected stimuli orchestrated by another predictive process elsewhere in the brain (Egner and Summerfield, 2013). This debate has open a new line of research in which expectation and attention are combined to discover the role of attention in prediction (Jiang et al., 2013; Kok et al., 2012; Feldman and Friston, 2010; Wyart et al., 2012).

The basic principle of the predictive coding theory is depicted in FIGURE 4 {p. 18}. Inferred causes of sensory input are encoded in representational units (R). R-units are high-level modules that send predictions to lower levels where they are compared against the actual sensorial input by Prediction

Error (PE) units. If the prediction and the actual input do not match, PE units send this information to higher-level units so that the internal representation of the world can be updated. Higher area units receive input only from Prediction Error units (black line in FIGURE 4), which decreases the amount of sensory data that is further passed, limiting such information only to the parts that have discrepancies, or parts that are not correctly represented in the internal model.

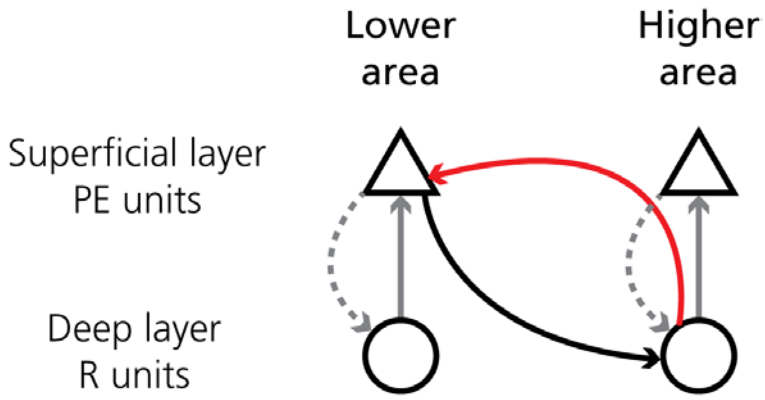

Figure 4 Schematic representation of information passing between units in the predictive coding theory (Based in Schröger et al., 2015 and Friston, 2005).

Recapping, the Ideomotor Theory asserted that with action execution, the predicted effects of an action are internally represented. The comparison of information regarding the action and its expected consequences would allow the motor system to evaluate the action by comparing the predicted with the actual effects. Predictive systems, thus, are in charge of evaluating the amount of difference between actual and expected inputs. When the amount of difference exceeds certain threshold, the system further passes this *error* information to any systems that need the evaluation information in order to update the model of the outside world. Also, such comparison allows the system to have an evaluation of the current mapping between an action and its effects.

4 THE VISUAL SYSTEM

The visual system is of special interest because of its high complexity. It is presented in a whole new chapter because of its high relevance to the current work, the theories and studies mentioned before. Its complexity and its extension may be the reason why it is also the most studied sensory system. The visual system provides us with the richest information about the outside world. With a simple glimpse at an object, the visual system is able to gather and decode its position, colour, shape, texture, direction of motion and speed. The theories of action control (2 THEORIES OF ACTION CONTROL {p. 6}) and the models in charge of prediction (3 ACTION PREDICTION: FORWARD AND INVERSE MODELS {p. 13}) that we presented earlier, apply and fit eye actions (movements) in a very similar way. But how does the impression of seeing emerges?

Our impression of *seeing* seems to us as a stable and continuous process, even though the actual retinal input is constantly changing, either interrupted by blinks, smeared by eye and head movements, or shifted by saccades. Saccades are fast eye movements that locate the fovea on objects of interest, and occur on average 3–4 times per second. The fovea is the central region of the retina and comprises around 1 degree of the visual field (VF) with approximately 1.5 mm in diameter. Despite its small area, the fovea has the highest density of photoreceptors, which makes it the zone with the finest resolution on the retina. Interestingly, the fovea entails 1% of the surface of the retina but its neural projections represent up to 50% of the visual cortex input (Kandel et al., 2000). Eye movements and visual attention are closely intertwined. It is known that perception is better at attended locations of the visual field compared to unattended locations. This has been measured either by subjects RT to an object unexpectedly appearing in the VF, or by their ability to perceive a just-noticeable stimulus. Saccades are, in consequence, of great importance for attention, perception and action control.

In this section we will describe the basic brain circuitry involved in seeing and in eye movements, the pathways taken by retinal information in order to reach high-level areas of the brain, and how is visual stability achieved. That will lead to trans-saccadic correspondence, or how the scene before a saccade is matched to the scene after a saccade, without having the impression that the world moved. Finally, we will end up seeing saccades as actions in order to integrate the Ideomotor Theory into the visual system. This will allow us to assume the existence of a Forward Model that deals with visual stability and trans-saccadic correspondence by making predictions about objects' location (in time and space) and features.

4.1 Information flow throughout the Visual System

Information within the visual system is passed onto different brain regions that interact and give rise to the impression of seeing. Interestingly, what we call *seeing* is a complex concatenation of multiple and parallel information processed by the visual system, and it all begins in the eye.

4.1.1 The Retina

The eye is filled with a fluid called *vitreous humour*. At the centre of the eye and enclosed by the cornea is the pupil, which permits the light to enter to create an image on the retina, the back-most part of the eyeball. The *retina* is a specialised region that contains neurons that are sensitive to light and transmit visual signals to specific centres of the brain. The process of transforming light into neural signals is called *phototransduction*.

The pupil and the lens are responsible for the refraction[6] necessary for a neat focused image on the retina. The shape of the lens can be modified when viewing close or distant objects, a process termed *accommodation*. When viewing close objects, the lens thickens and becomes rounder, allowing it to have the most refractive power. When viewing distant objects, the lens is thin and flat, and loses its refractive power. Accommodation is possible thanks to the tension that the ciliary muscles exert on the lens (see FIGURE 5 {p. 21}).

The *fundus* is the surface of the retina. Located near the centre of the retina is the *macula lutea*, a 3 mm circular region containing a yellow pigment. The macula is the region of the retina which has high visual resolution, and therefore has the ability to provide detailed information about objects and scenes. The *fovea* is the part of the macula which has the highest acuity and it is located at its centre as a small depression.

[6] The change of direction of a ray of light.

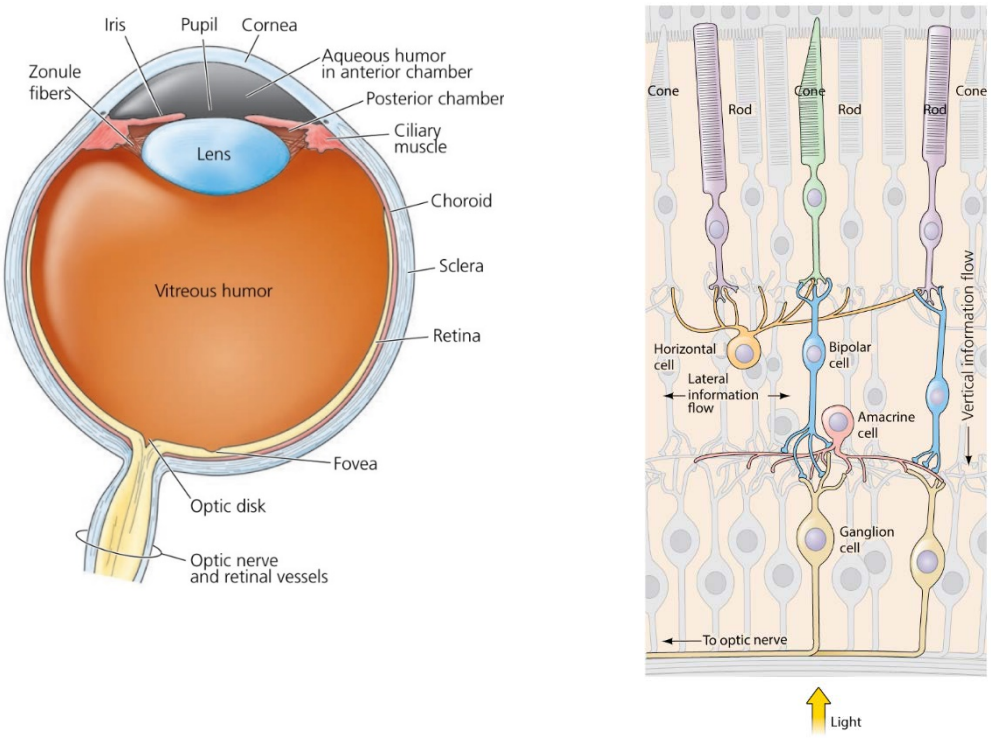

Figure 5 (left) Anatomy of the eye. (right) Internal structure of the retina.

Even though the retina has a peripheral location (in the eye, which is "outside" the head), it is part of the Central Nervous System (CNS). The retina is formed by five types of neurons, all of which have an important and entangled part in vision. The five neuron types are: photoreceptors (which in turn can be of two types: rods or cones), bipolar cells, ganglion cells, horizontal cells, and amacrine cells. Despite the indispensable role of bipolar, horizontal and amacrine cells, we will only focus on and describe photoreceptors and ganglion cells. These five types of cells are organised in layers, forming a three-neuron chain: a photoreceptor cell connecting to a bipolar cell which in turn connects to a ganglion cell. This comprises the most direct path of information form photoreceptors to the optic nerve. Horizontal cells, as their name imply, connect to photoreceptors horizontally, and help in the regulation and integration of photoreceptors' output. The optic nerve carries information about retinal stimulation issued by ganglion neurons to the CNS (Masland, 2012).

Rods photoreceptors have very low spatial resolution, but are very sensitive to light; their lack of detail is compensated by their high sensitivity. There are around one hundred million rod cells in the human retina; and they compose the second most numerous set of cells in the human body (Masland, 2012), after red blood cells (Bianconi et al., 2013). Responsible for colour vision, cone photoreceptors

have a very high spatial resolution, but are barely sensitive to light; their compromise is acuity at the expense of sensitivity. There are around 4.5 million cones in the human retina. Rods can respond to a single photon of light, while a cone would require more than a hundred photons to create a response similar to that of a rod (Sakmann and Creutzfeldt, 1969). In the fovea, however, cone density increases intensely, until reaching the central 300 μm where there are only cones. As cone density decreases with increasing eccentricity, the amount of convergence of ganglion cells increases, which results in a severely reduced acuity in peripheral vision. At about 6° of eccentricity, acuity is reduced by 75%. Given that high resolution vision is limited to a small area of the retina, humans (and many animals) perform constant eye (and head) movements in order to direct the fovea of both eyes to regions of interest; such "action" is termed *foveation*.

Ganglion cells receive indirect input from rods and cones alike, but this depends on the level of illumination. Ganglion cells convey information about image-like and non-image-like forms from the retina to other brain regions. Stephen Kuffler studied individual ganglion cell responses to small flashes of light using a cat's retina. He found that ganglion cells responded to stimulation of a small circular patch of the retina, which defines the cell's receptive field (Kuffler, 1953). These different responses made him distinguish two types of ganglion cells, *ON-centre* and *OFF-centre*. Flashing the light in the receptive field of an ON-centre neuron produced a burst in the cell's responses. When the same flash was presented to an OFF-centre cell, the response rate was reduced, and when the flash was extinguished the cell's response rate burst. Kuffler concluded that ON-centre cells rise their discharge rate to increases in luminance in the receptive field centre, while OFF-centre cells increase their discharge rate to luminance decrements in the receptive field centre. These findings imply that increments and decrements in luminance are transferred separately and independently into the brain by ganglion cells.

4.1.1 The Primary Visual Pathway

Ganglion cells' efferent information follows a path known as the *primary visual pathway*. Ganglion cells' axons exit the retina through the optic disk and all axons together form the optic nerve. These axons continue to the optic chiasm, where around 60% of them cross it (contralaterally), and the remaining 40% continue towards the thalamus and mid-brain regions (ipsilaterally). Regardless of ganglion cells' origin, passed the optic chiasm all axons together form the *optic tract*, which transfers information from both eyes.

Before continuing into cortical areas, retinal information passes through another important target: the *pretectum*; a region that regulates pupillary light reflex. Throughout the visual pathway, this is

the first module in which ganglion-cell information can be "sent back" to retinal, low-level areas. The pretectum is also the first unit in the primary visual pathway in which retinal information is decoded. Neurons in the pretectum deal with luminance, and determine if the entering amount of light passing onto the retina should be adjusted by contracting or detracting the pupil via the *pupillary constrictor muscle*[7]. The pretectum's task is to issue a motor command to contract or detract the pupil with respect to the level of luminance that —few milliseconds ago— passed through the retinas. This maximises acuity and avoids extreme values of luminance, which could seriously impair visual acuity or even temporarily impair vision.

The axons of ganglion cells connect into several structures of the diencephalon and the mid-brain. In the diencephalon the main target is the (dorso) Lateral Geniculate Nucleus of the thalamus (d LGN), which in turn makes synapses with the primary visual cortex (V1; or striate cortex; Brodmann's area 17). The LGN is a laminated structure that contains six layers of different neuron kinds all of which receive input from ganglion cells. The first two layers of the LGN are composed of *magno* cells. Magno cells receive input from parasol ganglion cells (Perry et al., 1984; Rodieck and Watanabe, 1993), and are big neurons with large receptive fields which particularly respond to luminance and do not respond to colour (Callaway, 2005). Layers 3 through 6 are constituted by *parvo* cells: small cells with short receptive fields which receive input from midget ganglion cells. Parvo cells respond to colour and provide information about motion and depth (Atkinson, 1992). In between each of the layers exists another layer called *konio*cellular layer, which carries cone low-frequency information to the superficial layers of the visual cortex (Wallace et al., 2016). Based on these types of neurons, a *magno*- and a *parvo*- pathways have been suggested to carry different types of segregated information from each eye. A recent article suggests that the LGN forms a subcortical region of binocular integration composed of koniocellular layer cells, a property until recent years only found in rodents (Zeater et al., 2015).

An important discovery about the information processed along the primary visual pathway was made by Hubel and Wiesel (1962). By means of microelectrode recordings in cats and monkeys, they found that neurons in the Lateral Geniculate Nucleus (LGN) behave much like retinal ganglion cells, that is, with a centre-surround receptive field that responds to luminance changes. Neurons in the cortex, however, did not respond to spots of light, but to bars and edges; and peculiarly, the firing rate depended on the orientation of the bars. This showed that the response of these cortical neurons

[7] In combination with the Edinger–Westphal nucleus and the ciliary ganglion.

was tuned to orientation and based in luminance (Hubel and Wiesel, 1977, 1962) with all orientations being represented equally within that cortical region.

After passing the LGN, ganglion cell axons reach other important targets, like the superior colliculus (SC). The superior colliculus is in charge of coordinating head and eye movements to visual targets, among other things. This spatial coordination can occur even in the absence of the cerebral cortex (Sprague, 1996). The ability to keep its function in the lack of cortical input indicates that spatial processing is already possible at the SC using purely ganglion-cell information. Space emerges in the superior colliculus where other low-level information may be combined (Klier et al., 2001).

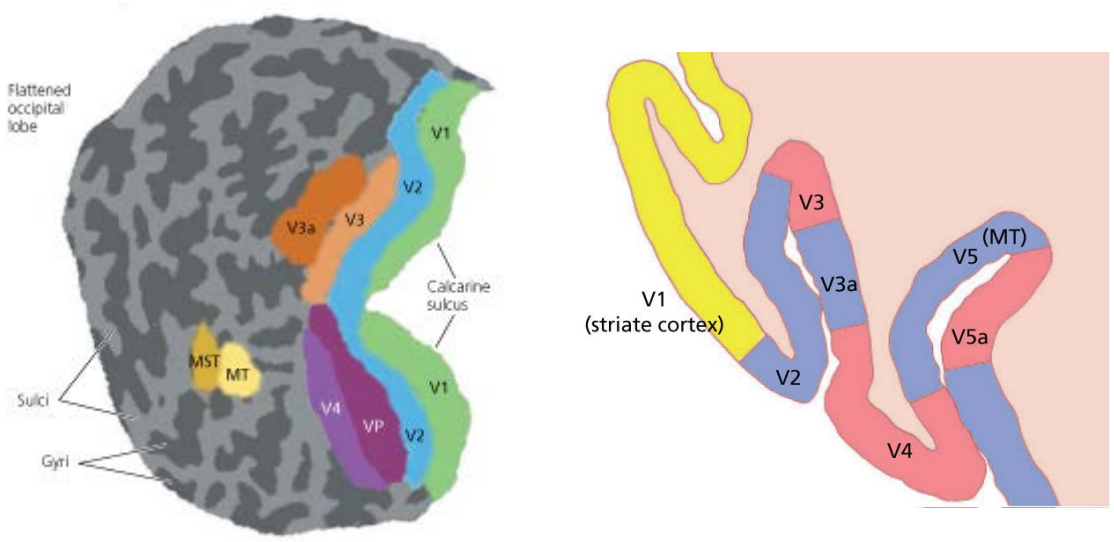

Figure 6 (left) Flattened image of a dorsal view of the cortex in which visual cortex expands. (right) Location of the primary visual cortex (V1) and the secondary visual cortex: V2, V3, V3a, V4, V5 & V5a.

Information eventually reaches V1, where neurons very precisely encode retinotopically the visual scene. Organised in six layers, the primary visual cortex segregates populations of neurons that have different properties and connection patterns. This organisation resolves into sparse clusters of neurons which respond to edges and orientations of each point of the visual field; this way, all possible orientations (0°-180°) are represented in a distributed fashion by different neighbourhoods of neurons. Neurons in the superficial layers of V1 project to extra-striate cortical areas, while neurons in deeper layers send axons to sub-cortical regions, including the LGN and the SC.

Then, the primary visual cortex sends synaptic information into the secondary visual cortex (V2) which serves as a feedback relay to V3, V4 and V5 with V1. Distributed along and across in both hemispheres, V2 contains a map of the visual field. V2 neurons, as V1's, respond to simple visual

features like spatial frequency, colours or orientation; but in more complex arrangements than V1. The deeper the layer in the secondary visual cortex, the more complex visual stimuli it will likely respond to. V2 is known to contribute to binocular disparity (Hegdé and Essen, 2000; von der Heydt et al., 2000) and to more complex combination of orientations (Anzai et al., 2007). A study conducted on rats showed that cells in the sixth layer of V2 contribute to long-term visual memory (López-Aranda et al., 2009). The third visual cortex (V3) is subdivided into several sub-units (V3a, V3b, VP). V3's function is still under debate because its subdivisions remain also under debate. In general, V3 does not project strongly into V1, and instead creates dense connections into the inferior temporal cortex, for example. Among the properties of V3 are its full representation of the visual field, and its response to global (Braddick et al., 2001) and coherent motion (Lui et al., 2006).

Eventually, information reaches ulterior cortical areas of visual cortex, which comprises brain regions of the occipital, temporal and parietal lobes. FIGURE 6 {p. 24} shows the extension of the primary visual pathway into the above-mentioned lobes in a "flattened" brain in which the main areas of the visual cortex can be seen. The temporal lobe is mostly involved in object recognition, and the parietal lobe with motion and spatial location, as will be detailed in 4.4 VISUAL OBJECT RECOGNITION {p.30}. Information ultimately reaches V4, to be further transferred to the medial superior temporal area (MST) and the middle temporal area (MT/V5).

V4 receives input from V2 and from foveal V1, sending output to the inferior temporal gyrus. V4 exhibits low connectivity to V5, and like V2 neurons, is tuned to orientation, spatial frequency, but other "intermediate" complex visual properties have been found to be treated in V4 like shape & temporal frequency (Basole et al., 2003; Brincat and Connor, 2004; Michel et al., 2013). It has also been found to respond to geometric shapes and forms recognition, as well as being involved in long-term plasticity (Schmid et al., 2013) and in object saliency gated by FEF input connections (Moore and Armstrong, 2003). MT/V5 receives input from koniocellular layers of the LGN (Sincich et al., 2004) and from cortical regions of V1, V2 & V3 (Felleman and Van Essen, 1991; Ungerleider and Desimone, 1986) . Cortical neurons in the middle temporal visual area (V5/MT) are tuned to the direction of a moving stimulus, and fire faster when it moves in a particular direction. For instance, MT neurons have been found to be direction-selective (Born and Bradley, 2005). Some neurons in visual area V4, on the other hand, respond selectively to the colour of a stimulus, regardless of motion or shape (Zeki, 1980).

4.2 Actions within the Visual System

Motor behaviour is guided by a series of brain modules from the primary motor and the pre-motor cortices down to a specific muscle. The execution of actions does not finish there, as the system needs to update its representation of the outside world after the action has been executed. As we saw in 1 NEURAL BASES OF MOTOR ACTION {p. 1}, once a movement has ended, proprioceptive signals are sent back encoding the immediate movement consequences. As depicted in FIGURE 1 {p. 2}, the just-executed action is initially guided by visual cortices, where space —among other things— is incorporated into the "parameters" that will guide the next movement. The way in which visual information is integrated into the system, is by considering the current available information in visually-concerned areas (V1, V2, MT, etc.), to refine or change the following action based on the previous one.

Visual information can be voluntarily updated by scanning areas of interest in the visual scene. Such exploration involves movements executed by the eyes. Eye movements are controlled by three pairs of antagonistic muscles, termed *extraocular muscles*. The *lateral* and *medial rectus muscles* which control horizontal eye movements; the *superior* and *inferior rectus muscles*; and the *superior* and *inferior oblique muscles*. Vertical eye movements are controlled jointly by the rectus and the oblique muscles. Different combinations of these muscles' contractions or detractions can create different eye movements.

In essence, extraocular muscles follow the same synaptic flow as a limb movement, except for some omissions that will be clarified below. There are two basic types of actions that extraocular muscles can create. The first, is to change the direction of gaze. Gaze-change involve *saccades*, *smooth pursuit* movements and *vergence* movements. The second are stabilizing actions which involve *vestibulo-ocular* and *optokinetic* movements. By far, saccades have been the most studied eye action, not only because of their ubiquity, but also because of the visual changes they evoke. Thus, for brevity and for the aim of this thesis, we will describe saccades thoroughly and other actions briefly.

Saccades have as function to orient the foveae toward areas of interest for closer examination. The selection of such interesting regions in the visual scene is guided by non-foveal/peripheral areas, which implies that these peripheral areas have sufficient spatial resolution to precisely locate a saccade's target. It takes around 200 ms for the eye to move after a saccade command has been issued. During this time-lapse the target's position is computed, and the difference between the current and the intended retinal positions is converted into a *motor command* that will *minimise the distance* until the target position is located on the fovea. Such motor command activates the

extraocular muscles in order to move the eye into a specific direction and distance. During the eye movement, changes in the target's position or changes to its visual features are very difficult to detect. Although saccade latencies are small, ranging from 15 to 100 ms, once the motor command is issued, it cannot be modified on the fly. Therefore, if during this latency the target changed its position and this change is detected upon saccade end, a supplementary saccade is needed to correctly foveate the displaced target.

Smooth pursuit movements are slow-tracking eye movements, whose function is to keep a moving object foveated. The impression of movement can originate in an object actually moving in the environment, or in a head or body movement. Vergence movements align the fovea of each eye with targets located at different distances from the observer. Vestibulo-ocular movements and optokinetic eye movements stabilise gaze relative to the external world, and compensate for head movements.

The saccadic system needs to resolve two spatial problems before issuing the motor command: first, to determine the distance to the target, that is, how far away it is from current foveal position; second, to determine in which direction the target is located. These two pieces of information conform the basic *vector* of a saccade's motor command: amplitude and direction.

Amplitude is eventually encoded by the duration of neuronal activity in lower motor neurons of the oculomotor nuclei, and correlates with the firing patterns in the abducens neurons (Fuchs and Luschei, 1970). Direction is determined by local circuit neurons in the gaze centres (located in the reticular formation) that control eye movements in the horizontal and vertical axes (using lateral, medial, superior and inferior rectus muscles). Horizontal saccades are generated by the paramedian pontine reticular formation (PPRF, the horizontal gaze centre), located in the pons, while vertical saccades are generated by the rostral interstitial nucleus (vertical gaze centre) located in the rostral part of the midbrain. By combining the activation of the horizontal and vertical gaze centres, saccades can indeed be made to any direction and magnitude with respect to the fovea.

4.3 Saccadic Control: SC and FEF

Saccades, however, must be precise in order to minimise the workload and to correctly update the visual scene. Correctly updating the visual scene should, in principle, minimize the number of misplaced saccades. The superior colliculus (SC) and cortical areas of the frontal and the parietal lobes project into the gaze centres and thus are important for the initiation and the accurate execution of saccades. One of cortical regions that project into gaze centres is located in the frontal lobe region and is known as the Frontal Eye Field (FEF); conveniently located close to the motor cortex.

Neurons in the SC and the FEF both contain a topographical map of eye movement vectors, and increase their firing rate just before the saccade initiates. These topographical maps activate certain sites of these two modules to elicit a saccadic eye movement with a certain amplitude and direction. Stimulating a specific site of the Frontal Eye Field (or the SC) would provoke a specific saccadic movement with a specific direction and amplitude, which is independent of the position of the eye in the orbit (Foulsham, 2015; Masland, 2012). The SC and the FEF both have neurons that can be activated by visual stimuli. This is important because specific retinal information can be transferred to these two regions to commence an eye movement.

The topographical organisation of superior colliculus neurons instantiates a visual map of space. Neurons in a particular region of the SC are activated by visual stimuli in a limited and specific portion of visual space (Schiller and Stryker, 1972). The relevance of the SC is of major importance, as it also responds to auditory and somatosensory stimuli. Spatial location is consequently mapped in line with the visual maps in the SC. By integrating sensory and motor maps into the topographic disposition, the superior colliculus allows the transformation of sensory input into movements which involve spatial information.

The initial role of the superior colliculus was thought to be only the control of eye movements. However, recent experiments using finer stimulation and recordings found that the SC is rather involved in the production of gaze shifts by combining head and eye movements. The superior colliculus, thus, plays a crucial role in directing behaviour and movement *intention* toward specific objects in the animal's space, while object spatial recognition is not a central part of its roles (Sprague, 1996).

The notion that the SC's function involves a motor map rather than a space map was further developed by David Sparks. He showed that retinal information is not sufficient for localising saccade targets (Sparks et al., 1987; Sparks and Mays, 1983). Sparks trained monkeys to perform a voluntary saccade to a target after a cue was on. Right before the saccade was initiated, neurons in the deep layers of the superior colliculus were stimulated. This stimulation induced the saccade away from the fixation dot. Eye movements were recorded to check if the alteration of the saccade had an influence in the *intended* saccade's movement vector (amplitude and direction). If saccade vectors were created solely by means of retinal information, the monkey would make a saccade towards the previously trained saccadic cue. In such case, the saccade should miss the target by the amount of deviation that the stimulation had induced. He found, nevertheless, that monkeys compensated the deviation by performing a correcting saccade towards the true saccadic target. The compensatory saccade used

information about the location of the retinal image and information about the position of the eyes in the orbit, and not only retinal information.

In general, it can be said that the FEF and the SC both control eye movements, however FEF's control has been associated with visual attention (Schall, 2004). As the SC, FEF stimulation elicits a saccadic eye movement represented in retinotopic coordinate system (Bruce et al., 1985). The FEF has been also associated with smooth pursuit movements (Mustari et al., 2009; Tian and Lynch, 1996), and can control eye movements independently of the SC. This can happen because efferent signals from the FEF project directly into the SC (which also projects directly into the PPRF), but also directly into the PPRF. This result was discovered using monkeys with induced lesions in the SC (Schiller et al., 1980). Saccades' latencies and accuracies were diminished, yet they still occurred. This result suggested that the SC and the FEF both provide independent pathways for controlling saccades. Eye movements, like other movements, are also under the control of basal ganglia and the cerebellum.

4.3.1 The Corollary Discharge

Given the high similarity with the human visual system, numerous neural and physiological studies rely on recordings of monkeys. Muscles that control saccades obtain their motor signals from pons and mesencephalic neurons. Provided that saccades can be both reflexive and voluntary, these muscles' activation can also come from a network of cortical and subcortical areas that control the saccadic system via the Superior Colliculus (SC). The main brain regions involved in the generation of saccades in monkeys are the brainstem, the SC, the FEF and the Lateral Intraparietal Area (LIP) (Andersen, 1989; Komatsu and Wurtz, 1989). Additionally, in humans, it has been shown that frontal and parietal cortices are involved in the spatial selection and preparation of saccades, particularly the FEF, the Supplementary Eye Field (SEF) and the Intraparietal Sulcus (IPS) (Curtis and Connolly, 2008). Signals from the eye have been considered critical on granting visual stability, because changes in the image of the retina can be a consequence of the eye moving towards a new position, or consequence of an external object moving. Visual change is termed *reafference* when it's self-generated, and *exafference* when it's not (no eye movement). There are three basic signals that contribute to visual stability (Wurtz, 2008). When the eye moves, a reafferent signal from the retina indicates the eye movement. Once eye muscles have contracted, proprioceptors signals are triggered announcing that the eye has moved. Proprioceptor signals have been shown not to contribute much to visual stability as they do not contain substantial information regarding the eye's new position (Lewis et al., 2001; Sparks et al., 1987; Wang et al., 2007). The third signal (an extraretinal one) is a copy (or "corollary") of the motor command that enabled the eye movement; probably the most

relevant of the three, this signal is often referred as Corollary Discharge (CD) or Efference Copy (EC) (Sperry, 1950; von Holst and Mittelstaedt, 1950). Basically, at the same time that sensorimotor information is issued to command muscles to produce an eye movement, a *copy* of such information is also sent to other brain regions to notify about the movement's characteristics and the eye's future position (the saccade's vector). The superior colliculus (SC), the medial dorsal nucleus of the thalamus (MD) and the Front Eye Field (FEF) are among the regions believed to be on the path of the Corollary Discharge (Lynch et al., 1994).

The systems that control eye movements follow the same principles as systems that guide movements in other parts of the body. The difference lies in the location of lower-motor neurons. For hand movements, for example, these neurons are located in the spinal cord, while for eye movements such neurons are located in the reticular formation of the brainstem.

Once an object's visual information has been segregated by primary and secondary visual cortices, the visual system continues to treat the information in order to identify an object of interest, both by its visual features and by its position in space.

4.4 Visual Object Recognition

So far, we saw that information coming from the retina is separated into different processing streams among different brain regions. Much of this information goes to the primary visual cortex (V1) to be further transferred in the visual cortex and beyond onto the parietal and temporal cortices, both involved in different aspects of visual object recognition.

Based on many studies, extra-striate cortical areas have been organised into two separate systems, or pathways. The *ventral system* includes area V4 ending in the inferior part of the temporal lobe. This system is known for high-resolution form-vision and object recognition. The *dorsal stream* includes the middle temporal area (MT/V5) in the striate cortex, ending in the parietal lobe. This system is responsible for the spatial aspects of vision, such as motion and positional relationships between objects in the visual scene. As I will expose below, these two streams are of sum importance to the present research.

The ventral stream is specialised in object perception and recognition, that is, in determining *what* we are seeing. The dorsal stream specialises in analysing the spatial configuration of objects in space, that is, determining *where* the object is. Colloquially, these two pathways have been termed the *what stream* and the *where stream*, respectively (see FIGURE 7 {p. 32}). The *where* and *what* information

will lead to identify an object and to create its internal representation. As described later (in 4.6 VISUAL STABILITY AND TRANSSACCADIC CORRESPONDENCE {p. 36}), Object File Theory proposes a mechanism for object identity which includes attention-enhanced visual- and spatial-features of a stimulus.

Walter Pohl discovered that monkeys with bilateral lesions to the temporal lobe (ventral "what" stream) were very bad at discriminating different shapes (Pohl, 1973). When animals suffered a parietal lobe lesion (dorsal "where" stream), they had trouble localising an object with respect to other objects; discrimination between two similar objects, however, remained unaffected. Mishkin and Ungerleider formalised previous findings and suggested that the ventral and the dorsal streams were in charge of processing different types of information in object recognition (Mishkin et al., 1983). Almost a decade later Goodale and Milner (1992) suggested that the ventral "what" stream is mainly involved in the perceptual identification of objects, while the dorsal "where" stream is in charge of sensorimotor transformations for visually guided actions directed at objects. Further studies found that the ventral and the dorsal systems both received input from parvocellular and magnocellular layers, information which remained relatively separated until V1 or V2. Information segregation between parvo and magno cells was not as exclusive as originally thought, because V4 cells as well as V5/MT cells were found to receive information from magno and parvo cells (Livingstone and Hubel, 1988; Maunsell et al., 1990; Schiller and Logothetis, 1990). The parvocellular layer appears to be selective for form and colour, while the magnocellular seems to be in charge of depth and movement. After Livingstone and Hubel's article, the dorsal pathway which was previously associated with spatial aspects of object recognition, was renamed as the "how" stream which; it would be in charge of visuomotor control rather than only spatial representations (Freud et al., 2016). I will continue to use the terms 'what' and 'where' streams/systems, regardless of the goal of the information carried through the visual cortex.

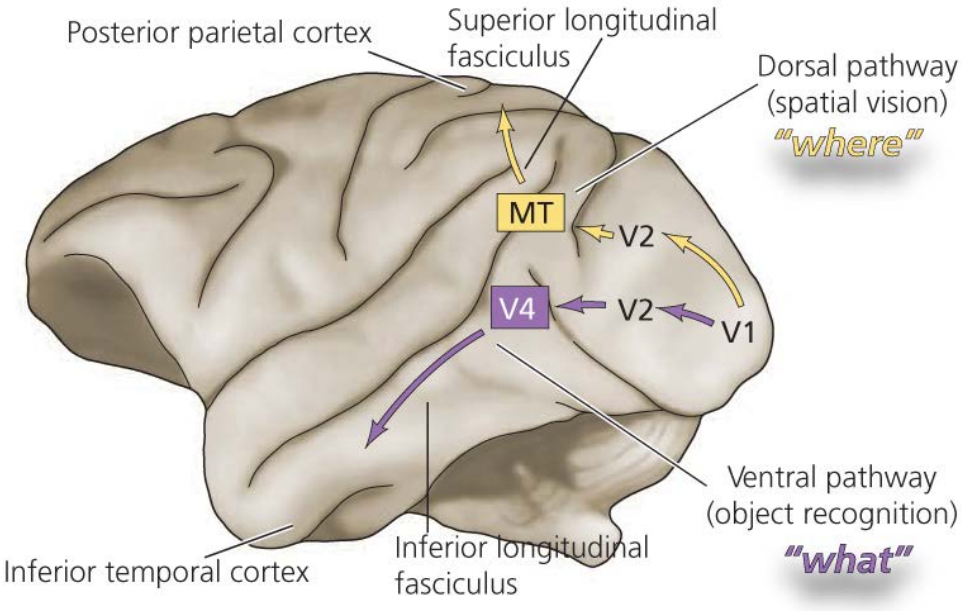

Figure 7 Dorsal and ventral pathways (in this example of the macaque brain) carry information to extra-striate areas to perform different tasks. The dorsal stream carries information in order to recognise the spatial properties an object in space. The ventral stream carries information in order to recognise the surface features of an object. Both streams work independently of each other. Based on Gazzaniga and Ivry, 2013.

The receptive fields of neurons from the temporal and the parietal lobes are both large, suggesting that both areas still treat low-level features of retinal input. However, there are many differences in the way they behave and thus in the things they encode and decode. Dorsal stream cells, for example, respond similarly to many different stimuli (Robinson et al., 1978). Around 40% of dorsal stream neurons have receptive fields near the fovea, while the remaining 60% exclude the foveal region[8]. This eccentrically arranged proportion of neurons gives the dorsal system an appropriate way for detecting the presence and the location of an object that suddenly appears inside the visual field. Parvo and magno cells exhibit a somehow similar aspect of laterality: layers receive contralateral- and ipsilateral input from the eyes. This knowledge was based on findings regarding the binocular receptive fields in mice and cats (Howarth et al., 2014; Sanderson et al., 1971) especially for koniocellular layers of the LGN, which revealed that binocular vision was decoded before the striate cortex (Wallace et al., 2016). Neurons of the ventral system have all receptive fields that include the

[8] See "ON THE 40%–60% RATIO" {p. 119} in the general DISCUSSION.

fovea; this foveal-exclusive property indicates that in order to find out "what" an object is, one must first locate it, to then foveate it in order to grasp as most details as possible.

Remarkably, the *what stream* creates afferent connections on the temporal lobe —a region in charge of visual memory—, which in turn contacts the hippocampus —a region explicitly dedicated to memory. The *where stream* instead, synapses through the superior longitudinal fasciculus (SLF), a tract that interconnects the occipital, the temporal and the frontal lobes (Makris et al., 2005), among others. SFL I (one of the three functional divisions of the SFL) transfers occipital information to the superior parietal cortex (Brodmann's areas 5 & 7) which encodes locations of body parts in a body-centric coordinate system.

Neurons in the posterior part of the temporal lobe ("what" stream), have preference for simple features such as edges. In this sense, the perception of an object depends on the decomposition of the shape using features like colour or texture. Amazingly, when these features are not present in the visual stimulus, object recognition is not impaired. Simple features as lines can be gathered as oriented edges, then grouped into parts, and finally ensembled into "objects". Other neurons have preference for more complex features. This suggests that, as information flows into these object recognition areas, features are grouped together into more complex percepts.

The following figure depicts this idea by showing the regions involved in the ventral and dorsal pathways in the macaque brain (FIGURE 8 {p.34}). Once retinal information arrives in the occipital cortex, shapes are decoded. Information then splits and exits the occipital cortex leading into the temporal cortex where shapes are ensembled to decode a more complex object, like a face.

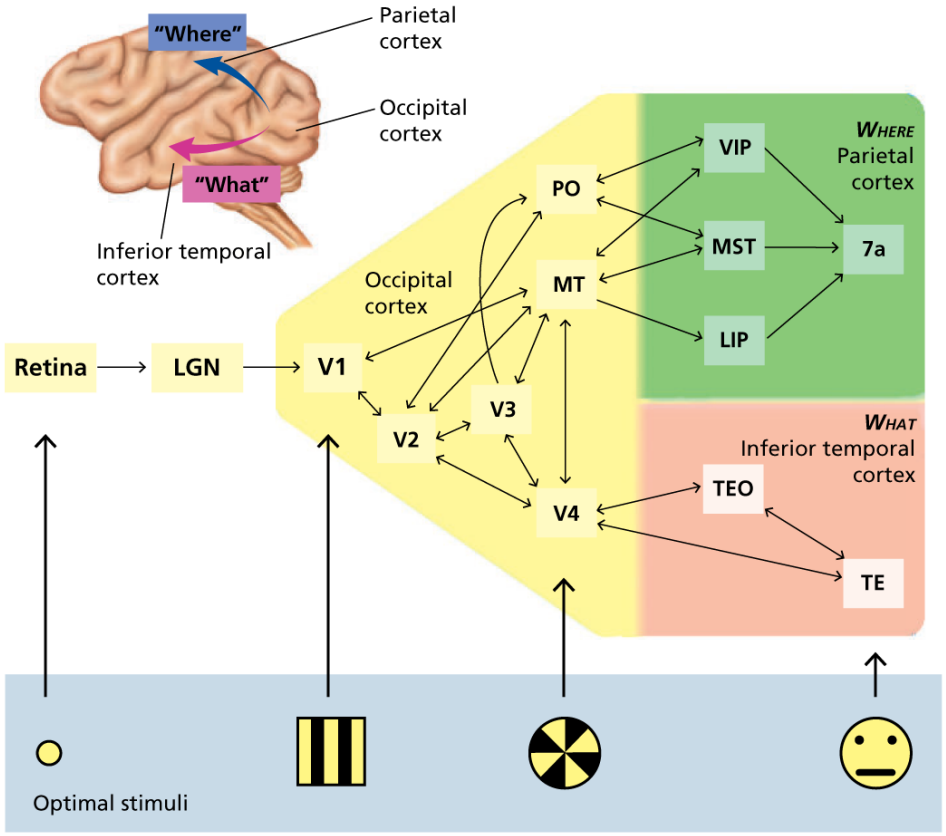

Figure 8 Visual processing of stimuli, taking the ventral or dorsal streams after leaving the occipital cortex in the macaque brain. Each of the areas shown contains its own retinotopic map.

In a series of studies, neuronal responses of regions in the inferior temporal (IT) cortex were recorded. Such studies revealed that these neurons selectively respond to complex shapes, like hands (Gross et al., 1972) or faces (Desimone, 1991; Desimone et al., 1984), and different stimulus sizes (Ito et al., 1995). Interestingly, IT neurons always include the fovea, a property stressed above in the recognition of details. Confirming and complementing previous findings, Sheinberg and Logothetis found the inferior temporal cortex (IT) to be also involved in object recognition. They proposed that the inferior temporal cortex was linked to the *percept* of the stimulus, rather than to the low-level features (e.g. shape) of the stimulus (Sheinberg and Logothetis, 1997). Trying to determine which mental operations are used when viewing familiar, novel and random shapes, Kanwhisher and colleagues found an increased BOLD signal in the lateral occipital cortex (LOC) when viewing familiar and novel objects, as opposed to random non-object stimuli (Kanwisher et al., 1997). This and other studies suggest that shape is critical for object recognition, being shape the most relevant feature treated by the lateral occipital cortex (LOC) (Martin, 2007).

4.4.1 Comparison between the Ventral and Dorsal streams

The following table briefly displays the main differences between the ventral "what" and dorsal "where" systems (Based on Norman, 2002).

	Ventral system (what)	**Dorsal system (where)**
Function	Recognition & identification	Visually guided behaviour
Sensitivity	High spatial frequencies	High temporal frequencies
Memory	Long-term stored representations	Short-term storage
Speed	Relatively slow	Relatively fast
Consciousness	High in general	Low in general
Reference frame	Allocentric or object-centred	Egocentric or viewer-centred
Visual input	Mainly foveal or parafoveal	Foveal and peripheral

Table 1 Comparison of characteristics between the ventral and the dorsal systems.

4.5 Divided Visual Attention

It is well known that orienting attention to a specific location of the visual field facilitates information processing, as compared to non-attended locations. Attention, seems to be a 'resource' that can have multiple regions in sensory space: divided attention. There exist three main models of spatial divided visual attention. The first uses the "spotlight" metaphor initially proposed by William James, stating that visual attention behaves like a spotlight which *illuminates* a single and contiguous region of the visual field (James, 1890). This model implicitly suggests that spatial attention is processed serially. The studies behind the spotlight theory have demonstrated that the deployment of attention can be stimulus-driven (bottom-up), or by means of voluntary control in a goal-driven manner (top-down) (Posner and Cohen, 1984; Posner and Petersen, 1990). It has also been found that the set of goals of an observer (the Attentional Control Settings: ACS) can override the stimulus-driven manipulation of attention (Folk and Remington, 2006; Pratt and Hommel, 2003). Imaging studies have shown that shifts of the attention spotlight can be observed in striate and extra-striate visual cortices, indicating that retinotopic representation of the attended regions exhibit increased

activation (Kastner et al., 1998). Shulman and colleagues suggested that the attentional spotlight could rapidly switch between multiple objects (Shulman et al., 1979), and thus a pseudo-parallel processing could be achieved in order to explain spatial divided attention. In the spotlight approach, in which divided attention can be processed serially by rapidly switching attention from one location to another, studies have found that a 100-500 ms time window is required to switch attention, depending of the type of task (Awh and Pashler, 2000; McMains and Somers, 2004; Reeves and Sperling, 1986).

The second model is colloquially called the "zoom lens". Originally published by Eriksen and St. James, this model proposes that the spotlight has a variable power (zooming) which allows the parallel processing of two or more objects at a cost of less resolution or attentional resources (Eriksen and St. James, 1986).

The third and more recent theory states that spatial attention can actually be *split* into several regions of space, while *neglecting* unimportant regions in between (Adamo et al., 2008; Awh and Pashler, 2000; Itthipuripat et al., 2013). Regardless of whether these theories complement each other or whether they are mutually exclusive, there is still an ongoing debate about how spatial divided attention is achieved (Jans et al., 2010).

4.6 Visual Stability and Transsaccadic Correspondence

One of the oldest questions about vision is how visual stability is accomplished, despite the persistent change of the visual scene. Over the past decades, a lot of evidence has been accumulated trying to explain this phenomenon. Assuming that the visual scene is blurred by the speed of saccadic eye movements, there was a lot of intrigue as to why the blur does not reach conscious awareness. Several theories have tried to solve this issue.

4.6.1 Remapping

According to the remapping point of view, a motor signal is combined with the retinal consequences of the saccade so as to establish a correspondence between two locations. The initial proposal was that the motor signal "cancelled out" the self-caused retinal signal, as if the motor signal were a kind of negative image of the retinal input. Although such a negative image-like signal has been shown to prevent self-stimulation in electric fish (Bell, 1981), in primates this is unlikely (Sommer & Wurtz, 2004). Later it was instead proposed that a copy of the saccade command (the CD, which includes

magnitude and direction) allows the activity evoked by a stimulus at a particular retinal location to be transferred or *remapped* to neurons whose classic receptive field (RF) is not at that location, but in whose RF the stimulus will fall after the saccade. In this way, remapping allows the past and present retinal locations of a visual object to be aligned; when remapping is predictive, it allows the current and future retinal locations of an object to be matched. Remapping has been shown to occur in neuronal and in behavioural studies (Duhamel et al., 1992; Rolfs et al., 2011; Sommer and Wurtz, 2006; Umeno and Goldberg, 1997; Walker et al., 1995).

The first innovative experiment to display remapping, showed that neurons in the parietal cortex of the monkey brain adjust their receptive field (the retinal response area of sensitivity) in anticipation of an upcoming eye movement (Duhamel et al., 1992). Monkeys were trained to do fixation and saccade tasks, during which activity from LIP neurons was recorded. During fixation, neurons of the LIP were stable and their RFs were encoding the stimulus at a given position. Right before and during the saccade, LIP neurons stopped responding to the stimulus relative to initial fixation, and began to respond to its future location (post-saccadic); that is, their RF reacted to location at which the stimulus (i.e. the saccadic cue) would be upon landing. This finding shed light into a possible link between the corollary discharge and visual stability. The change in spatial-sensitivity was called *remapping* or *spatial update*, and is also referred to as *shifting receptive fields* based on these cells' neural activity.

Based on Duhamel's work, the first hypothesis to explain visual stability stated that a *retinotopic map* of retinal input is updated with each saccade. The way these neurons increased their sensitivity to future stimuli, implied that they must receive information about the amplitude and direction of the saccade before the saccade began, because their future-RF's sensitivity was pre-located at the site where the it would be after a saccade. It could be considered that after a saccade, visual stability is granted whenever the information in the current-RF and the future-RF is congruent; this matching of pre- and post-saccadic information is called *trans-saccadic correspondence*. When there is a self-generated (reafferent) change of the visual image, two increased activities take place to reflect the change: an increased sensitivity in the future-RF and another as the RF is moved onto the object. However, these neurons do not get activated when the object appears in the visual field (exafferent), which confirms the importance of the corollary discharge to distinguish between reafferent and exafferent visual changes. The "retinotopic hypothesis" of visual stability is supported by a) the proof that neurons in LIP, FEF (Umeno and Goldberg, 1997) and SC (Walker et al., 1995) show shifting receptive fields that could contribute to updating or remapping of the visual field with each saccade; b) by the identification of the CD through the SC–MD–FEF pathway that reaches these neurons (Lynch et al., 1994); and finally c) by the fact that the CD is necessary for the shift to occur in FEF (Sommer and Wurtz, 2006).

In physiological terms, remapping could occur because each neuron is potentially activated by a stimulus at all possible retinal locations; the connections that transfer activity would vary according to planned movements. Remapping could also result from lateral connections that transfer activity from one retinal location to the other within a given brain area. This lateral transfer of activity is more physiologically plausible than massive feed-forward connections (Casagrande & Kaas, 1994).

4.6.2 Spatiotopic Maps

Another hypothesis about visual stability theorises the existence of a high-level spatiotopic map in the brain, which would be updated as the retinal image changes. It is based on evidence showing that as visual information arrives at brain regions supporting visual perception (see VISUAL OBJECT RECOGNITION {p. 30}), the representation of the visual world is no longer in retinotopic but in spatiotopic coordinates[9]. Neuronal evidence for this lies on two types of neurons: *gain field neurons*, whose visual responses are modulated by the position of the eye in the orbit, and *real position neurons*, whose responses are independent of eye-orbital position (Galletti et al., 1993). A vast number of studies supports the idea that neurons in diverse brain regions could contribute to the update of a spatiotopic map (Andersen, 1989; Bridgeman, 1973; Turi and Burr, 2012).

For example, Anderson and Mountcastle reported that neurons in the Posterior Parietal Cortex (PPC) responded under attentive fixation to a stimulus, and that these responses were strongly influenced by the angle of gaze. Still retinotopically organised, PPC neurons showed a gain in their visual response modulated as the eye changed position to another region of the visual field (Andersen and Mountcastle, 1983). Other studies have also found that neurons in the Parieto-Occipital area (PO) keep a distributed spatial map stored in retinotopic (gaze relative) and spatiotopic coordinates ("real" position) (Galletti et al., 1995). Some other studies have claimed spatiotopic-dedicated networks (notably in V3A and V6) which keep detailed spatial information of a stimulus across saccades (Duhamel et al., 1997; Habak et al., 2004; Melcher, 2005). Finally, fMRI studies have confirmed the existence of head-centred networks particularly in the Lateral Occipital Complex (McKyton and Zohary, 2007) and in the Middle Temporal Cortex (d'Avossa et al., 2007). If it truly exists, a dedicated spatiotopic map remains to be found.

[9] In the sense that relevant information is not only "marked" by its retinal position, but by other reference frames.

4.6.3 Visual Short-Term Memory & Trans-Saccadic Memory

More recently Deubel and others hypothesised that visual stability across saccades is achieved by the use of *trans-saccadic memory*. They proposed that the features of the saccadic target and of objects immediately surrounding it are stored in memory. Trans-saccadic memory would be updated on each saccade, and its items retrieved to determine if, after the saccade, the target and its surroundings matched those stored before the saccade. If the comparison is a match, the assumption of a stable world is kept —thus granting perception with continuity. If they do not match, the stable world assumption is abandoned, and it is assumed that the target moved (Deubel et al., 2004, 2002). Many experiments have built upon this theory, and have shown that trans-saccadic memory can store around 4 items, similar to visual working memory (Irwin and Gordon, 1998; Irwin and Andrews, 1996; Prime et al., 2007; Vogel et al., 2001).

Object-File Theory

An alternative account of visual stability is that it relies on visual information about the pre- and post-saccadic scenes stored in Visual Short-Term Memory (VSTM) (Hollingworth et al., 2008). A rich body of literature has shown that the quality of information retained across a saccade is coarse and abstract (Henderson, 1994; O'Regan, 1992; Pollatsek et al., 1984), which limits the computational load. The *Object-File Theory of Trans-saccadic Memory*, introduced by Irwin proposes 4 levels of representation (Irwin, 1996): a) a *feature map*, (`FeatMap`) onto which object features of the visual field are represented encoding various features like colour or shape, with a coarse location information assumed to be encoded with each feature; b) a *master map of locations* (`MML`) which encodes precise spatial location of each feature; c) an *object-file* (`ObF`), which is a temporary representation that includes an object's identity, location and features updated whenever the object changes in features or space, and d) an *abstract long-term recognition network* (`LTRN`) which groups the former three. Perception emerges when attention is directed to a location in the visual field. This updates the `FeatMap` and the `MML` including features and precise spatial location to integrate an `ObF`. There are some limitations for the updating procedure. First, target proximity; given that attention precedes the eyes moving to a saccade target, `ObF`s of objects located near the saccade target are more likely to be created than those of further locations. Because of short-term memory limitations, there is also a limited number of around 3 or 4 `ObF`s that can be kept when saccades occur. The theory assumes that the links between the `FeatMap` and the `MML` are disrupted with saccades, such that precise spatial information about the locations of the features is lost. Irwin proposed that trans-saccadic memory consisted of Object-Files that are generated before a saccade and of "residual activation"

in the FeatMap and in the LTRN. He suggested that visual stability may be the opposite of what one might intuitively think:

« *the world appears stable (across saccades) not because a detailed memory exists, but because very little is remembered from one fixation to the next.* »[10]

The base argument for this was that fixation durations are ten times longer than saccade durations, indicating that in 90% of a person's perceptual experience a relatively stable image of the world is present on the retinas. Finally, according to this theory, instability across saccades is experienced only when one of the objects encoded in trans-saccadic memory is changed *during* the saccade; this, as we will see in the next section (4.7 SACCADIC SUPPRESSION OF DISPLACEMENT {p. 42}) is of great importance for the present study. The saccade target itself benefits from prioritized encoding into VSTM (Irwin and Gordon, 1998), and a variant of the object-file theory proposes that a search process attempts to locate the target object after every saccade, based on remembered visual features of the pre-saccadic object (Deubel et al., 2002; McConkie and Currie, 1996). However, attempts to measure such a search have not been successful (Eymond et al., 2016; Jonikaitis and Theeuwes, 2013). According to the VSTM account of visual stability, establishing object correspondence across a saccade depends on similarity between the visual features before and after the saccade.

4.6.4 Remapping Attention Pointers

The attention pointers theory is yet another theory derived from Deubel's trans-saccadic memory theory in order to explain trans-saccadic correspondence. The Remapping Attention Pointers approach is a modern instantiation of the remapping account of visual stability (Cavanagh et al., 2010). In this theory, the CD signal allows activity in a saccade map to be transferred from neurons whose RF corresponds to the pre-saccadic retinotopic location of the object to neurons whose RF corresponds to its post-saccadic location, and that activity feeds back to visual brain areas (V1–V4, MT), thus associating two retinotopically distinct locations into a single object representation and allowing features at the post-saccadic locations to be enhanced. Only attended objects are thought to be remapped (Joiner et al., 2011), which limits the amount of processed information. The attentional pointer account is similar to several other proposals of how a descending saccade signal determines transsaccadic correspondence and the content of attention (re-entry hypothesis: Hamker, 2003; visual attention model: Schneider, 1995).

[10] Irwin, 1996; p. 99

According to these proposals, the descending activity is blind with regard to features; it selectively activates neurons with a particular retinotopic RF, independently of their feature selectivity. According to the remapping account of visual stability, establishing object correspondence across a saccade depends on the spatial match between the actual post-saccadic target and its predicted location, which in turn depends on the CD. Therefore, the Remapping Attentional Pointers Theory explains trans-saccadic correspondence as predictive shifts of location *pointers* to attended targets (Cavanagh et al., 2010). This theory is sustained by the existence of shared brain regions which handle spatial attention and saccade control, mainly SC, LIP and FEF (Corbetta and Shulman, 2002). Neurons in these networks not only encode the location of a stimulus, but also (some of) its features aided by attention. Cavanagh and his colleagues argue that for trans-saccadic correspondence to work using attentional pointers, first a spatial shift of attention takes place to attend the saccadic target. Once this is done —and imminent to the saccadic movement—, information regarding the object's features and location is "transferred back" from peripheral vision to fovea, and then the saccade execution triggers the CD to shift receptive fields and finally execute the saccade. Evidence of attention remapping can be found in experiments of apparent motion (Rock and Ebenholtz, 1962; Verstraten et al., 2000), in experiments that test the limit of features that can correctly be accumulated across saccades showing that an attention-based mechanism is involved (Hayhoe et al., 1991) and cueing experiments (Golomb et al., 2008; Mathôt and Theeuwes, 2009). The remapping of visual features from pre-saccadic to post-saccadic target locations has also been found mainly for motion and aftereffects adaptations (Burr et al., 2007; Ezzati et al., 2008; McKyton and Zohary, 2007; Melcher, 2007, 2005). Nevertheless, other studies have not been able to replicate such findings (Gardner et al., 2008; Knapen et al., 2010, 2009; Wenderoth and Wiese, 2008).

4.6.5 Combination of Visual Stability Theories

To conclude, remapping cannot be the only mechanism behind visual stability, because CD signals are notoriously imperfect in space and time (Grüsser et al., 1987; Honda, 1991, 1989; Matin, 1982). Even small discrepancies between the corollary discharge and the actual saccade should seriously affect visual stability, but this virtually never happens.

These accounts of visual stability make different hypotheses as to the causal or consequential role of features in visual stability: according to the attention pointers account, features are matched across a saccade as a consequence of a transsaccadic correspondence between locations. According to the VSTM account, a transsaccadic correspondence between features causes locations to be matched.

In all likelihood, visual stability depends on some combination of remapping and visual memory. Indeed, the saccade target is prioritized, but objects close to the saccade target are recalled better than objects farther away, which means that retinal location information (and spatial information in general) still plays an important role (Irwin and Gordon, 1998). Moreover, in free viewing situations in which the target is not specified in advance, attention benefits correlate with the actual landing position of the saccade (Doré-Mazars et al., 2004), suggesting that transsaccadic correspondence relies on the specific metrics of the saccade and not only on the target's location (which can differ because of oculomotor noise; see Collins et al., 2009).

4.7 Saccadic Suppression of Displacement

Although at first glimpse the hypotheses described above could seem to deal with different concepts, they complement each other (Sun and Goldberg, 2016). Saccadic Suppression is a known phenomenon in which any changes of the visual scene *during* a saccade are discarded before reaching awareness. This term was coined in the 1970's by Bridgeman and colleagues, who showed that the *identification* of a target object is strongly *suppressed* during saccadic eye movements (Bridgeman et al., 1975). Given that saccades occur very often during waking life, and that many of them are not consciously made, it is often acknowledged that the main function of saccadic suppression is to maintain visual stability. Saccade durations can be as short as 15 ms, which leaves very little time for retinal information to reach high-level areas, let alone conscious awareness.

In the laboratory, visual stability is often probed with the *inflight displacement task*, in which participants make a saccade to a peripheral target; if during the saccade the target is displaced, participants are notoriously bad at discriminating its direction or even detecting its displacement (Bridgeman et al., 1975). The Saccadic Suppression of Displacement (SSD) is the phenomenon in which when a visual stimulus is displaced in the same direction as the saccade during the saccade, small displacements of the target are not perceived; as large as 10%–30% of the saccade's amplitude.

In 1996, Deubel, Schneider, and Bridgeman, discovered what they called the *temporal blanking effect*, which inhibited Saccadic Suppression of Displacement (Deubel et al., 1996). They found that subjects were no longer bad at detecting the direction of the displacement if it was displaced ≈200 ms after saccade landing, which did not occur when the target was displaced during the saccade, immediately after the saccade ended, or with smaller time gaps. Without the 200 ms blank, the pre-saccadic properties of the target match (with an error tolerance) the post-saccadic properties,

therefore small changes go unnoticed; with the extra time gap, a mismatch is easily detected and the target object is treated as a "surprising" or "new" object which has different spatial properties (Deubel et al., 2004, 1998). SSD can also be relieved when the displacement is orthogonal to the saccade (Wexler & Collins, 2014). Thus, when there is a discrepancy between the post-saccadic location of the target and its remapped location, visual stability suffers. Furthermore, the specific metrics of the saccade are known to the visual system because oculomotor noise is not confounded with object displacement (Collins, 2014; Collins et al., 2009). These results argue in favour of the importance of spatiotemporal information in establishing transsaccadic correspondence, in line with the remapping account.

However, SSD is also relieved when some (high-level) visual features of the object change simultaneously with the displacement. When a change in location is accompanied by a change in shape (Demeyer et al., 2010), polarity, or identity (Tas et al., 2012), the displacements are perceived. In other words, transsaccadic correspondence is enhanced when the pre- and post-saccadic objects share features, which argues in favour of transsaccadic feature correspondence as gating location correspondence, in line with the VSTM account.

The in-flight displacement task is consequently an important paradigm for understanding the interplay between remapping and VSTM in visual stability.

5 The Ideomotor Theory for Eye Movements

We have seen so far that the Ideomotor Theory proposes that perception and action are guided by the anticipation of the sensory consequences of an action. This anticipation can indeed be seen as an *expectation* of the action's sensory consequences (see 2 Theories of Action Control {p. 6}). If in fact actions are cognitively represented by their sensory effects, the Common Coding principle introduced by Prinz (1990) allows for a single internal representation for actions and effects, which supports Action-Effect associations' automaticity needed for learning. The notion of *anticipation* in the Ideomotor Theory is adopted as *prediction* by the Predictive Coding Theory and the Common Coding Principle. These two theories emphasize the *bidirectional* aspect of *common information* flow between high- and low-level networks of visual processing.

5.1 Motor Action Control

Right before an action is executed, its associated effects are activated at their respective sensory brain regions *in advance*. After execution, those *pre-activated* brain regions receive as input the action's *actual and current*[11] sensory consequences (the action's effects). Those sensory brain regions, consequently, get activated twice: first with the expected consequences and then by the real, environmental ones. This double activation gives the executing agent a) the "ability" to identify the action as its own based on the comparison between predicted and actual effects: Intentional Binding, and b) an "intensity" or "surprise" level of the expected sensory input: Sensory Attenuation. Both Sensory Attenuation and Intentional Binding phenomena are known to change as the time-lapse between pre-activation and activation varies, highlighting the crucial role of time within perception and action control. This information —the expected and the actual— is passed onto networks of increasing complexity until reaching a high-level unit whose task is to *evaluate* how different the estimated sensory feedback predicted by the Forward Model and the actual sensory feedback are. Then, the evaluation can be used to 1) update an AE mapping, 2) correct an ongoing action, 3) update the internal representation of the world, or 4) to alert the agent about a truly, and extremely unexpected sensory income.

O'Regan & Noë (2001) proposed that *seeing is a way of acting*. In other words, eye movements can simply be viewed as actions by which we explore the environment, like moving the hand to get the keys from the pocket. Considering eye movements as actions, they should follow the same principles specified by the Ideomotor Theory. And in fact, they do (Herwig and Schneider, 2014).

5.2 Oculomotor Control

In principle, what was described above can also happen for eye movements (saccades, specifically). The most relevant difference between a normal action in the Ideomotor Theory and a saccade lies in the kind of information the sensory effects are composed of. With each saccade, the *whole* visual field changes, regardless of the saccade's direction and magnitude. It should also be expected that extraocular muscle information be part of the pre-activated sensory effects. Let's assume, however, that sensory information regarding eye muscles for a 15^{dva}-saccade has little difference from a 2^{dva}-saccade, such that non-muscular information plays a more important role in the pre-activation of the effects. Without loss of generality, we can safely assume that muscle information is disregarded most

[11] That is, "real", because it is actual in time and current in sensory features.

of the time, leaving visual information as the most important sensory information for eye movements. Importantly, before a saccade is executed, the receptive fields of every neuron in the visual field are *pre-activated* with what they are to expect upon saccade execution: the shifting receptive fields. Once the eye has foveated a new region of the visual scene, these pre-activated neurons receive the *actual* sensory consequences of the saccade. This information, too, is passed onto visual networks of increasing complexity until reaching a high-level unit that we have been calling the *visual Forward Model* (vFM)[12]. It is important to say that the evaluation between expected and actual visual input is not executed for every shifted receptive field; such task would require a huge amount of computational resources. Therefore, the vFM must receive not only the new visual scene, but also relevant regions to evaluate within it (possibly "marked" by some attentional process), because without such marked relevant regions the vFM would have to compare the whole *previous* and *actual* visual scenes. Then, the vFM's output can be used to make another saccade (or different eye movement), or to modify behaviour.

To summarise: the vFM is constantly making predictions about objects in space and, to some extent, concerning their features. These predictions are based on foveal *and* peripheral information, enhanced by attention, that predisposes the vFM on what to expect after saccade execution. As in the Ideomotor Theory, past (pre-saccadic and peripheral) information is compared with the current (post-saccadic, foveal and peripheral) information. When the real visual scene matches the predicted post-saccadic scene, we perceive stability. Indeed, the visual system can be seen as a system containing a Forward Model that manages spatial predictions of objects in the visual scene to grant stability and guide attention. As mentioned before, the visual effects would be pre-activated by using attention-enhanced peripheral information about the pre-saccadic scene, including stimulus's features tightly linked to spatial locations (attentional pointers). Because the visual scene may constantly change —independently of eye movements—, the mapping between saccades and effects is stored in a trans-saccadic memory buffer (Prime et al., 2007), which is constantly updated and not learnt for long-term usage. Therefore, the essential usage of visual prediction is the correct identification of unique objects in the observer's visual field (Irwin and Gordon, 1998).

[12] For clarity and hereon, when we refer to the "visual Forward Model", we refer to the "Predictive system" that contains the Forward and the Inverse Models, the motor system, and the discrepancy unit that compares predicted versus actual sensory feedback. See FIGURE 3 {p. 15}. Especially taken into consideration in the next section, THE VISUAL FORWARD MODEL.

5.3 The Visual Forward Model

Considering that the vFM is in charge of predicting the correct location of objects followed by their features, Saccadic Suppression of Displacement could be explained in the following way:

1. Fixation: the observer foveally sees stimulus $\mathcal{S}$ at initial position, P_0.
2. Saccade: $\mathcal{S}$ disappears from current foveal position and reappears 10^{dva} to the right of the initial position; this cues the observer to execute a saccade to the new position, P_{new}.
3. Attention treats $\mathcal{S}$ peripherally obtaining its location and relevant features.
4. Imminent to saccade execution, receptive fields are shifted and updated, and the vFM receives as input (within the CD) the 10^{dva} magnitude and the direction of the saccade, and the features and locations of the expected object ($\mathcal{S}$) in the post-saccadic scene.
5. Displacement during the saccade[13]: $\mathcal{S}$ is displaced d^{dva} (distance of the displacement in degrees of visual angle) in the direction of the saccade.

    a. If d^{dva} does not exceed the tolerated saccadic error (Wexler and Collins, 2014), the vFM receives the post-saccadic scene, and evaluates the expected vs. actual comparison as "matched". Displacement goes unnoticed.
    b. If d^{dva} exceedes the tolerated saccadic error, the vFM receives the post-saccadic scene but evaluates the expected vs. the actual as "mismatched". The mismatched "error" is further passed until reaching consciousness making the observer aware of the displacement.

6. Displacement with blank: $\mathcal{S}$ is displaced d^{dva} in the direction of the saccade T milliseconds after saccade landing.

    a. If d^{dva} does not exceed the tolerated saccadic error and T is ≈50 ms, the vFM receives the post-saccadic scene and evaluates it as "matched", because saccadic suppression is still going on (Bridgeman et al., 1975; Deubel et al., 1998, 1996). Displacement goes unnoticed.
    b. If d^{dva} does not exceed the tolerated saccadic error but $T \geq 200\,ms$, the vFM receives the post-saccadic scene and evaluates the prediction as "matched", because saccadic suppression has terminated. Foveally, $\mathcal{S}$ is displaced d^{dva}. Changes in the current foveal scene trigger "error" or "surprise" which is further passed until reaching awareness; displacement is noticed.

[13] Or upon saccade landing, i.e. with no "blank".

c. If d^{dva} exceeds the tolerated saccadic error —regardless of T— vFM evaluates a mismatch and the displacement is noticed.

SSD using forward models assumes that with the blank gap the vFM has finished the current expected-to-actual evaluation when the target is displaced. The vFM would need to be re-activated to begin a new prediction if the observer makes another saccade towards the displaced target. In this sense, each eye movement activates the vFM discretely. This, however, does not impede that information from the previous vFM activation be present in the current activation.

The next section presents the experimental part, where we tested what type of information Forward Models take as input to make predictions, in the oculomotor and manual motor systems.

II EXPERIMENTS

Ideomotor Theory research has gathered much evidence of the existence of motor anticipatory mechanisms; however, research has mainly focused on the features that play a role within: the "what" content. Less is known about the way in which spatial information is used, the "where" content. Studies have not yet thoroughly investigated the spatial dynamics of Action-Effect associations. For example, it has not been investigated what kind of spatial information is used when the expected sensory effects of an action are evoked.

Riechelmann et al. recently reported relevant results for this work regarding Action-Effect anticipation in the oculomotor system (Riechelmann et al., 2017). Participants executed free-choice saccades to a peripheral traffic light, either to the left or to the right of a central fixation point. A saccade (the action) induced a light-switch (the effect) in the upper or the lower part of the traffic light. The switch of light could be either immediate or delayed (0, 100 or 300 ms). They found that saccades were shifted towards the location of the learnt light-change, which proved anticipation of the effect, because the location of the effect —which is spatial information— was anticipated upon saccade execution. However, within their paradigm, it is not clear under which spatial reference frame such anticipations are made.

In the first series of experiments (ACTION-EFFECT EXPERIMENTS {p. 49}), we tried to shed some light as to what information the FM uses to make predictions. The main purpose of those experiments is to find out how spatial and feature information are treated by the forward model after Action-Effect associations have been established, but also during their creation.

Transsaccadic correspondence experiments have found that colour, texture, etc. play an important role in visual stability. Specifically using the Saccadic Suppression of Displacement phenomenon, this has been tested by many studies. Nevertheless, to our knowledge, orientation as a feature has not been yet tested. In the second set of experiments (SACCADIC SUPPRESSION EXPERIMENTS {p. 88}), we tried to determine if orientation information can inhibit SSD, and to what degree visual features of a target influence its spatial perception. In other words, we tested if orientation is taken into account by the vFM to make a pre- and post-saccadic scene correspondence. We replicated previous findings on SSD to which the vFM apply as well.

1 Action-Effect Experiments

In the current study we addressed the question of whether Action-Effect associations are spatially specific. We did this by using a classic motion discrimination task in which we associated a direction of motion with a particular action. Previous work has investigated the temporal dynamics of Action-Effect associations. Desantis, Roussel and Waszak traced the time-course of the perceptual consequences of Action-Effect anticipation (Desantis et al., 2014). In an acquisition phase, subjects' actions (one of two keypresses) induced the appearance of a Random Dot Kinematogram (RDK) at threshold dot-coherence. Each of the actions was separately associated to an upwards or downwards dot motion. Later, in the test phase, subjects were instructed to report if the motion of the dots was upwards or downwards. On 50% of the test trials the previously used action-effect association was inverted, these were the incongruent trials. They found higher sensitivity to motion discrimination in a time-window of 220 ms before and 280 ms after the action. This result suggests that action-effect prediction modulates perception at later stages of motor preparation (Desantis et al., 2014).

We examined in what spatial reference frame Action-Effect associations might be specific: spatiotopic or retinotopic. A spatiotopic reference frame refers to a coordinate system of the outside world, for example, relative to the centre of the screen; a retinotopic reference frame refers to positions relative to the centre of the retina. The spatiotopic location of a stationary visual object does not change, but its retinotopic location changes with every eye movement.

1.1 Coherent Motion Detection Experiment

1.1.1 Methods

Subjects

Fifteen healthy human adults participated in the experiment (8 men, mean age: 29.4 years; standard deviation: 6.4 years). All reported normal or corrected-to-normal vision and no oculomotor or neurological impairments. They received payment of 10€/h for their time, and provided informed consent.

Apparatus

Stimuli were presented on a `Sony Multiscan CPD E400` CRT monitor with 1280×1024 (width by height) pixel resolution at an 85-Hz refresh rate inside a 352×264 mm area. Participants used forehead and chin rests, and viewed approximately 57 cm away from the screen. Responses were given via a `Logitech Wireless F710` gamepad. Saccadic eye movements were recorded using an `EyeLink 1000` (SR Research Ltd., Mississauga, Ontario, Canada) operated in monocular mode at 1kHz, tracking the right eye, with the `EyeLink` toolbox extension for `Matlab` (Cornelissen, Peters, & Palmer, 2002). Stimuli were generated using the `Psychophysics Toolbox` extensions for `Matlab` (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007).

Stimuli & Procedure

The fixation stimuli were three dots (radius=0.25^{dva}) along the horizontal axis of the screen, 7.4 ^{dva} apart and the central one at the centre of the screen. Peripheral stimuli were Random Dot Kinematograms (RDK). RDKs had identical number of dots (130), size (3 ^{dva} radius), dot-size (0.12 ^{dva} radius), speed (10 dva/s), colour (white) and duration (275 ms). The coherence of an RDK is the percentage of dots that have the same direction of motion. In a 100%-coherence RDK, all dots move in the same direction from one frame to the next, while in a 0%-coherence RKD each dot moves at a random direction sampled from a uniform distribution between 0 and 360°.

Eye movement acquisition

Viewing was binocular and movements of the right eye were monitored with an `EyeLink 1000` (SR Research, Mississauga, Ontario, Canada) at a 1kHz sampling rate. At the beginning of each block, the eyetracker was calibrated using a task-specific thirteen-point procedure. The eye movement traces were stored for offline analysis. For the offline analyses, the eye movement samples were smoothed with SR Research's proprietary algorithms. Instantaneous velocity and acceleration were computed for each data sample and compared to a threshold (30°/s and 9500°/s^2). Saccade onset was defined as two consecutive above-threshold samples for both criteria. Saccade offset was defined as the first sample of a 20-ms period of below threshold samples.

Design and Task

We used a randomised design which yielded a total count of 1024 trials. We had two different values for congruency (`congruent` versus `incongruent`) and for saccade (`fixation` versus `saccade` conditions), and four different positions. We repeated each of these combinations 64 times, giving the 1024 total split into 16 blocks. Each block consisted of an acquisition and a test phase. The test phase could be a fixation or saccade test, as illustrated in FIGURE 9. The test phase consisted of four randomised sub-blocks in which subjects were instructed before each sub-block to execute a saccade or to maintain fixation. The first block's acquisition phase had 100 trials, and was followed by two saccade test sub-blocks and two fixation test sub-blocks. Further acquisition blocks had 50 trials and the test phase sub-blocks' order was randomised. At the beginning of each block, the eyetracker was recalibrated.

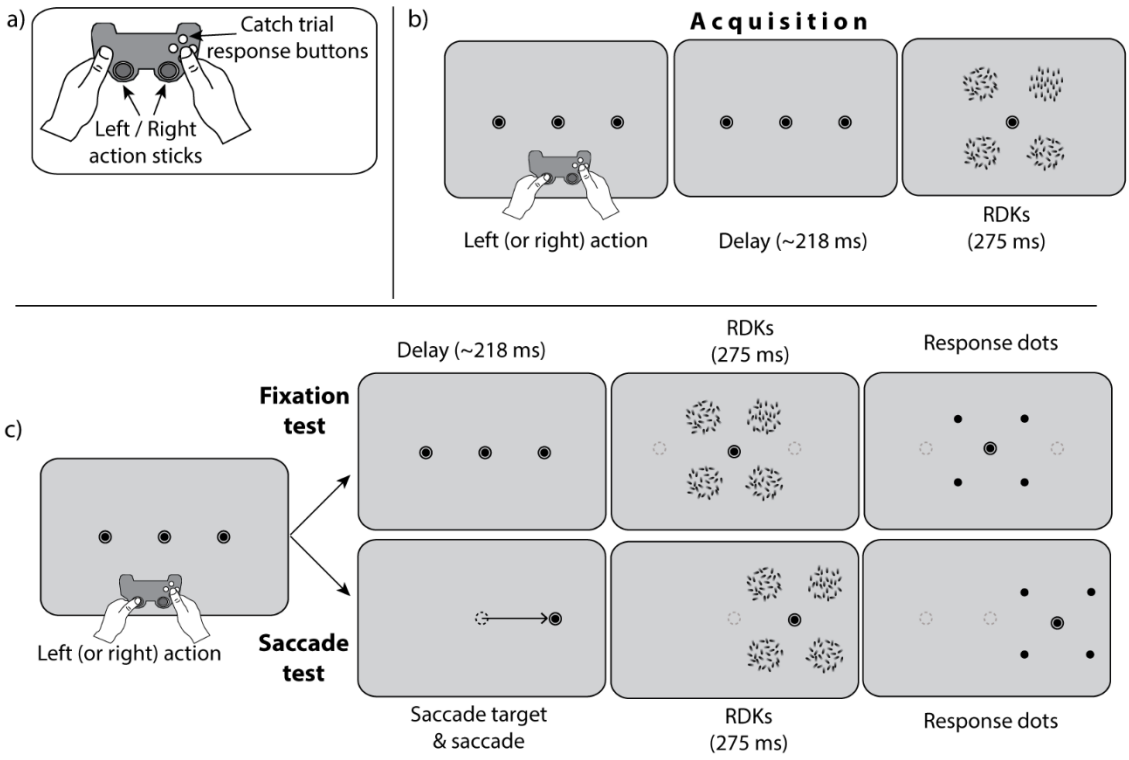

Figure 9. a) Gamepad. b) Acquisition phase. Subjects generated a left or right action using the sticks. After a ≈218 ms delay, four RDKs were presented, one 100% coherent and three 0% coherent. On 88% of trials, they simply observed the RDKs. In the remaining 12% catch trials, they had to press any of the four buttons to report they had seen the horizontal dot motion. C) Test phase. Subjects generated a left or right action. In fixation tests (top), 218 ms after the action 4 RDKs appeared for 275 ms, after which they were replaced by four response dots. Subjects were instructed to saccade to the dot at the location that the coherent RDK had been. In the saccade condition, subjects generated the left/right action and immediately executed a saccade towards the peripheral dot. Upon landing, the RDKs appeared for 275ms, after which they were replaced by the four response dots. Greyed dashed fixation dots were not present, they are displayed for geometrical reference.

Acquisition Phase

The purpose of this phase was to build action-effect associations. The action was a left- or a right-hand movement using the sticks of the gamepad, and the effect was the appearance of the stimuli. Subjects fixated the central fixation dot and performed the action. The action triggered the appearance of four RDKs (see FIGURE 9 {p. 51} and FIGURE 10), equidistantly from fixation, one of which had 100% coherent upwards or downwards motion. The other three RDKs had 0% coherence. The location at which the coherent RDK appeared (the *acquisition position*) was fixed for one subject, but counterbalanced across subjects. The action-effect association (e.g. left stick – upwards motion) was also fixed for one subject and counterbalanced across subjects. The delay between action and effect was determined individually (see ACTION-EFFECT DELAY {p. 54}).

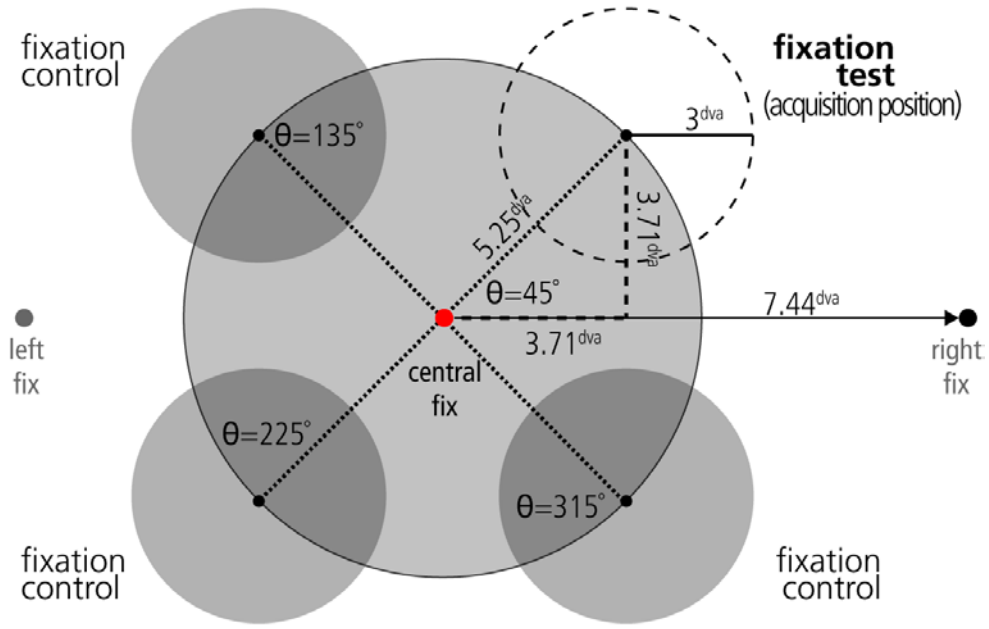

Figure 10 Task Geometry. RDKs shown relative to central fixation point. The same disposition applies if the RDKs are displayed relative to left or right fixation points (i.e.: saccade condition). The big grey circle has a 5.25° radius. Grey small circles denote RDKs with 3° radius. The location of fixation test RDK relative to central fixation starts at 45° and is consequently shifted 90° counter-clockwise for the remaining three RDKs. The size of the saccade that subjects would make was 7.44 dva, starting at central fixation and landing either on left or right fixations. The dashed line in the fixation test position indicates also the position at which the coherent RDK was displayed in the acquisition phase, the acquisition position.

Central fixation was checked throughout the duration of RDK presentation. When broken, subjects received feedback and the trial was rerun. Subjects were instructed to generate an equal number of

left and right actions. Whenever the difference between left and right actions was greater than 8%, subjects were reminded to balance their choices as close as possible to 50%–50%.

On 12% of the trials, the *catch trials*, the coherent dot motion was horizontal. Subjects' task was to press a key on the gamepad as soon as they detected horizontal movement, whatever its direction. Subjects had to respond to catch trials with the buttons. When they actioned one of the sticks and not the buttons, the trial was marked as incorrect. Catch trials ensured that subjects were paying attention to the coherent motion direction at their own acquisition position. Performance on catch trials was 83.1±22%, suggesting that subjects were paying attention to the RDK at the acquisition position.

Test phase

In the fixation condition, subjects' actions induced the appearance of four RDKs: the coherent RDK and three noise-RDKs. The noise-RDKs had zero coherence, but dot motion parameters were identical for the three. The coherent RDK could appear equiprobably at any one of the four locations. The coherence of this RDK and the delay between action and effect were determined individually with a staircase procedure and action delay measurement as described in ACTION-EFFECT DELAY {p. 54} and in COHERENCE THRESHOLD {p. 54}. After the 275 ms of display, the RDKs were extinguished and replaced by small response-dots at their previous centres. Subjects' task was to perform a saccade towards the response dot at the same location at which the coherent RDK had been. In the saccade condition, subjects were instructed to make a left/right action with the gamepad and to immediately make a saccade to the lateral fixation dot. Immediately after saccade landing, i.e. with 0 ms delay, the four RDKs appeared at the same positions as in the fixation condition but relative to the new fixation position (the saccadic cue). Subjects again had to indicate the location at which they had detected the coherent RDK by making a saccade to the corresponding response dot.

In congruent trials, the action-effect association used at the acquisition phase was maintained (e.g. left stick – upwards motion), while in incongruent trials the association was inversed (e.g. left stick – downwards motion). Whatever the congruency, the task was to detect the location of the coherent RDK, and to perform a saccade towards that location as a response. To facilitate response saccades, the four response dots were displayed until the saccade towards one of them was detected. Trials in which no saccade was detected within 2.25^{dva} of any of the response dots were discarded and rerun.

Eye position was monitored online, and trials were discarded and rerun if the subject made a saccade before performing the action, or if the latency between gamepad stick action and saccade landing

was greater than 300 ms. In both cases, feedback was provided to help subjects achieve the correct timing.

Action-effect delay

In the saccade test, subjects performed an action and then a saccade. RDKs were presented at saccade offset. Thus, there was delay between action and effect equivalent to Saccadic Reaction Time (SRT). We wanted to match this delay across all other conditions and phases (acquisition and fixation). To do so, before running the experiment we estimated individual SRT in a pre-test. This pre-test consisted of 50 trials in which subjects had to saccade from fixation to a peripheral dot on the same side as their saccade side in the experiment proper. The individual mean SRT from this pre-test was used in subsequent acquisition phases and fixation tests. At the end of each saccade test sub-block, a new SRT mean was measured and the action-effect delay updated accordingly.

Coherence Threshold

To determine each subject's coherence threshold, we ran a staircase procedure after the SRT pre-test. Subjects fixated the central dot and viewed an RDK in the periphery. On every trial, the RDK could appear equiprobably to the left or to the right of the fixation dot (i.e. the same left or right locations as in FIGURE 10 {p. 52}). Subjects had to *discriminate* its motion (upwards or downwards) by using the sticks. The staircase was a one-up-two-down procedure. The level of coherence was adjusted for at least 80 trials, and then until subjects reached ≈75% correct discrimination performance. This coherence threshold was used for the coherent RDK in the test sub-blocks. The mean coherence threshold found across subjects was 55.9±23%.

Task Geometry: Retinotopic vs Spatiotopic positions

Our design distinguished spatiotopic from retinotopic coordinates. In this experiment, each location was matched with a control location, as illustrated in FIGURE 11 {p. 55}.

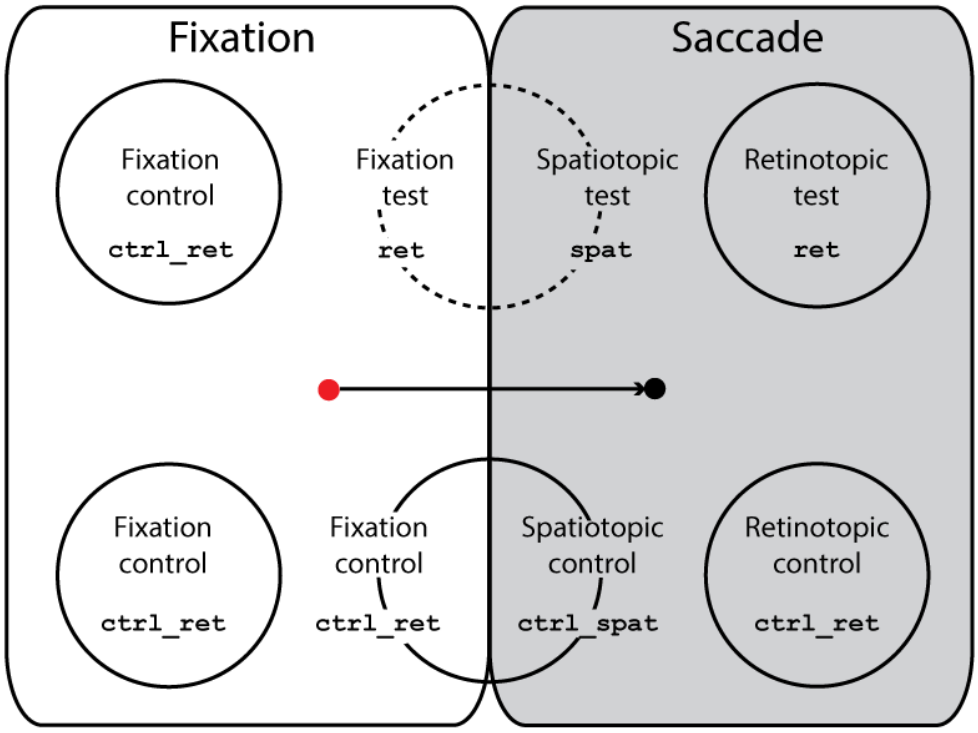

Figure 11 Retinotopic and spatiotopic positions. In this example, the acquisition position is centred at the top (indicated by the dashed circle). On the right (grey background), positions tested in the saccade condition; on the left (white background), positions tested in the fixation condition. The labels (`ret`, `spat`, `ctrl_ret`, `ctrl_spat`) will be used in the results section for clarity.

1.1.2 Hypotheses

We hypothesized that congruency would influence performance, and that the location of the coherent RDK would be less well detected when its motion was congruent with the acquired direction. We also hypothesized that the information learnt during the acquisition phase would include spatial information, and thus, that in the fixation condition detection would be worse at the acquisition position relative to control positions. In the saccade condition, we did not have any strong hypothesis about whether spatial effects would manifest in a spatiotopic or retinotopic reference frame.

We also hypothesized that subjects would take more time to report the appearance of an incongruent trial, than the previously learnt association, and that location would have a relevant role, in such a way that the acquisition position would take less time to be reported than others.

1.1.3 Data analysis

Our paradigm introduced changes in the visual stimulus contingent on the execution of a saccade. To unify the acquisition positions across all subjects, we collapsed leftwards and rightwards saccades. Data was pre-processed in order to remove trials which met the following criteria. For the online detection of saccades, gaze-contingent changes were triggered when the current eye position crossed a boundary 2.25^{dva} away from the central fixation dot. On average, saccade duration was 52.13ms±28, which left ample time for the saccade to be detected and the gaze-contingent change to occur within the ≈12 ms granted by an 85-Hz refresh rate. We eliminated trials in which the eye movement occurred too early or too late. Online, a trial was marked as `discarded` whenever one of the following conditions was met: 1) Fixation was not maintained on the central fixation dot during at least 200 ms before right/left action. 2) A saccade towards any of the RDKs was performed during the right/left action, or during the 275 ms of simulated motion. 3) In saccade trials, if a saccade towards the saccadic cue was started before the generation of right/left action. 4) The latency of the saccade was greater than 300 ms. The average number of (online) discarded trials was 361±219.

Offline, a trial was marked as `rejected` whenever the saccade towards the saccadic cue was shorter than half the expected saccade length (on average 3.42^{dva}) plus the RDK radius (3^{dva}), yielding on average a cut-off of 6.42^{dva}. On average 27±42 trials were rejected per subject (<3.4% of the trials).

We also filtered response saccades if: 1) Response saccade began before the response dots appearance. 2) Response saccade did not enter a 2.3 ^{dva} "imaginary circle" surrounding the response dots. If the response saccade was not filtered by these two criteria (i.e. did not land in the imaginary circle), the response was marked as `boundary`. Four boundaries (north, east, west, south) were defined for the subjects to be able to respond to a tentative area, while not being a proper response. The four boundaries, although present during the experiment, were not visible to subjects and they did not have any knowledge about their presence or usage. On average 107±58 trials per subject were marked as `boundary`.

After filtering, the average amount of valid trials was 867±107 (≈85%). Subjects executed 33% more of the 1024 trials, thus on average the task contained 361 extra trials discarded or rejected because of one of the criteria listed above.

For the statistical analyses we used R's `lme4` package to run a Linear Mixed-Effects Model (LMEM or "mixed-model") on the data. In general, two models were run separately for fixation and for saccade trials. This was necessary because under fixation trials, there was no distinction between spatiotopic

and retinotopic coordinates, while after an eye movement the two coordinate systems could be distinguished. However, all analyses were also run without doing such distinction, and fixation and saccade trials were all input together to a "joint" mixed model.

We used $maximal$ mixed-effects models as proposed by Barr et al. (2013). According to a series of simulations they performed in which power vs Type-I error ratio was measured and compared, mixed-effect models will explain data in the best way whenever random factors include as much information as possible from the data collection process (the experiment). Barr and colleagues state that random effects should be specified in a way according to the experimental design that is in line with the dataset. One of the main motivations to do so is that in many laboratory experiments, dependent variables are very well controlled during the data collection process. This causes statistical analyses to have few relevant theoretical variables which may have been rendered irrelevant through randomisation and counterbalancing. Yet another reason to use maximal mixed-effects models is that designs that only specify random-intercepts tend to have an increase of Type-I errors, which should be avoided. The model's predictor variable should be influenced by a clustering variable (`subject` in our case and in general) to adjust values for each cluster. Termed random-intercept, it allows the intercept term to vary across subjects relative to the predictor variables. This, however, does not allow to generalise results to the population from which subjects were randomly drawn. Accounting for this variation typically reduces random-effects residual error. Finally, given that mixed-models which only have random-intercepts do not account for by-subject variability, it is suggested that the model contain random-slopes as well. Otherwise, a subject's outlier values could create a cluster in the sample. Random-slopes, thus, allow subjects to vary also around the fixed effects.

As previously mentioned, we split the data into fixation and saccade trials, but performed a "joint" mixed-model in which data was not split. Performance (d') was measured following the method described by Green & Dai for mAFC tasks (Green and Dai, 1991), for the split-trials mixed models and for the joint-trials mixed models. Each set of data was then used to run a mixed model. For fixation trials (in the split approach), independent variables were position type (*retinotopic* versus *controls*) and congruency (*congruent* versus *incongruent*). For the saccade condition, the mixed-model included position type (*retinotopic*, *spatiotopic, retinotopic control and spatiotopic control*) and congruency. Both sets used subjects as grouping factor. We considered effects with a t-value greater than |2| to be significant, nevertheless p-values and their significance are also given for each corresponding t-value.

1.1.4 Results

d' is given for each position in both saccade and fixation conditions in FIGURE 12.

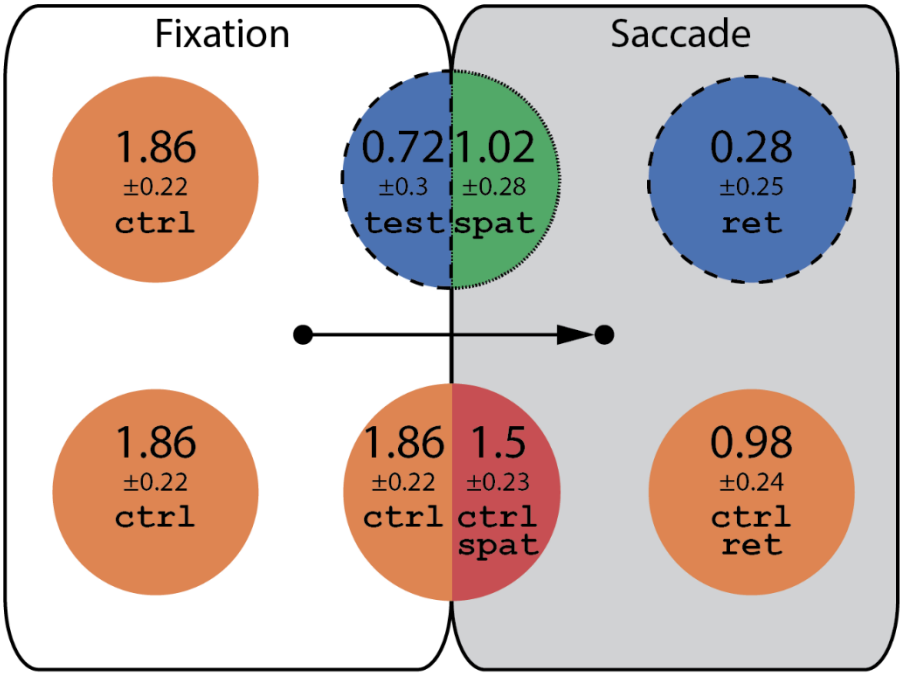

Figure 12 d' averaged across subjects for fixation and saccade conditions over the tested positions. The dashed circle represents the acquisition position.

Subjects show a clear advantage on fixation trials in detecting the coherent RDK at the fixation control positions (d'=1.86), as opposed to an almost null ability to detect it at the fixation test position (d'=0.72). In saccade trials, the retinotopic test (d'=0.28) position is still not detectable, while the spatiotopic test (d'=1.02), the retinotopic control (d'=0.98) and the spatiotopic control (d'=1.5) all show higher d' values.

Mixed-Models

The mixed model was run using d-prime (d') as the observed variable. For both models, results are given relative to the control position (`ctrl_ret`), which serves as intercept. In fixation trials, only control (`ctrl_ret`, is displayed as intercept) and retinotopic (`ret`) positions are distinguished. The acquisition position is `ret` in fixation trials, and by `spat` in saccade trials (see FIGURE 11 {p. 55} and section TASK GEOMETRY: RETINOTOPIC VS SPATIOTOPIC POSITIONS {p. 54}).

Fixation	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	1.879	0.1837	10.2273	0	***
ret	-1.1684	0.2637	-4.431	0.0008	***
congruent	-0.0665	0.0748	-0.8894	0.2571	ns
ret × congruent	0.1167	0.0907	1.2866	0.1682	ns
Saccade	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.9843	0.247	3.9859	0.0007	***
ctrl_spat	0.5775	0.244	2.3662	0.0287	*
ret	-0.7027	0.2109	-3.3325	0.0034	**
spat	0.0823	0.2783	0.2958	0.3772	ns
congruent	-0.0198	0.1122	-0.1761	0.3885	ns
ctrl_spat × congruent	-0.078	0.1501	-0.5194	0.3433	ns
ret × congruent	0.1119	0.1501	0.7456	0.2965	ns
spat × congruent	0.0787	0.1501	0.5244	0.3424	ns

LMEM 1 Mixed-model on fixation trials (top) and saccade trials (bottom) for the detection experiment.

The mixed-model for fixation trials (upper part of LMEM 1) showed that at the retinotopic position (ret) subjects significantly detect worse the coherent RDK (estimate=-1.168, t-statistic=-4.43, stderr=0.3, p-value<0.001) as compared to the control positions (all labelled under ctrl_ret). There was no significant contribution of congruency (t-statistic=-0.89, p-value=0.3) and no interaction between congruency and position type (t-statistic=1.28, p-value=0.2). The random-effects (LMEM 13 {p. 143}), however, showed a high variance for the retinotopic position ($\sigma^2 = 0.98$) and a non-negligible variance for control positions ($\sigma^2 = 0.475$). The residual variance was relatively low ($\sigma^2 = 0.03$). Correlation for the random-effects were also low. The high variance of the retinotopic position may be explained by the fact that in each trial control positions were more likely to contain the coherent RDK than the retinotopic position.

The mixed-model for saccade trials (lower part of table LMEM 1) revealed a significant lower sensitivity to test positions (ret and spat) compared to control ones (ctrl_ret & ctrl_spat) all relative to the retinotopic control (spatiotopic control: t-statistic=2.36, stderr=0.2, p-value=0.029; retinotopic test: t-statistic=-3.3, stderr=0.2, p-value=0.003; spatiotopic test: t-statistic=0.29, stderr=0.3, p-value=0.4). The difference between spat and ctrl_spat (t-value=-2.31; p-value=.03) was significant (see TABLE 1 {p. 148}). The retinotopic position was significantly different from all other positions (see TABLE 1; t-values=-3.3, 3.8 & 2.4; p-values=0.003, 0.001 & 0.03). Congruency was not significant (t-statistic=-0.176, stderr=0.1, p-value=0.4). No significance was found on any of the possible interactions of independent variables. The random-effects (LMEM 15 {p. 144}) in saccade trials did not show any relevant correlation between variables. There were nonetheless high variances for all positions, being the retinotopic the lowest one ($\sigma^2 = 0.4$). The

residual variance ended up being 0.085, which relative to the before mentioned variances is quite low. The low value of the standard error in the fixed-effects (0.075) and of the variance in the random-effects for congruency ($\sigma^2 = 0.097$) are justified by the fact that congruent or incongruent motion was an irrelevant feature of the task. The big amount of error that the position type leaves in the fixed-effects (on average 0.25) is the largest of all, yet, only two positions remain significant as a main effect. These high values for standard error and variance for position type suggest an increase of uncertainty when subjects reported the coherent RDK explained by the fact of having 4AFC task. A comparison between the observed d' (as displayed in FIGURE 12 {p. 58}) compared with the mixed-model's estimates (as shown in LMEM 1 {p. 59} for fixation and saccade trials) is presented in the appendix for reference in FIGURE 37.

Expanded Mixed-Model

In the mixed model presented for fixation trials (LMEM 1 {p. 59}), there are two distinguishable position. The `fixation test` (named `ret` in the mixed model's output), which corresponds to the acquisition position, and the control positions, which correspond to any position other than the acquisition position (named `ctrl_ret`, serving as intercept in the mixed-model). Although this is conceptually true, there are physically four positions, as can be viewed in FIGURE 13 {p. 61}. We reasoned that in reality fixation trials did contain four distinguishable positions, even though conceptually they served as controls and were averaged together. Next, we ran the mixed-model making the distinction among positions within fixation trials: this is the "expanded" mixed model. The computed values by position & condition for the expanded fixation trials are shown in FIGURE 13 {p. 61}.

The mixed model for the expanded fixation trials is shown in LMEM 2 {p. 61}. The retinotopic position (that is, the acquisition position, `fixation ret` in the figure, `ret` in the mixed model output) is the only one to show a significant contribution in decreasing subjects' performance (estimate=-1.223, stderr=0.2, t-statistic=-5.59, p-value=0). Other positions do not show significant contributions. Specifically, the difference between the retinotopic control (`ctrl_ret`) and the spatiotopic controls (`ctrl_spat`) is not significant (t-statistic=0.087). As expected, congruency did not influence subjects' ability to detect coherent motion. None of the predictors interacted between each other.

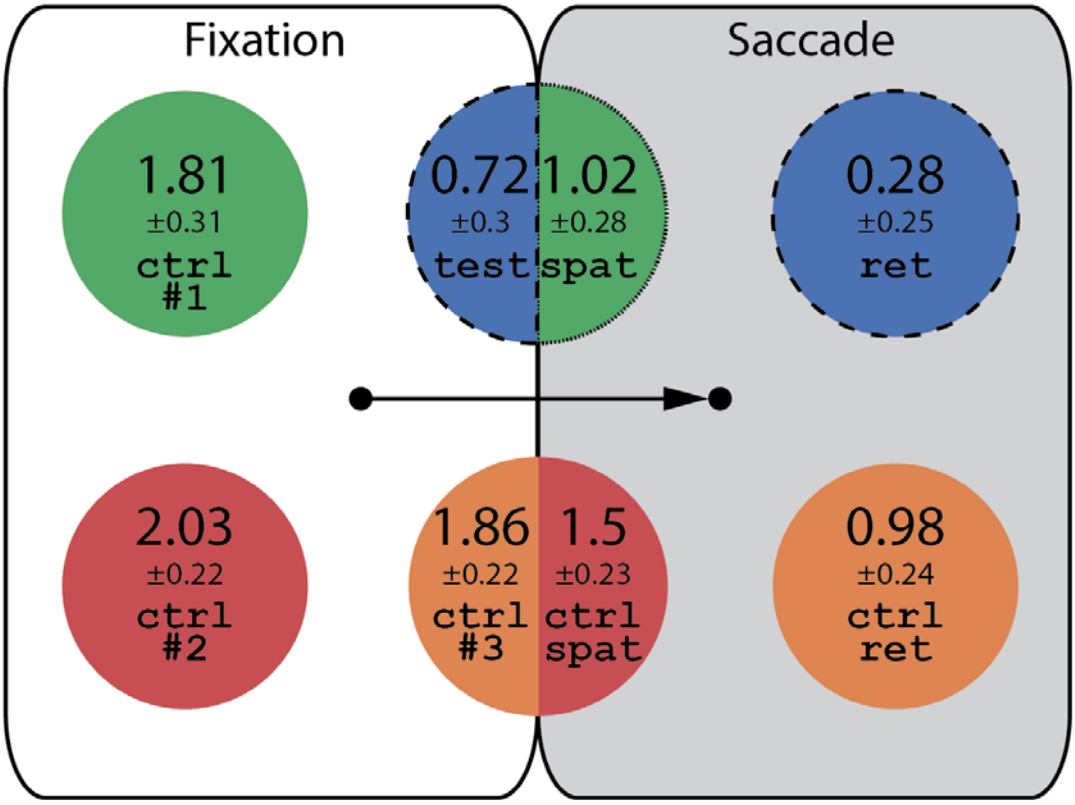

Figure 13 Performance for the expanded fixation positions. The fixation test is named `ret` *in the mixed model, while the fixation controls are considered separately and labelled as in saccade trials: based on* FIGURE *12 control #1 is labelled* `spat`*, control #2* `ctrl_spat` *and control #3* `ctrl_ret`*.*

Fixation	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	1.9331	0.2123	9.104	0	***
ctrl_spat	0.0872	0.2341	0.3724	0.3674	ns
ret	-1.2225	0.2187	-5.5898	0	***
spat	-0.0341	0.3218	-0.1058	0.3925	ns
congruent	-0.1316	0.0897	-1.4673	0.1349	ns
ctrl_spat × congruent	0.1693	0.1174	1.442	0.1397	ns
ret × congruent	0.1818	0.1174	1.5479	0.1202	ns
spat × congruent	-0.007	0.1174	-0.0594	0.3941	ns
Saccade	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.9843	0.247	3.9859	0.0007	***
ctrl_spat	0.5775	0.244	2.3662	0.0287	*
ret	-0.7027	0.2109	-3.3325	0.0034	**
spat	0.0823	0.2783	0.2958	0.3772	ns
congruent	-0.0198	0.1122	-0.1761	0.3885	ns
ctrl_spat × congruent	-0.078	0.1501	-0.5194	0.3433	ns
ret × congruent	0.1119	0.1501	0.7456	0.2965	ns
spat × congruent	0.0787	0.1501	0.5244	0.3424	ns

LMEM 2 Mixed model results for the expanded version of the fixation trials.

The difference between the retinotopic position (`ret`) and the spatiotopic position (`spat`) for fixation trials was significant (see TABLE 2 {p.149}) (estimate=1.1885, t-value=-3.056; p-value=0.0065). The spatiotopic control (`ctrl_spat`) was significantly different to the retinotopic position (estimate=1.3097, t-value=3.9973; p-value=0.0007). Random-effects for the expanded fixation trials (LMEM 14 {p. 144}) decrease the variance of the retinotopic position and it is spread into the differentiated control positions. We observed that the controls now display higher variances ($\sigma^2 = 0.6246$ & $\sigma^2 = 0.7842$) than the other positions ($\sigma^2 = 0.6407$ & $\sigma^2 = 0.4503$), specifically the spatiotopic one. Residual variance ($\sigma^2 = 0.057$), however, was higher than the unexpanded version.

Joint Mixed-Model

Later we reasoned that performance should also be measured taking fixation and saccade trials together, to determine the influence of the saccade condition on subjects' overall performance. We joined fixation and saccade trials and ran the mixed-model using the same predictors and the same predicted variables. This mixture was possible by introducing as a predictor the `saccade` condition (fixation vs saccade).

Positions in fixation trials are considered as in EXPANDED MIXED-MODEL, thus controls are not averaged together. Predicted results are shown in LMEM 3 {p. 63}. Overall, retinotopic test position (`ret`) shows a significant decline in performance (estimate=-1.223, stderr=0.2, t-statistic=-6.012, p-value=0). None of the other positions contribute to subjects' performance (t<2). Performance significantly decreased by 0.95 under saccade trials (estimate=-0.949, stderr=0.2, t-statistic=-5.747, p-value=0). The spatiotopic control (`ctrl_spat`) interacts with the `saccade` factor, indicating that this position is significantly different from the retinotopic position (`ret`) when executing or not a saccade. The retinotopic position shows a significant interaction with the `saccade` predictor, demonstrating that retinal information still influences participants' decision regardless of the saccade. The differences of the estimates (TABLE 3 {p. 149}) were significant between `ctrl_spat` and `ret` (t-value=4.016), and between `ret` and `spat` (t-value=2.782). Random-effects (see LMEM 16 {p. 144}) showed the spatiotopic position (`spat`) as having a very large variance ($\sigma^2 = 1.2$) which is a result of the significant difference with the retinotopic position and the non-significant difference with the other positions. Variances for retinotopic and the spatiotopic control remain at the high levels (`ret`: $\sigma^2 = 0.4$; `ctrl_spat`: $\sigma^2 = 0.68$) with same magnitude as the intercept ($\sigma^2 = 0.67$). High correlations ($\rho \approx 0.7$) were found between `spat` and `ctrl_spat` ($\rho = 0.77$) and between `spat` and `congruent` ($\rho = 0.72$). The residual variance for the random-effects suggests that little unexplained data was left ($\sigma^2 = 0.1$).

	estimate	stderr	t-value	p-value	sign
(Intercept)	1.9331	0.227	8.5002	0	***
ctrl_spat	0.0872	0.245	0.3564	0.3714	ns
ret	-1.2225	0.203	-6.0118	0	***
spat	-0.0341	0.302	-0.1127	0.3938	ns
congruent	-0.1316	0.121	-1.0902	0.2174	ns
saccade	-0.9487	0.165	-5.747	0	***
ctrl_spat × congruent	0.1693	0.167	1.0126	0.2358	ns
ret × congruent	0.1818	0.167	1.087	0.2181	ns
spat × congruent	-0.007	0.167	-0.0417	0.396	ns
ctrl_spat × saccade	0.4903	0.167	2.9322	0.0074	**
ret × saccade	0.5198	0.167	3.1086	0.0048	**
spat × saccade	0.1164	0.167	0.6962	0.3096	ns
congruent × saccade	0.1118	0.167	0.6688	0.3155	ns
ctrl_spat × congruent × saccade	-0.2473	0.237	-1.0458	0.2279	ns
ret × congruent × saccade	-0.0698	0.237	-0.2952	0.379	ns
spat × congruent × saccade	0.0857	0.237	0.3624	0.3705	ns

LMEM 3 Mixed model results for the joint (saccade and fixation trials) version of the data. Fixation trials positions are not averaged together.

Response Times

RTs in this experiment measure the response-saccade latency: the time between the appearance of the response dots and the saccade's landing on one of the response dots. RTs show an inverse but similar behaviour to d-primes. Congruency did not show any significance. On average subjects took 309±298ms to make a valid saccade to a response dot. Some subjects executed the saccade very fast (11 ms) which resulted in an incorrect response. For fixation trials, correct responses were faster than incorrect ones (incorrect=359.8 vs correct=313.8, t-statistic=5.7, p-value=0). Saccade trials exhibit the opposite behaviour (incorrect= 264.1 vs correct=303.9, t-statistic=-5, p-value=0). On fixation, the retinotopic position showed very elevated RTs as compared to the other positions (fixation trials are expanded). On the other hand, in saccade trials the retinotopic position decreases dramatically, but remains the highest of all four. At the spatiotopic control (`ctrl_spat`) subjects respond the fastest. The retinotopic control and the spatiotopic test display more or less the same average RT. These results again rectify the facility that subjects show to the spatiotopic control (in saccade trials), and a trace of uncertainty whenever the coherent RDK contains the acquisition conditions. Comparing only fixation versus saccade trials, fixation trials exhibit significantly (t-value=7.74, p-value=0) slower speeds (328.14 ms) than saccade trials (286.24 ms).

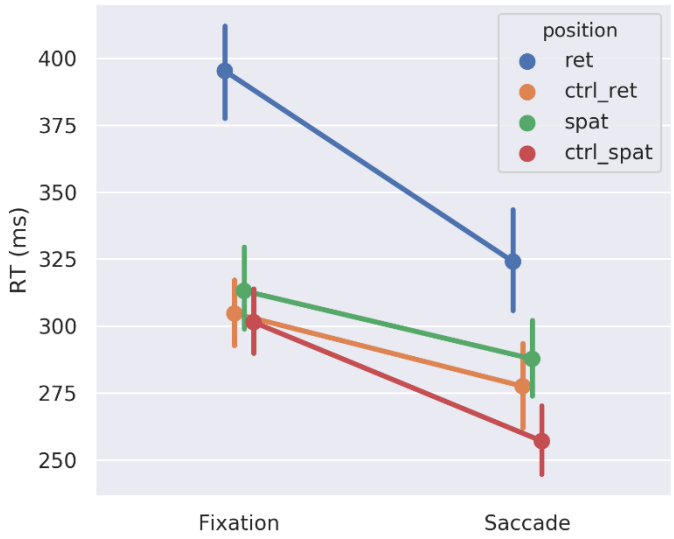

Figure 14 Response Times for fixation and saccade conditions. Retinotopic test position (`ret`) is the slowest to report, while the spatiotopic control the fastest.

We ran a mixed-model using RT as dependent variable. We applied the same procedure as with previous mixed-models; however, fixation positions are expanded. For fixation trials (see LMEM 4) the retinotopic position (`ret`) showed a significant effect (estimate=90.25, stderr=28.9, t-statistic=3.1, p-value=0.005). In saccade trials, none of the positions showed a significant effect. For both models, congruency and positions' interactions were not significant.

Fixation	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	300.3226	25.535	11.7612	0	***
ctrl_spat	4.7722	23.8141	0.2004	0.3866	ns
ret	90.2466	28.9	3.1227	0.0056	**
spat	20.579	31.1059	0.6616	0.315	ns
congruent	4.0345	14.3481	0.2812	0.3789	ns
ctrl_spat × congruent	-16.7568	19.2989	-0.8683	0.2682	ns
ret × congruent	0.6646	19.4842	0.0341	0.3946	ns
spat × congruent	-17.4214	19.3365	-0.901	0.2605	ns
Saccade	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	285.6387	36.3671	7.8543	0	***
ctrl_spat	-31.9901	24.0629	-1.3294	0.1624	ns
ret	53.0895	27.9124	1.902	0.0683	ns
spat	7.6679	20.5922	0.3724	0.3674	ns
congruent	-18.6223	14.9068	-1.2492	0.1796	ns
ctrl_spat × congruent	14.4964	19.7926	0.7324	0.2995	ns
ret × congruent	-10.6437	19.9963	-0.5323	0.3409	ns
spat × congruent	-5.1931	19.9064	-0.2609	0.3811	ns

LMEM 4 Fixed effects for the RTs (in milliseconds) in the motion detection experiment for expanded fixation trials and saccade trials.

Another mixed-model was run with positions labelled as *control* or *test* and distinguished between reference frame —spatiotopic or retinotopic— (see LMEM 5). Therefore, positions previously labelled `ret` and `spat` became `tests`; positions labelled `ctrl_ret` and `ctrl_spat` became `controls`; finally, positions in fixation were all labelled as `ret_rf` while positions in saccade trials were labelled `ret_rf` if they were retinotopic (`ctrl_ret` or `ret`) or `spa_rf` if they were previously spatiotopic (`ctrl_spat` or `spat`). The hypothesized change in RT was found, with the additional constraint that the controls displayed a significant difference with respect to the tests. For fixation trials all reference frame labels were —arbitrarily chosen to be— retinotopic; in saccade trials, positions aligned with the acquisition's x-axis were assigned to a spatiotopic reference frame, and equal retinal positions as in the acquisition phase were assigned to the retinotopic reference frame. Positions of type `test` took significantly more time to be reported (estimate=63.8, stderr=19.7, t-statistic=3.24, p-value=0.005) as compared to control positions (the intercept). The spatiotopic reference frame, on the other hand, decreased RTs by approximately 50 ms (estimate=-50.41, stderr=19.14, t-statistic=-2.63, p-value=0.017). These results confirm what performance tests revealed: test positions (`ret` & `spat`) take more time to be detected and it is easier for subjects to detect coherent motion under the retinotopic reference frame.

	estimate	stderr	t-value	p-value	sign
(Intercept)	303.5831	20.35	14.9198	0	***
test	63.7995	19.7	3.2388	0.0049	**
congruent	-9.3377	7.594	-1.2296	0.1831	ns
ref_frame	-50.4065	19.14	-2.633	0.0175	*
test × congruent	1.5601	12.29	0.1269	0.3904	ns
test × ref_frame	-23.0838	17.68	-1.3058	0.1667	ns
congruent × ref_frame	5.2827	16.69	0.3166	0.3735	ns
test × congruent × ref_frame	-19.9971	24.88	-0.8036	0.2819	ns

LMEM 5 Fixed-effects joint mixed-model estimating RT (in milliseconds). Positions are classified into control (`ctrl_ret` & `ctrl_spat`) and test (`ret` & `spat`). A category reframe classifies trials as testing the retinotopic coordinate system (all positions for fixation trails, and `ret` and `ctrl_ret` for saccade trials) or the spatiotopic coordinate system (`spat` & `ctrl_spat` for saccade trials).

1.1.5 Discussion

In this experiment we saw that, regardless of eye movement and congruency, the acquisition position (retinotopic) was the hardest to detect. For fixation trials on the one hand, the *control* position showed an increase of 2.5 times the retinotopic sensitivity value, demonstrating that subjects were very good at locating the coherent RDK whenever it was *not* at the acquisition position. Judging by this result, it would appear that acquisition conditions were being neglected. Moreover, there was

no difference among control positions. In saccade trials we found the same pattern: subjects were not able to detect the coherent RDK when it appeared at the retinotopic position (i.e.: the retinal location of the acquisition position); they performed better in the retinotopic control and the spatiotopic positions, but not significantly different from retinotopic.

Performance was best when the tested position was the spatiotopic control, far away from the retinotopic position. The simplest way to explain this result is by affirming that sensory attenuation phenomenon took place. Moreover, such explanation would be in line with the pre-activation hypothesis (Roussel et al., 2013). According to this hypothesis, the execution or the preparation of an action pre-activates sensory networks which contain the expected perceptual consequence; in our case, the expected consequence is a coherent RDK at the acquisition position. During each acquisition phase, these networks increase their mean level of activity to a "pedestal" level (Cardoso-Leite et al., 2010; Desantis et al., 2014; Roussel et al., 2013). The detected decrease in sensitivity (d') for the acquisition position is due to the big difference in signal with respect to the beforementioned pedestal level. In other words, it is easier for the system to detect the *non-acquired* signals whose pre-activations are closer to the baseline level than to detect the signal at the pedestal level. In this sense, sensitivity at the acquisition position is reduced because the mean of the pre-activated distribution has increased.

As depicted in FIGURE 15 {p. 67}, activation signals are far away from baseline (which is near to or at zero level). At the acquisition phases, as the experiment progresses, Action-Effect associations are created. Later and still during acquisition, such actions begin to be predicted as the sensory consequences of each action. With the prediction, the corresponding sensory networks get activated prior to action execution. Such prediction and pre-activation involve a specific motion (up or down according to the left or right action) located precisely at the retinotopic position (top right in the figure). When tested during test phases, such acquired pre-activation is present regardless of the tested position. When the actual stimulus signal is seen, subjects' ability to detect the position of the coherent RDK depends on the difference between the pre-activation and the stimulus signals. Therefore, when the tested position is the acquisition one (retinotopic), the beforementioned difference is smaller, inducing a reduce sensitivity at that precise location and thus subjects' inability to detect it. On the other hand, when other positions are tested, the pre-activation signals are closer to the corresponding stimulus's baseline, and consequently the difference in distance between the actual signal (either S1 or S2) and the pre-activation signals is larger, making subjects more easily perceive coherent motion at such locations. The same applies for fixation trials; however, the distance between pre-activation and actual signals at the retinotopic position is smaller, making the observed d' larger than in saccade conditions.

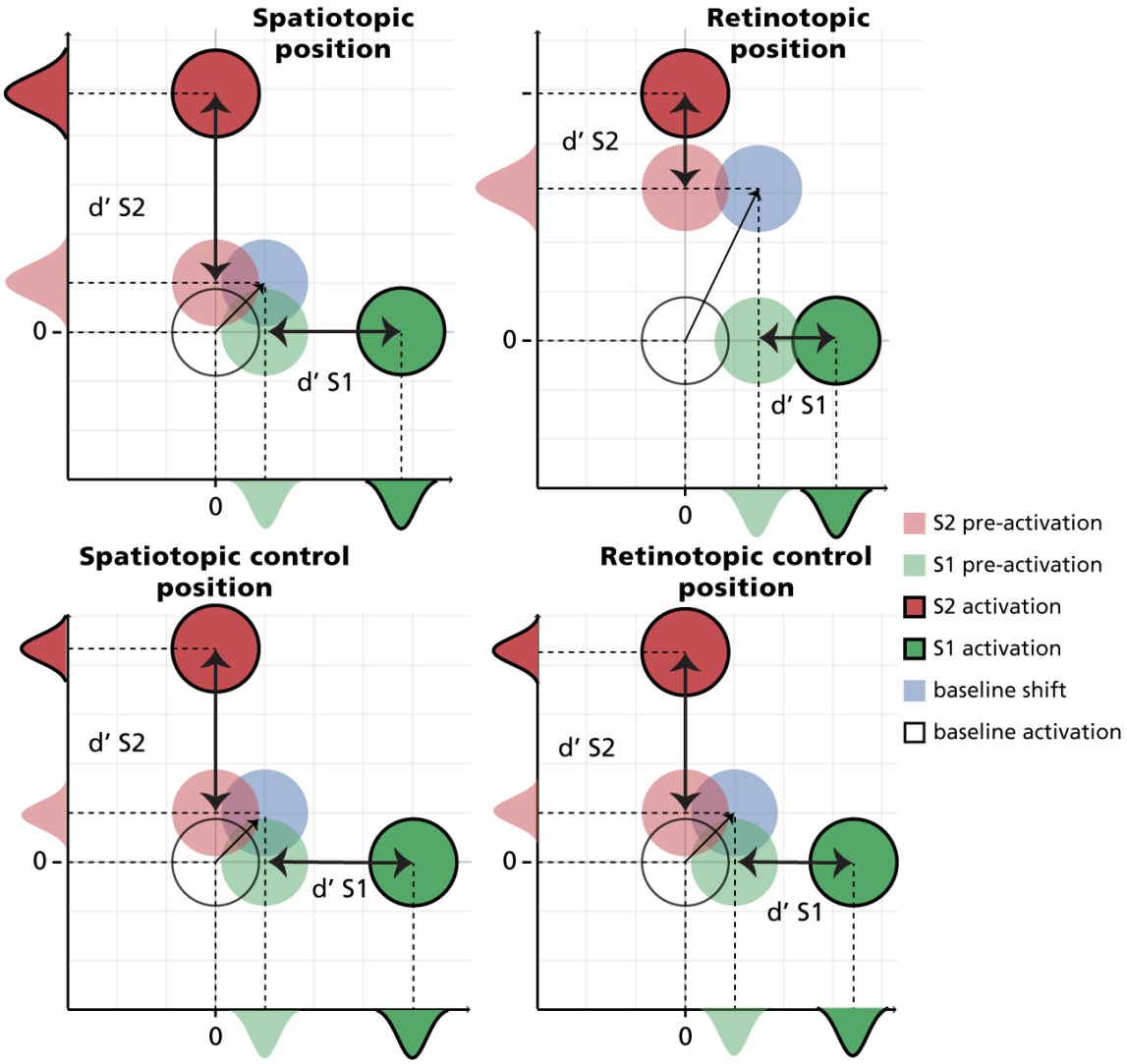

Figure 15 Schematic illustration of the pre-activation hypothesis for trials under saccade condition for all four tested positions. S1 and S2 are the respective up and down signals. Positions are located following the same order as displayed in FIGURE 11. The pre-activation signals are shown in transparent background, all above zero. Stimuli signals are displayed in non-transparent background. In red, signals for downwards motion and in green for upwards motion. Each of the pre-activation signals has a different distribution for the corresponding dot motion.

Another aspect that influences these results is the task's geometry, notably the distances between positions relative to the retinotopic position (see FIGURE 10 {p. 52}). The distance between the retinotopic position and the retinotopic control or the spatiotopic is 7.44^{dva}, while the distance with the spatiotopic control is 10.52^{dva}. Although these distances are equal relative to fixation (5.25^{dva}), it appears that attentional resources are not being deployed equally in space[14]. Results, however, do not allow us to determine if attentional resources are distributed under retinotopic or spatiotopic coordinate systems. The spatial redistribution of attentional resources after a saccade are

[14] Independently of the pre-activation explanation.

geometrically and topographically the same, although it is clear that attention suffers when a second action —the saccade— is executed.

To be able to report the location of the coherent RDK, subjects had to divide their attention into four locations. By dividing their attention equally, subjects' probability of correctly detecting the coherent RDK was ¼. If attention was being serially switched from one position to another as suggested in the spotlight approach (James, 1890; Posner and Cohen, 1984; Posner and Petersen, 1990), subjects had a 68.75 ms attentional time-window at each location. If we consider control positions separately in fixation trials (the "expanded" version), their values do not differ significantly among each other, suggesting that subjects were particularly neglecting attention to the acquisition position. This, in turn, allowed them to reduce the probability of being correct to ⅓ and widened their serial processing time-window to 91.6 ms. The ratio between retinotopic and the other positions (both, expanded or collapsed) is also approximately ⅓. This approach, however, does not explain in itself why the acquisition position was 'neglected'.

Furthermore, the observed behaviour took place at the first stages of the experiment in testing phases. In fixation trials, the retinotopic position displays a significantly poor performance as opposed to the spatiotopic and the retinotopic control positions. The spatiotopic control, on the other hand, displays a high performance from the beginning of the experiment, after later reaching a plateau. Performance maintains stable after the first block, for all positions as displayed in XXX. The same behaviour was found for saccade trials and for the ensemble of trials where performance fluctuates less XXX.

Finally, RTs completed the findings. It appears that subjects had equal speed at reporting the retinotopic control position in fixation trials and spatiotopic control in saccade trials. This can be seen by the limits of their confidence intervals. The overall decrease of speed on fixation trials relative to saccade trials is to be expected, as saccade latencies tend to decrease as a sequence of saccades lengthens (McSorley et al., 2019).

1.2 Motion Discrimination Experiment

1.2.1 Methods

Subjects

Twelve healthy human adults (6 women, mean age: 28.9 years old; standard deviation: 5.9 years) participated in the experiment. All reported normal or corrected-to-normal vision and no oculomotor or neurological impairments. They received payment of 10€/h for their time, and provided informed consent.

Apparatus, Stimuli & Procedure

This experiment used the same apparatus and configuration as in the COHERENT MOTION DETECTION EXPERIMENT (EXPERIMENT 1 {p. 49}). Stimuli measures and parameters were the same as in EXPERIMENT 1, except for the spatial configuration, as detailed below.

Eye movement acquisition

Eye movement acquisition was the same as for COHERENT MOTION DETECTION EXPERIMENT {p. 49}. On average, saccade duration was 57.43ms±46. Trials were discarded and rerun following the same rules as in EXPERIMENT 1. The number of online discarded trials was on average 248±241. On average, subjects executed 43% more of the 576 trials, thus the task contained 248 extra trials discarded because of one of the beforementioned criteria.

Design and task

One experimental session consisted of 576 trials, resulting from the combination of three factors: stimulus position (3 locations, see FIGURE 17 {p. 71}), congruency (motion direction congruent or incongruent with the acquisition phase), eye-movement condition (fixation versus saccade), and 48 repetitions per condition. These trials were split into 8 blocks. At the beginning of each block the eye tracker was calibrated. The experiment was split into acquisition and test phases, with two types of tests: fixation and saccade. Each of the eight blocks was further divided into 4 sub-blocks, two saccade tests and two fixation tests. Each of the 8 blocks started with an acquisition phase; 100 trials for the first block, and 50 in the remaining seven blocks. The four test sub-blocks were run in

randomized order, except for the first block in which the sub-block order was two saccade tests followed by two fixation tests. FIGURE 16 {p. 70} illustrates the task.

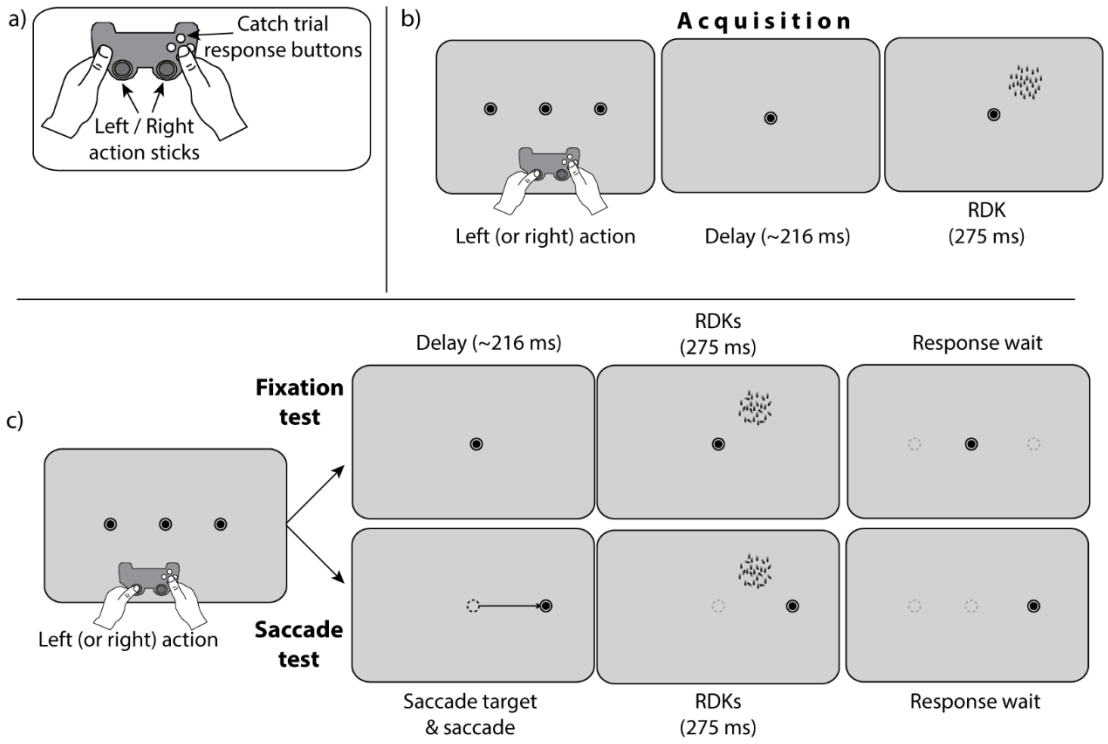

Figure 16 ***a)*** *Gamepad used by subjects.* ***b)*** *Task for the Acquisition phase. During the acquisition phase subjects generated randomly left or right action using the sticks; when a catch trial was on, they were asked to report its presence using the any of the four buttons.* ***c)*** *Test phases. Subjects generated a free-choice action between left and right to induce the appearance of an RDK. On fixation tests, subjects remained fixated on the central fixation point during the 275 ms of RDK motion, to answer by performing up or down stick movements to the dot motion direction. In saccade tests, subjects generated left/right action right before executing the saccade towards the saccade cue dot. Upon landing, the RDK appeared for the same period of time, after which they reported the RDK motion direction. Greyed dashed circles were not visible, they are displayed for geometrical reference.*

Task Geometry: Retinotopic vs Spatiotopic positions

Subjects were asked to report the motion direction of an RDK stimulus, which could appear at one of three positions defined relative to the fixation dot and the acquisition position (i.e. the position at which subjects had evoked the appearance of an RDK during the acquisition phase; FIGURE 16). We differentiated spatiotopic, retinotopic and control positions. In the fixation test, retinotopic and spatiotopic positions were confounded. Indeed, when the coherent RDK appeared at the same position on the screen as in the acquisition phase, the RDK had the same coordinates both in space and on the retina. There were two control positions, one in the opposite hemifield and one in the upper visual field. In the saccade test, positions were defined relative to the new, post-saccadic fixation dot. When the RDK appeared at the same location on the screen as in the acquisition phase, it had the same spatiotopic coordinates, but different retinotopic coordinates, since the retina had

moved due to the intervening saccade. In retinotopic trials, the RDK appeared at a location that had the same retinal location as the acquisition position, but it was at a different position on the screen. The control location was in the upper hemifield. See FIGURE 17 {p. 71}.

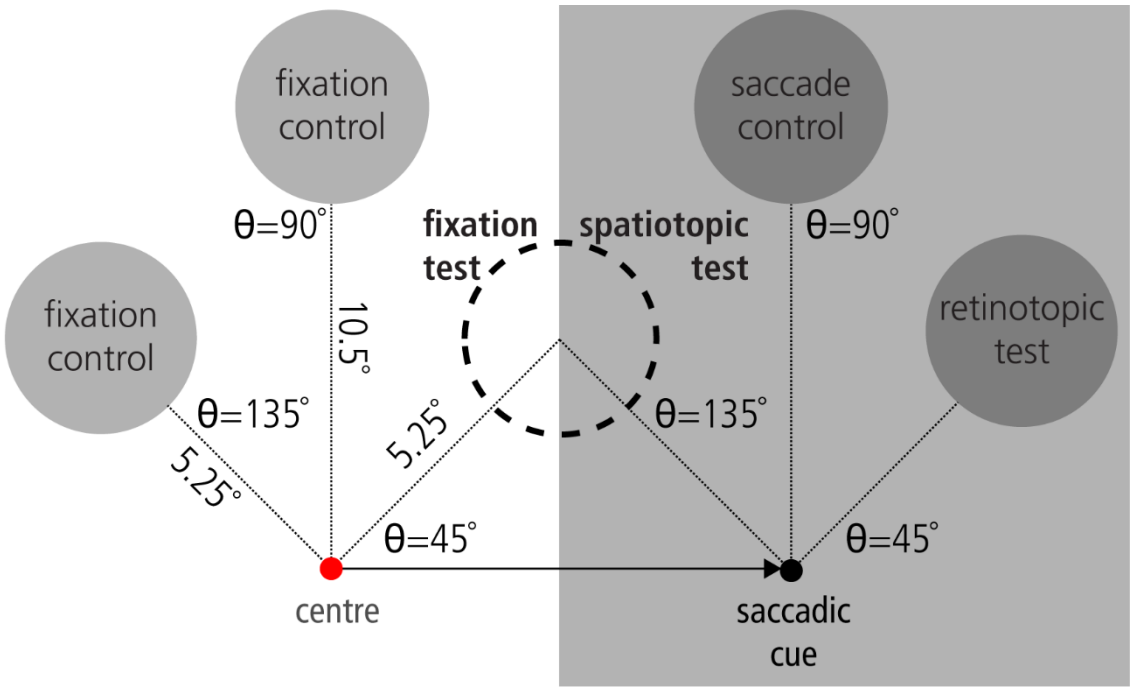

Figure 17 The task geometry relative to central fixation and fixation after a right saccade. On the left, positions tested under fixation trials. Relative to central fixation point (red dot), the acquisition position (dashed line) is at the right, other positions are considered controls. Spatiotopic and retinotopic reference frames are confounded under the Fixation condition for all three positions. On the right of the figure (grey background), positions after a rightwards saccade from central fixation to right fixation point has been executed. This time relative to the right fixation which served as saccadic cue, the position which used to be at the acquisition position becomes now the retinotopic test position (right, angle 45°). Also relative to new fixation the position to the left is the spatiotopic (at 135°). After saccade execution, the fixation test matches spatially the spatiotopic test.

Acquisition Phase

As in EXPERIMENT 1 {p.49}, the purpose of this phase was to create action-effect associations. The action was a left- or a right-hand movement using the sticks of a gamepad, and the effect was the presentation of an RDK. Subjects fixated the central fixation dot and generated the action, which triggered the appearance of an RDK (see FIGURE 16 {p. 70}) which had 100% coherent upwards or downwards motion. The location at which the RDK appeared —the *acquisition position*— was fixed

for one subject, and counterbalanced across subjects. The action-effect association (e.g. left stick – upwards motion) was also fixed for one subject and counterbalanced across subjects. The delay between action and effect was determined individually as described in REAL SACCADE LENGTH {p. 73}.

Central fixation was checked throughout the duration of RDK presentation. When broken, subjects received feedback and the trial was rerun. Subjects were instructed to generate an equal number of left and right actions. Whenever the difference between left and right actions was greater than 8%, subjects were reminded to balance their choices as close as possible to 50%–50%.

As in EXPERIMENT 1 {p. 49}, this phase included 12% of catch trials, in which dot motion was horizontal. Subjects' task was to press any button on the gamepad as soon as possible when they detected horizontal movement, whatever its direction. Catch trials ensured that subjects were paying attention to the motion direction at their acquisition position. Performance on catch trials was 87.5%±20, suggesting that subjects were paying attention to the RDK's motion at the acquisition position.

Test phase

Each acquisition phase was always followed by four tests: two fixation tests and two saccade tests. In fixation tests, subjects' action evoked the appearance of an RDK. The tested RDK could appear equiprobably at any of three locations (FIGURE 17 {p. 71}). This RDK's coherence was determined individually in a staircase procedure (see COHERENCE THRESHOLD for the previous experiment {p. 54}). The delay between action and RDK onset was also determined individually (see ACTION-EFFECT DELAY for the previous experiment {p. 54}). After the 275 ms duration of motion, the RDK was extinguished. Subjects' task was to report whether dot motion was upwards or downwards. In saccade tests, subjects were instructed to make their left/right action with the gamepad and to immediately make a saccade to the lateral fixation dot. Immediately after saccade end, i.e. with 0 ms delay, the RDK appeared at one of the same test positions as in the fixation condition but relative to the new fixation position. The side towards which subjects made a saccade was fixed, but left and right sides were counterbalanced among subjects. The saccade side was on the same side as the acquisition position (e.g. a subject who made saccades towards the right as illustrated in FIGURE 17 always had an acquisition position on the right {p. 71}).

In congruent trials, the action-effect association was maintained (e.g. left stick – upwards motion), while in incongruent trials the association was inverted (e.g. left stick – downwards motion). Whatever the congruency and the location of the RDK, the task was to discriminate dot motion direction (up versus down).

Real saccade length

Although subjects were instructed to perform a saccade from central fixation dot to a peripheral dot, saccades real amplitude tend to be about 10% shorter than stimulus's eccentricity (Kapoula and Robinson, 1986; Toyomura, 2003). With this in mind, we wanted to match the exact retinotopic position of the coherent RDK with respect to the actual saccade landing position. To do this, during the ACTION-EFFECT DELAY {p. 54} pre-test we also measured the actual saccade amplitude to account for the undershoot. On saccade trials, the RDK position was shifted by the measured difference, ensuring the retinotopic location to be exact relative to landing position.

1.2.2 Hypotheses

Based on previous findings (Desantis et al., 2014), we hypothesized that congruency would influence performance: the direction of motion would be better discriminated when its direction was congruent with the acquisition direction than when it was incongruent. Concerning the effect of tested position, in the fixation condition we expected to find better discrimination sensitivity at the acquisition position relative to control positions. In the saccade condition, we had no specific hypotheses as to whether the change in sensitivity at the acquisition position would manifest itself under the spatiotopic or retinotopic test location.

On response times (RT), we hypothesized that subjects would take more time to report the appearance of an incongruent trial than a congruent trial, and that the acquisition position would be reported faster.

1.2.3 Data analysis

Data was pre-processed in order to remove invalid trials. Trials were removed following the same criteria as for EXPERIMENT 1 {p. 49}. A trial was marked as `rejected` whenever the saccade towards the saccadic cue was shorter than half the measured saccade length (on average 3.42^{dva}, see REAL SACCADE LENGTH {p. 73}) plus the RDK radius (3^{dva}), yielding on average a cut-off of 6.42^{dva}. On average 0.83±1.8 trials were rejected per subject (<0.15% of the trials). To unify the acquisition positions across all subjects, we collapsed left and right saccades. On average, SRT was 134±80ms and average RT for responses was 587±339ms (from RDK offset to button press). Saccade duration was 57.43ms±46 on average. Overall, subjects executed 43% more of the 576 trials, thus the task contained 248 extra trials discarded because of one of the discard criteria.

1.2.4 Results

Discrimination performance (d') for each of the conditions is illustrated in FIGURE 18. Because we found no effect of congruency (t-statistic=1.284), d' is collapsed across congruent and incongruent trials. On average for fixation trials, d' at the acquisition position (dashed circle) was 1.23±0.6 (mean ± standard error of the mean), while d' at the two control positions averaged together was 0.95±0.36. The right part of figure shows sensitivity in the saccade condition. At the *retinotopic test* position (bottom right) d' was 1.15±0.8; at the *spatiotopic test* position (lower left) d' was 1.41±0.53; d' at the control position was 0.96±0.55.

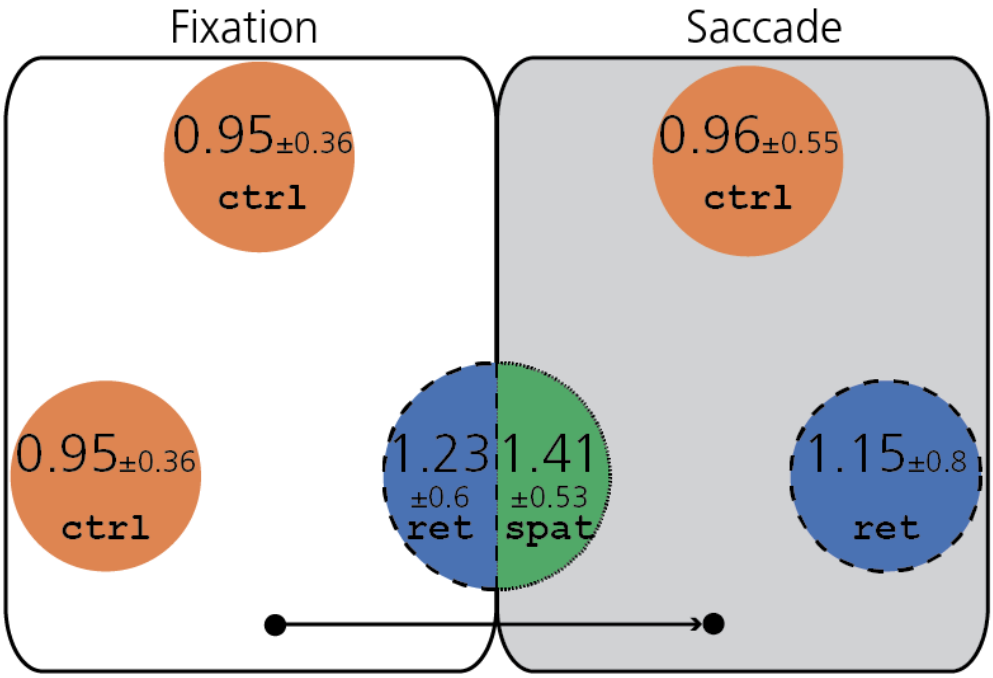

Figure 18 D-prime results for control vs test positions under the saccade (greyed area) and fixation conditions. For the Saccade condition, the spatiotopic position shows a significant increase in sensitivity compared to the retinotopic and the control positions. Fixation condition also displays this behaviour for the test vs control positions.

Mixed-Models

We analysed the data following the same procedure as in EXPERIMENT 1. We separated fixation and saccade trials using the maximal mixed-model approach previously described. Later we used the "joint" and the "expanded" approaches. Predictors were congruency and RDK position, and additionally saccade for joint mixed-models. In the fixation condition, there were 2 possible RDK positions (acquisition/test versus control positions), whereas in the saccade condition, there were 3 possible RDK positions (retinotopic test, spatiotopic test & control) as shown in FIGURE 17 {p. 71}. Mixed-model predictor variables were labelled in a similar way as for the previous experiment: `control` (the intercept), `ret` & `spat`. As described before, the acquisition position is labelled `ret`

in fixation trials, and in saccade trials labelled `ret` when it shared retinotopically the same coordinates as the acquisition position, and `spat` when it shared the same spatial coordinates as the acquisition position.

The fixed-effects for fixation trials (top of LMEM 6) displayed no effect of congruency (t-statistic=1.18), but there was a main effect of the retinotopic position (`ret`) (estimate=0.44, stderr=0.12, t-statistic=3.53, p-value<0.005): performance was higher at the acquisition location relative to the control location, in line with our hypotheses. Moreover, we found an interaction between position type and congruency (estimate=-0.30, stderr=0.14, t-statistic=-2.234 and p-value<0.05), indicating that subjects were worse by 0.3 at discriminating motion direction at the retinotopic position when trials were congruent. The random-effects for this mixed-model (displayed in LMEM 18 {p. 145}) revealed that the retinotopic position and congruency contained low variances ($\sigma^2 = 0.0741$ and $\sigma^2 = 0.07$, respectively) relative to the control ($\sigma^2 = 0.457$). This confirms that this mixed-model explained adequately the data, leaving $\sigma^2 = 0.056$ as residual variance.

Fixation	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.8675	0.1295	6.6982	0	***
ret	0.4381	0.1243	3.5263	0.0042	**
congruent	0.1579	0.1332	1.1857	0.1895	ns
ret × congruent	-0.3042	0.1361	-2.2349	0.0413	*
Saccade	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.9201	0.1597	5.7599	0	***
ret	0.2511	0.1752	1.4329	0.1409	ns
spat	0.4822	0.1515	3.1818	0.0059	**
congruent	0.0489	0.1871	0.2614	0.3792	ns
ret × congruent	-0.0989	0.1906	-0.519	0.3412	ns
spat × congruent	-0.0262	0.1906	-0.1373	0.3892	ns

LMEM 6 Fixed-Effects for fixation trials (top) and saccade trials (bottom).

In saccade trials (bottom of LMEM 6), there was again no effect of congruency (t-statistic=0.26). Performance was higher at both retinotopic (estimate= 0.25, stderr=0.17, t-statistic=1.433 & p-value=0.141) and spatiotopic (estimate=0.48, stderr=0.15, t-statistic=3.2 & p-value=0.0058) positions relative to control. However, the retinotopic increment was not significant. Interactions between positions and congruency showed that subjects performed better within incongruent trails, although these changes were not significant (t-statistics=-0.519 & -0.137). The difference between retinotopic and spatiotopic (refer to TABLE 4 {p. 150}) was not significant (estimate=-0.23113, t-statistic=-1.057, p-value>0.1). Random-effects (see LMEM 19 {p. 145}), on the other hand, show the spatiotopic position having very low variance ($\sigma^2 = 0.057$), while retinotopic and control positions,

and congruency were quite high. The high anti-correlation between `spat` and `ret` ($\rho = -0.79$) and their difference in variability, explains other positions' variances being high. The residual variance ($\sigma^2 = 0.089$) was of the same magnitude as the spatiotopic variance, indicating good numerical approximation of the data. A comparison between observed and predicted d' values is displayed in FIGURE 37 {p. 143}.

Expanded Mixed-Model

As for EXPERIMENT 1, we ran a mixed-model without averaging fixation positions. As depicted in FIGURE 17 {p. 71} and in FIGURE 19, the fixation `ret`inotopic test position is located to the right of the fixation dot (at angle 45°), the `control` is at angle 90° above the central fixation dot, and the third left of the fixation dot (at angle 135°) which has been labelled `fixspat`, similar to the saccade condition. Analysing how these three distinct positions influenced the observed d' value by means of another mixed model, unveiled interesting results.

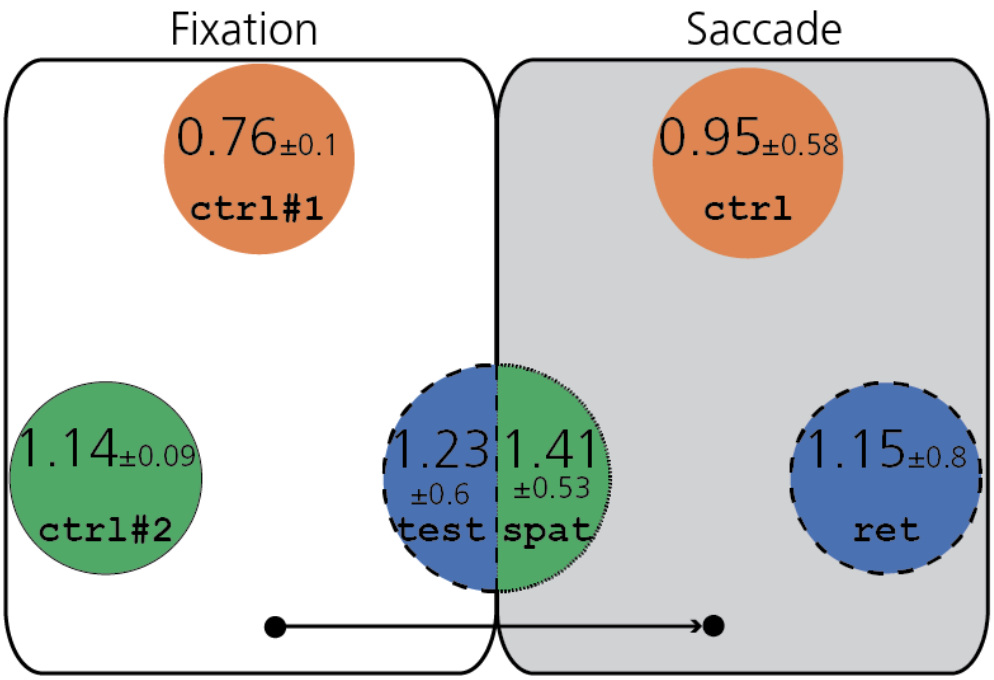

Figure 19 d' values for fixation and saccade trials with expanded fixation positions (not average together).

The fixed-effects for fixation trials (top of LMEM 7) exhibited the retinotopic position (`ret`) to contribute significantly to d' (estimate=0.65, stderr=0.13, t-statistic=4.85 and p-value=0.0001). The spatiotopic position (`fixspat`; previously averaged) also provides a (less, but yet) significant effect to the observed sensitivity (estimate=0.42, stderr=0.16, t-statistic=2.62 and p-value<0.05). Given that we established that t-statistics whose absolute value greater than 2 would be significant, this position is indeed explaining subjects' sensitivity. These results suggest that subjects are sensible enough to a position that is —theoretically— a control, when no saccade is present. Additionally, the

differences in estimates (see TABLE 5 {p. 150}) showed that `fixspat` (labelled `spat` in the table) and `ret` were not significantly different from one another (estimate difference=-0.2273; t-statistic=1.3257, p-value=0.162). The difference between `fixspat` and `control` was significant (estimate=0.4247; stderr=0.1624; t-value=2.616; p-value=0.0187). In this sense, `fixspat` is more likely to be considered a test than a control position. The "expanded" results just mentioned are plotted in FIGURE 19 {p. 76}, similar to the plots in FIGURE 18 {p. 74} but for the expansion of fixation positions.

Fixation	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.6536	0.1627	4.0176	0.001	***
ret	0.652	0.1345	4.847	0.0002	***
spat	0.4247	0.1624	2.616	0.0187	*
congruent	0.2184	0.1417	1.5413	0.1211	ns
ret × congruent	-0.3647	0.1791	-2.036	0.0552	ns
spat × congruent	-0.1004	0.1791	-0.5606	0.3334	ns
Saccade	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.9201	0.1597	5.7599	0	***
ret	0.2511	0.1752	1.4329	0.1409	ns
spat	0.4822	0.1515	3.1818	0.0059	**
congruent	0.0489	0.1871	0.2614	0.3792	ns
ret × congruent	-0.0989	0.1906	-0.519	0.3412	ns
spat × congruent	-0.0262	0.1906	-0.1373	0.3892	ns

LMEM 7 Mixed-model for fixation trials in the expanded version.

The random-effects for the expanded fixation trials (LMEM 21 {p. 146}) displayed the same behaviour as in the simple mixed-model that considers only control and test positions (LMEM 4 {p. 64}).

Joint Mixed-Model

We ran the mixed-model with fixation and saccade trials together, with averaged fixation positions (not expanded), results are shown in LMEM 8. Performance increases at the retinotopic position (estimate=0.438, stderr=0.161, t-statistic=2.716, p-value=0.014), while the spatiotopic shows even higher estimates (estimate=0.482, stderr=0.157, t-statistic=3.063, p-value=0.006). As before, congruency had no significant role in determining subjects' performance. Neither did saccade display a main effect (t-statistic=0.35). None of the interactions was significant; and the previous interaction between the retinotopic position and congruency found for fixation trials appears to be broken under saccade trials, because in this mixed-model such interaction is not significant (t-statistic=-1.575). The random-effects (LMEM 20 {p. 145}) shows minimal variances for position type (`ret`: $\sigma^2 = 0.088$,

spat: $\sigma^2 = 0.074$), congruency ($\sigma^2 = 0$) and saccade ($\sigma^2 = 0.047$). Residual variance is drastically high ($\sigma^2 = 0.98$) compared to other predictors. We also found a high correlation ($\rho = -0.9$) between congruency and the spatiotopic position, which can explain the low variance of congruency.

	estimate	stderr	t-value	p-value	sign
(Intercept)	0.8675	0.136	6.3836	0	***
ret	0.4381	0.161	2.7159	0.0136	*
spat	0.4822	0.157	3.0629	0.0062	**
congruent	0.1579	0.167	0.9467	0.2501	ns
saccade	0.0526	0.15	0.3497	0.3709	ns
ret × congruent	-0.3042	0.193	-1.5747	0.1155	ns
spat × congruent	-0.0262	0.193	-0.1355	0.3914	ns
ret × saccade	-0.1871	0.193	-0.9683	0.2449	ns
congruent × saccade	-0.109	0.193	-0.5644	0.3352	ns
ret × congruent × saccade	0.2053	0.273	0.7514	0.2957	ns

LMEM 8 Mixed-model joining saccade and fixation trials.

Joint Expanded Mixed-Model

Furthermore, the expanded version of fixation trials is corroborated when taking fixation and saccade trials together to run the mixed-model using the expanded fixation trials, as shown in LMEM 9.

	estimate	stderr	t-value	p-value	sign
(Intercept)	0.6536	0.157	4.1649	0.0004	***
ret	0.652	0.165	3.9525	0.0006	***
spat	0.4247	0.165	2.5771	0.0181	*
congruent	0.2184	0.17	1.284	0.1725	ns
saccade	0.2664	0.158	1.6923	0.0963	ns
ret × congruent	-0.3647	0.213	-1.713	0.0932	ns
spat × congruent	-0.1004	0.213	-0.4716	0.3525	ns
ret × saccade	-0.401	0.213	-1.8835	0.0701	ns
spat × saccade	0.0575	0.213	0.2699	0.3808	ns
congruent × saccade	-0.1695	0.213	-0.796	0.2858	ns
ret × congruent × saccade	0.2657	0.301	0.8826	0.2656	ns
spat × congruent × saccade	0.0742	0.301	0.2465	0.3832	ns

LMEM 9 Mixed model combining saccade and fixation trials, and differentiating fixation positions.

The retinotopic position shows an increment of 0.65 with respect to the control (estimate=0.652, stderr=0.165, t-statistic=3.953, p-value=0.0006). The spatiotopic a 0.42 increment relative to the retinotopic position (estimate=0.425, stderr=0.165, t-statistic=2.577, p-value=0.018). As in previous results, congruency was not significant. The saccade condition increments by 0.26 the estimated

performance, however this increment was not significant. None of the interactions was significant. The random-effects displayed the same behaviour for the random slopes as in the expanded fixation trials (LMEM 22 {p. 146}). Differences between position predictors are shown in TABLE 6 {p. 150}.

Response Times

Response time was the delay between RDK offset and subjects' response (indicating direction of motion upwards or downwards using the gamepad). On average, subjects took 587±339 ms to report motion direction. Overall, subjects reported fixation trials faster than saccade trials (fixation=564.6 vs saccade=611.2; t-statistic=-5.67, p-value=0). Subjects also took more time to report incorrect trials than correct ones (incorrect=607.4 vs correct=578.8; t-statistic=3.25, p-value=0.001). As for performance, we ran two mixed-models for fixation and saccade trials separately, an expanded version of the fixation trials, the joint mixed-model, the joint expanded mixed-model and as in the previous experiment another mixed-model in which reference frame and test conditions were predictor variables.

In the mixed-models estimating RT for fixation and saccade trials, none of the predictors had a significant effect (main effect or interaction) on RTs (see LMEM 23 {p. 146}). The expanded version (see LMEM 24 {p. 146}) behaved in the same way. The joint mixed-model (LMEM 10 {p. 79}) found the saccade condition to significantly increase by 68 ms subjects' response times (estimate=68.45; stderr=20.18; t-value=3.391; p-value=0.0028).

	estimate	stderr	t-value	p-value	sign
(Intercept)	566.2724	47.97	11.804	0	***
ret	5.2148	21.98	0.2373	0.3838	ns
spat	-27.7026	25.93	-1.0685	0.2212	ns
saccade	68.4501	20.19	3.3906	0.0028	**
congruent	-3.4269	14.28	-0.2401	0.3835	ns
ret × saccade	11.9885	25.15	0.4766	0.3513	ns
ret × congruent	16.2432	23.3	0.6971	0.3077	ns
spat × congruent	21.3928	26.91	0.7951	0.2857	ns
saccade × congruent	-28.9571	23.32	-1.2417	0.1815	ns
ret × saccade × congruent	-18.6415	35.6	-0.5237	0.3429	ns

LMEM 10 Joint mixed-model for averaged fixation trials.

Next we executed the joint and expanded mixed-model (LMEM 11 {p. 80}). The results were like the averaged version of the joint mixed-model, with saccade adding 63 ms to subjects' response times (estimate=63.15; stderr=22.74; t-value=2.78; p-value=0.0116).

	estimate	stderr	t-value	p-value	sign
(Intercept)	571.5651	48.95	11.6756	0	***
ret	-0.0598	26.02	-0.0023	0.3954	ns
spat	-10.5708	24.32	-0.4347	0.3587	ns
saccade	63.1465	22.73	2.7775	0.0116	*
congruent	-3.6607	19.53	-0.1875	0.3883	ns
ret × saccade	17.3137	26.87	0.6443	0.3194	ns
spat × saccade	-17.1066	26.9	-0.636	0.3212	ns
ret × congruent	16.4552	26.89	0.6119	0.3261	ns
spat × congruent	0.4108	26.91	0.0153	0.3953	ns
saccade × congruent	-28.7483	26.91	-1.0684	0.2215	ns
ret × saccade × congruent	-18.8667	38.03	-0.4961	0.3483	ns
spat × saccade × congruent	20.9943	38.04	0.5519	0.338	ns

LMEM 11 Joint mixed-model for expanded fixation trials.

Finally, we ran a mixed-model (LMEM 12) for RTs using the reference frame vs test/control labelling previously used. Participants took 45 ms longer to report test positions compared to controls (estimate=45.45; stderr=20.5; t-value=2.2; p-value=0.04). The reference frame itself induced an increment of 68 ms when it was spatiotopic as compared with the retinotopic (estimate=68.43; stderr=18.5; t-value=3.69; p-value=0.002). Finally, there was an interaction between the tested position type (test vs control) and the reference frame (retinotopic vs spatiotopic) (estimate=-73.17; stderr=23.35; t-value=-3.13; p-value=0.006).

	estimate	stderr	t-value	p-value	sign
(Intercept)	566.2817	48.1	11.7739	0	***
test	45.4598	20.56	2.2106	0.0399	*
ref_frame	68.427	18.52	3.6941	0.0018	**
congruent	-3.4428	14.05	-0.2451	0.3815	ns
test × ref_frame	-73.169	23.35	-3.1342	0.0061	**
test × congruent	-7.5331	19.08	-0.3949	0.3628	ns
ref_frame × congruent	-28.9535	23.38	-1.2383	0.1812	ns
test × ref_frame × congruent	28.9782	33.04	0.8771	0.2648	ns

LMEM 12 Fixed-effects for the mixed-model labelling trials by reference frame and test/control.

The random-effects for all the RT mixed-models presented in current section are available in the MOTION DISCRIMINATION EXPERIMENT {p. 145} section of the APPENDIX in LMEM 25 through 31.

1.2.5 Discussion

In this experiment we found that action-effect associations contained spatial information, because subjects discriminated motion better when their action caused the stimulus to appear at the previously learnt position (the acquisition position); they were less sensitive when the tested position was opposite to the acquisition, and virtually insensitive when tested at the control position.

Comparing fixation and saccade trials allowed us to determine to what degree spatial information was in the retinotopic and the spatiotopic reference frames. However, we also saw that retinotopic and spatiotopic positions were not significantly different from one another, under both conditions. It is clear that eccentricity was affecting sensitivity, because the distance between the fixation dot and the control positions is twice as large as that of the retinotopic and spatiotopic positions.

We saw that retinotopic and spatiotopic information both influenced performance. In fixation trials the acquisition position (`ret`) was higher than its opposite position (`fixspat`). Such pattern is repeated in saccade trials, where sensitivity at the acquisition position (now `spat`) was higher than the (opposite) retinotopic's (`ret`). This could reflect the pre-activation of spatiotopic information in saccade trials. However, as we saw in FIGURE 19 {p. 76}, these differences are not significant. In sum, we can see that a) the *physical acquisition position* shows better performance under both conditions (fixation `ret`: 1.23 & saccade `ret`: 1.41 in FIGURE 19) relative to b) their counterparts which are virtually equal (`fixspat`: 1.14 & `spat ret`: 1.15; t-statistic=-0.046, p-value=0.96). This pattern could signify that the forward model uses a spatial information in a specific order. Retinotopic then spatiotopic, when there is no eye-movement, and spatiotopic then retinotopic, when there is. However, we cannot conclude whether information fed into the forward model is purely retinotopic, purely spatiotopic or a mixture of both. Therefore, this hypothesis does not stand.

The most plausible explanation for the non-significant difference between `ret` and `spat` (both in fixation and saccade trials) is that the Action-Effect associations began to build up *before the acquisition phase*. This reminds us Prinz's ideas (1990) on action-effect associations exposed in COMMON CODING THEORY {p. 9}, which suggested that the rules that will forge AE associations are created *before the experiment begins*, when the experimenter explains the task to the participant. In our case, action-effect associations begin to be created in the coherence threshold pre-test (COHERENCE THRESHOLD {p. 54}) at which subjects were exposed to two different stimulus positions (`ret` and `spat`). In this way, regardless of the action that would be present in acquisition and test phases, participants' visual system began to 'be aware' of those two relevant locations. Later, in the test phases, motion was easily discriminated at such locations, and not at the control location which was

not part of the pre-test. Finally, the numerical difference found between `ret` and `spat` may be due to the fact that `ret` was presented more during acquisition phases.

Although the acquisition phase associated a left/right action to a specific motion direction, participants performed equally well in congruent and incongruent trials. We had hypothesised based on another discrimination experiment (Desantis et al., 2014) that subjects would be better in congruent trials, but this was not the case. This can be initially explained by the several differences between Desantis et al.'s experiment and the present one. Their RDK was presented foveally while ours was presented peripherally; their stimulus was presented at constant location and variable timing, while ours was presented at variable timing *and* variable position. Furthermore, the timing of the stimuli presentation could change from sub-block to sub-block (see ACTION-EFFECT DELAY {p. 54} for details), while in their experiment an auditory cue preceded the onset of the stimulus appearance. Finally, their RDK was presented with random motion for 100 ms after which coherent motion was presented for 100 ms, to finish with 100 ms of extra random motion; ours was presented constantly during the 275 ms at threshold coherence.

But regardless of the experimental differences, why did congruency not influence discrimination performance? Was motion acquired? And if so, was it pre-activated?

It is not clear if the pre-activation of the sensory effects contained or not motion information. The task itself may require little or no motion information. First, because localising the object was 'easy', as only one stimulus was present at each interval, which left the ventral system the task to discriminate direction of motion. This may be explained by the fact that neurons in the middle temporal visual area (MT/V5) have direction and motion selectivity (Dubner and Zeki, 1971; Maunsell and Van Essen, 1983), and that information processing in the ventral and in the dorsal pathways is shared (Milner, 2017) and concurrent (Nassi and Callaway, 2009).

Concerning the action-effect association, the `ret`×`congruent` interaction shown for fixation trials (LMEM 6 {p. 75}) suggests that motion discrimination was indeed influenced by congruency at the acquisition position. Subjects' performance *decreased* by 0.3042 when congruent trials were presented at the acquisition position, i.e. when acquisition conditions were met[15]. This confirms the presence of motion in the Action-Effect association at the retinal acquisition position. However, the same cannot be confirmed for saccade trials because such interaction was not significant, and because of the high correlation between spat and ret ($\rho = -0.79$) combined with congruency's high

[15] Exactly as in the acquisition phase, i.e. in motion direction and location.

standard error ($\sigma^2 = 0.20$) in the mixed-model's random-effects (APPENDIX: LMEM 19 {p. 145}). Moreover, the "expanded" and the "joint" versions of the mixed-model showed this interaction to be non-significant.

The pre-activation hypothesis, which explained results of motion detection experiments, can partially explain the present results. Basically, at the beginning of the experiment both signals' distributions (upwards and downwards) are far away from the baseline (which is assumed reflects the activity without signal, and probably not exactly at zero level). As the acquisition phase progresses, action-effect associations start to build up; Sensory networks of the effects become pre-activated before each action, representing the predicted or expected consequences. After a while A-E associations are so strongly embedded that the pre-activated sensory networks reach a level very close to the actual signal, what others have been calling the "pedestal level" (Green and Swets, 1966; Cardoso-Leite et al., 2010; Desantis et al., 2014; Roussel et al., 2013; Waszak et al., 2012). If we consider that the difference between pre-activated and actual signals represents performance, the pre-activation hypothesis predicts that `ret` position pre-activation signal should have reached a "pedestal" level closer to the actual `ret` signal rendering it less easy to discriminate motion at, while not for other positions. Nonetheless, our results do not show such pedestal level pre-activation neither for `ret` nor `spat`, because discrimination is good at the retinotopic position *and* its opposite (`fixspat` or `spat`). This seems to contradict the pre-activation hypothesis, which predicts the contrary. This is clearly a design mistake. In order to isolate this effect, the coherence pre-test should have tested only at ret (the acquisition position), or the Action-Effect associations induced during the acquisition phase should make each action reflect two effects (upwards or downwards motion) at two locations (`ret` or `spat`).

Moreover, the observed behaviour took place at the first stages of the experiment in testing phases. In fixation trials, the retinotopic position displays a significantly higher performance as opposed to the two controls. This was true from the second until the sixth fixation test blocks, after which discrimination at the acquisition position decreases close to the spatiotopic (second control) levels, as shown in FIGURE 20. This was not the case for saccade trials, where performance up to the fifth test block is virtually equal for all three positions; after the sixth block performance at the spatiotopic position maintains stable in a plateau, while other positions decrease to the same level, as shown in FIGURE 35. A combination of the previously mentioned behaviour is found when trials are taken without making the distinction between fixation and saccade. The retinotopic (retinal acquisition position) increases after the first acquisition phase reaching almost a significant difference with other positions, while the other two positions reach a plateau that slightly increases across the experiment.

At the fifth block, performance at the retinotopic location decreases to finally join the performance levels of the other two positions, as displayed in FIGURE 36.

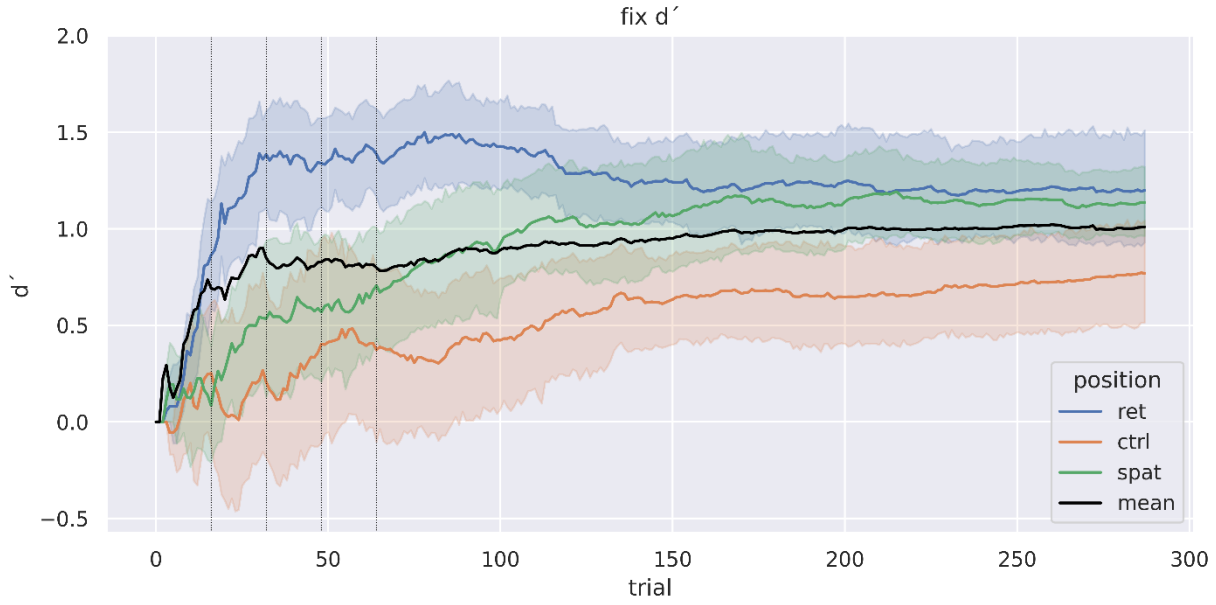

Figure 20 Performance across the fixation trials. In the first stages of the experiment, it is easier for subjects to discriminate motion at the acquisition position; moreover, performance is significantly different at this position, as opposed to the controls. After the sixth block, performance at this position begins to decay until reaching the same level of the spatiotopic's (second control) performance.

Finally, the facilitated speed with which subjects reported the spatiotopic position in saccade trials, reflects that spatiotopic information is more easily accessed by the predictive system, or that it is preferred. In any case, RTs do make a clear distinction as to which information is primarily used: with an eye movement, spatiotopic information is more quickly accessed. The confound in the fixation condition may be due to the fact that the acquisition phase did not involve a saccade action.

1.3 Discussion

In the current study we tried to determine if spatial information —in the form of effect within an AE association— is acquired by the visual system during an acquisition phase. This was determined by testing if such spatial information was pre-activated as prediction of the effects of an action and thus when changed, the system would reflect the mismatch. In this general discussion, we will try to explain results based on Forward Models.

In the MOTION DISCRIMINATION EXPERIMENT we implemented a 2-AFC task in which subjects' actions had as consequence the appearance of a single RDK at one of three equiprobable positions, inversing on

half the trials the acquired motion direction. We saw that subjects were very good at discriminating motion at the acquisition position, both with a saccade involved or not.

Studies have found regions of the brain that organise visual information in spatiotopic representations (Colby and Duhamel, 1996; Merriam and Colby, 2005), especially when the stimulus is complex or involves motion (Melcher, 2005; Melcher and Morrone, 2003). These findings, of course, do not deny the fact that retinotopic information is indeed used by lower-level regions of the brain. One can safely assume that when a prediction does not involve an eye movement, retinotopic information is used, unless a broader world-centred point of view is required. Despite the design issue discussed in the experiment's discussion, performance scores shed some light into the idea that both types of information are actually used. In fixation trials, we saw the sum of both types of information at the acquisition position, where subjects were very good at discriminating motion. The other relevant position, which we labelled `fixspat`, showed that subjects were also good at discriminating motion, but given that this position does not mirror the acquisition conditions, the predictive system has only access to spatiotopic information. Which means that the pre-activated signal is based on the retinal signal. On the other hand, with an intervening saccade, the predictive system reflects the same pattern of information usage. When tested at the retinotopic position, subjects are good at deciding whether motion is upwards or downwards, such that retinotopic information is used based on the learning phase. Nevertheless, subjects are even better at discriminating motion at the physical acquisition position (labelled `spat` in FIGURE 19 {p. 76}). However, at this position—in principle—the pre-activated effects are based only on acquisition-phase effects, thus retinotopic. The observed behaviour may be caused by the participation of a higher-level mechanism intervening whenever there is a change in gaze direction, such that retinotopic information is *transformed* into spatiotopic information[16], and thus, its usage is facilitated and provided to the predictive system. If such transformation indeed takes place, RT results point out the system's ability to transform it when it is required, as well as the resource overload required to make the transformation.

The observed results seem difficult to explain purely based on Forward Models if we do not have further knowledge about the entity in charge of making the comparison between actual and predicted effects (the discrepancy unit in FIGURE 3 {p. 15}). Interpreting exclusively the Forward Model's 'behaviour', we can deduce that given $any\ motor\ command$ (i.e. left or right stick in the gamepad) the predicted sensory feedback includes the retinotopic and the spatiotopic positions; with a slight

[16] If retinotopic information is indeed 'transformed' on-demand; it could be that both types of information are available with different priorities.

preference for the retinotopic without an intervening saccade; and a slight preference for the spatiotopic with an intervening saccade. Once again, such 'behaviour' may be caused by the pre-test which exposed the FM to both locations. In any case the difference with and without a saccade strongly suggests that the efference copy that the Forward Model receives includes eye movement information, hence the observed reference frame influence.

Concerning congruency, it appears that spatial information has higher priority than motion information. First, the task's intervals change according to position and a direction of motion. The Action-Effect association, nonetheless, only involves the effect of motion, because in the acquisition phase stimulus's location did not change. Given that subjects reported based on motion, we are not able to tell if position was as acquired as motion. Results, however, show that—regardless of whether it was acquired or not—motion congruency information is not needed or pre-activated as part of the action effects. Whether peripheral motion information is acquired or not remains to be tested. This could be done in an experiment in which the acquisition phase involved the creation of Action-Effect associations occurring spatiotopically *and* retinotopically, in other words, involving a saccade. Such an experiment would allow us to further distinguish under which reference frame A-E associations happen.

In the COHERENT MOTION DETECTION EXPERIMENT subjects had to detect the location of the a coherent RDK among four possible locations, making it a 4-AFC task. The acquisition phase, in contrast to the discrimination experiment, did include the simultaneous presentation of 4 stimuli.

Results showed very well the Sensory Attenuation phenomenon, because when tested at the acquisition position, subjects performed very poorly; that, regardless of whether a saccade was involved or not. The pre-activation hypothesis was confirmed. However, the interaction of spatial information was not as clear as in the MOTION DISCRIMINATION EXPERIMENT, mainly because non-acquisition positions displayed similar activations in saccade trials and different activations in fixation trails, with the attenuated effect present in both.

It would, thus, appear as if Sensory Attenuation depended also on the saccade, because results were saccade- and position-specific. If this selectivity is indeed happening, it would suggest that the Forward Model's predicted effects with a saccade are more pre-activated that without a saccade, which to our knowledge no studies have found. One might think that attentional resources are preferentially allocated at the retinotopic location especially after the saccade's execution. On the

other hand, divided attention[17] —and attention in general— seem to play a very important role in explaining participants' and the Forward Model's observed 'behaviour'.

[17] See "A Divided Attention Approach" in the general DISCUSSION {p. 115}

2 Saccadic Suppression Experiments

In this series of experiments, we are interested in the mechanisms that support transsaccadic correspondence—in other words, the mechanisms that match the various retinal locations of a given visual object across a saccade, giving rise to a unified and stable percept. There have been several theoretical proposals as to how this matching process occurs as described in Visual Stability and Transsaccadic Correspondence {p. 36}.

In the present study, we combined the in-flight displacement task with a feature change. In five experiments, participants had to make a saccade to a peripheral Gabor patch. During saccade execution, the patch could be displaced, its orientation changed, or both. In the first experiment (1A {p. 89}), we replicated saccadic suppression of displacement with the Gabor stimulus (no orientation shifts occurred). In the second and third experiments (1B and 1C {p. 94}), participants reported in which direction the patch had stepped, ignoring any orientation shifts. In the fourth and fifth experiments (2A and 2B {pp.102 & 105}), participants reported in what direction the orientation had shifted, while the stimulus displacements were task-irrelevant. In all experiments, we compared blank against no-blank conditions.

If matching features in VSTM drives visual stability, then disrupting visual stability with an orientation shift should expose transsaccadic displacements; we should observe an inverse relationship between displacement discrimination and orientation shift size (Experiments 1B and 1C {p. 94}). We compared the change with orientation shift to the change due to blanking, but had no specific hypotheses about any difference in magnitude between the two.

If a spatiotemporal remapping signal drives visual stability, then breaking visual stability with a displacement should unmask transsaccadic changes in orientation (Experiments 2A and 2B). Tas, Moore, and Hollingworth (2014) showed that when object continuity was broken by a blank, subjects were more able to correctly report the colour of the pre-saccadic object. Without a blank, they tended to report the colour of the post-saccadic target, as if the post-saccadic features had overwritten the pre-saccadic features (Tas et al., 2014, 2012). This idea is akin to masking: Post-saccadic features may mask pre-saccadic features, but this depends on the temporal proximity between the two sets of features. In our setup, breaking object continuity with a displacement should allow pre- and post-saccadic features to be accessed and compared better; we should observe an inverse relationship between the thresholds for orientation shift discrimination and displacement size. Again, we had no specific hypotheses about a difference in the magnitudes of change due to blanking or displacement.

2.1 Experiment 1A

Displacement (report displacement)

2.1.1 Methods

Subjects

Nine healthy human adults (six men, three women, 33 ± 12 years old) participated. All reported normal or corrected-to-normal vision and no oculomotor or neurological impairments, received payment of 10€/h for their time, and provided informed consent.

Stimuli and Procedure

The fixation stimulus was a Gabor patch with a diameter of 1 degree of visual angle (dva) and a spatial frequency of 4.5 cycles per degree (cpd). Its contrast was 100%, and the patch was presented on a light grey background (30 cd/m2). FIGURE 21 {p. 90} illustrates the experiment's geometry. Initially, the Gabor patch was randomly presented at position F inside an imaginary 2.45^{dva}-sided square at screen centre. Its orientation was always vertical. Once the eyetracker had detected fixation, the Gabor patch remained visible for 250–500 ms, after which it disappeared, to immediately reappear at position T either 6 or 8^{dva} to the left or right. Subjects were instructed to follow the target by making a saccade from F to T. At 20% of the distance between the saccade's onset position and the target's position, the target disappeared and reappeared in position TD, which was T displaced by one of eight possible magnitudes (represented as small squares in the figure; ±2^{dva}, ±1^{dva}, ±0.5^{dva}, or ±0.25^{dva}). Reappearance of the Gabor patch occurred either immediately (no-blank condition) or after a 250 ms delay, during which no stimuli were present on the screen (blank condition). The stimulus remained in its final position TD for 600 ms. Subjects were instructed to report whether the displacement from position T to position TD had been to the left or to the right, tapping the corresponding arrow key on the keyboard.

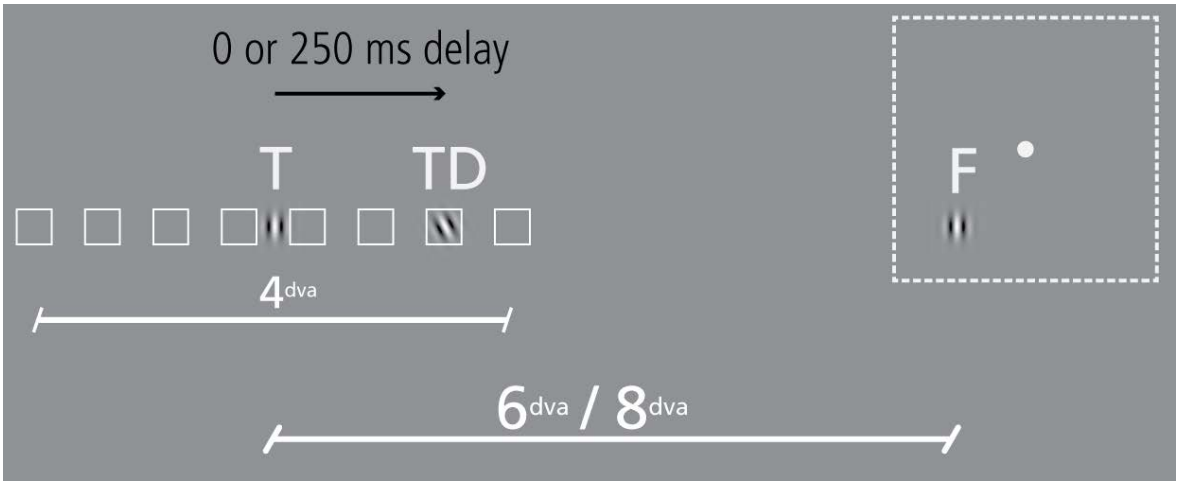

Figure 21 Example of a trial with a leftward saccade. Initially, the fixation Gabor is displayed within a cantered dashed square at position `F` *(the square is displayed for illustration purposes, and the white dot represents the centre of the screen). The fixation stimulus was replaced by a target stimulus either 6 or 8dva to the left at position* `T`*. When a saccade toward the target was detected, the target disappeared and was replaced, either immediately (no-blank condition) or 250 ms later (blank condition), by a stimulus at position* `TD` *(one of the eight final locations, shown as small squares). The subject's task was to report the horizontal direction of the final target step, responding either "left" or "right".*

The experiment was carried out in a randomized factorial design, with two saccade directions (left/right), two saccade amplitudes (6/8dva), eight target displacements (± 2.0, ± 1, ± 0.5, ± 0.25dva), two onsets (no blank, blank), eight repetitions in four blocks, yielding 520 trials per participant, divided into four blocks of 128 trials. Prior to the first block, participants performed a 20-trial practice, to get accustomed to performing the saccade and perceptual tasks.

We hypothesized that displacement discrimination performance would be better in the blank than in the no-blank condition, replicating previous work (e.g., Deubel et al., 1996). We examined the slope and the point of subjective equality (PSE, which corresponds to the stimulus level at 50%). We hypothesized that the PSE would depend on blanking, with PSEs being closer to zero in the blank than in the no-blank task, although this effect has been less robustly reported in the literature (Deubel et al., 1996; Wexler & Collins, 2014). Note that a negative PSE corresponds to a forward bias (see the DISCUSSION section {p. 112}). Additionally, we computed the threshold which was defined the difference in stimulus level between the 75% and the 50% points of the fitted psychometric curve.

Apparatus

Stimuli were presented on a Sony GDM-F520 CRT monitor with 1024 × 768 (width by height) pixel resolution at an 85-Hz refresh rate inside a 503×377 mm area. Responses were given by subjects via a standard keyboard. Saccadic eye movements were recorded using an EyeLink 1000 (SR Research Ltd., Mississauga, Ontario, Canada) operated in monocular mode at 1 kHz, tracking the right eye, with the EyeLink toolbox extension for Matlab (Cornelissen, Peters, & Palmer, 2002). Participants used forehead and chin rests and viewed 57 cm away from the screen. Stimuli were generated using the Psychophysics Toolbox extensions for Matlab (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007).

Eye Movement Acquisition

Viewing was binocular, and movements of the right eye were monitored with an EyeLink 1000 (SR Research, Mississauga, Ontario, Canada) at a 1kHz sampling rate. At the beginning of each experiment and each block, the eyetracker was calibrated using the standard nine-point EyeLink procedure.

Saccades were detected online when the horizontal component of gaze differed by more than 20% of the F–T distance (see FIGURE 21 {p. 90}). The eye movement traces were stored for offline analysis. For the offline analyses, the eye movement samples were smoothed with SR Research's proprietary algorithms. Instantaneous velocity and acceleration were computed for each data sample and compared to a threshold (30°/s and 8000°/s2). Saccade onset was defined as two consecutive above-threshold samples for both criteria. Saccade offset was defined as the first sample of a 20-ms period of below threshold samples.

We checked that gaze-contingent changes occurred during the saccade. On average, saccade duration was 43±3 ms, which left ample time for the saccade to be detected and the gaze-contingent change to occur within the ≈12 ms granted by an 85-Hz refresh rate. We eliminated trials in which the eye movement occurred too early (% eliminations given in the DATA ANALYSIS section of each experiment) or too late (< 1% for all experiments).

2.1.2 Hypothesis

We hypothesized that performance would be better in blank than in no-blank conditions, as expected from previous experiments.

2.1.3 Data Analysis

All data analyses were performed in `Python` (G. Van Rossum, 2003), `Matlab` (The MathWorks, Natick, 2014) and `R` (R Core Team, 2016). Perceptual performance was analysed by fitting individual psychometric functions to the proportion of "forward" responses. The difference in displacement size between the 0.5 and 0.75 points on the psychometric curve represented the perceptual threshold. We also examined the PSE and slopes. All fits were performed on individual data. The reported means are based on these individual fits averaged across subjects.

To assess the statistical significance of our effects, we used confidence intervals obtained by bootstrapping. We also assessed the evidence in favour of the null versus alternative hypotheses using Bayesian t-tests (Rouder, Speckman, Sun, Morey, & Iverson, 2009) with the `BayesFactor` `R` package (Morey & Rouder, 2015; R Core Team, 2016).

We eliminated trials on the basis of eye movement behaviour, so as to keep trials in which subjects performed the saccade task correctly. Subjects sometimes initiated the eye movement before the appearance of the target; we did not include those trials in the analysis (12%). We also eliminated all trials in which saccade amplitude or latency was more than two standard deviations away from the average (< 1%).

2.1.4 Results

The basic saccade characteristics were normal for visually guided saccades and did not differ between the blank and no-blank conditions. The average latencies were 180 ms (95% bootstrapped confidence interval: [162–201 ms]) in the blank condition and 177 [162–190] in the no-blank condition. The saccade amplitudes were 6.8 [5.9–7.7] in the blank condition and 6.8 [5.8–7.7] no-blank condition.

Figure 22A {p. 93} shows individual psychometric functions as well as the averaged resulting psychometric function for all subjects, and FIGURE 22B shows average PSEs, slopes and thresholds in the blank and no-blank conditions. On average, performance was better in the blank condition (threshold of 0.53 dva, 95% confidence interval [0.41–0.65]) than in the no-blank condition (1.10 [0.73 to 1.48] dva). A Bayes factor analysis comparing the hypothesis that performance was better in blank than in no-blank conditions to the null hypothesis of equal thresholds between blank and no-blank conditions revealed evidence in favour of the alternative hypothesis (BF = 1.78). Slopes

decreased under the blank condition, the same analysis comparing slopes yielded evidence in favour of the model that hypothesised a change in slopes (BF = 3.42).

The PSE also depended on blanking, with negative PSEs in the no-blank condition (– 1.23 [– 1.7 to – 0.74]) and PSEs close to zero in the blank condition (0.08 [– 0.23 to 0.38]). The Bayes Factor was 5.7—evidence in favour of the alternative hypothesis of more negative PSEs in the no-blank than in the blank condition, over the null hypothesis of no difference in PSEs between blank and no-blank. In sum, EXPERIMENT 1A showed that we were able to replicate the classic blanking effect in our setup.

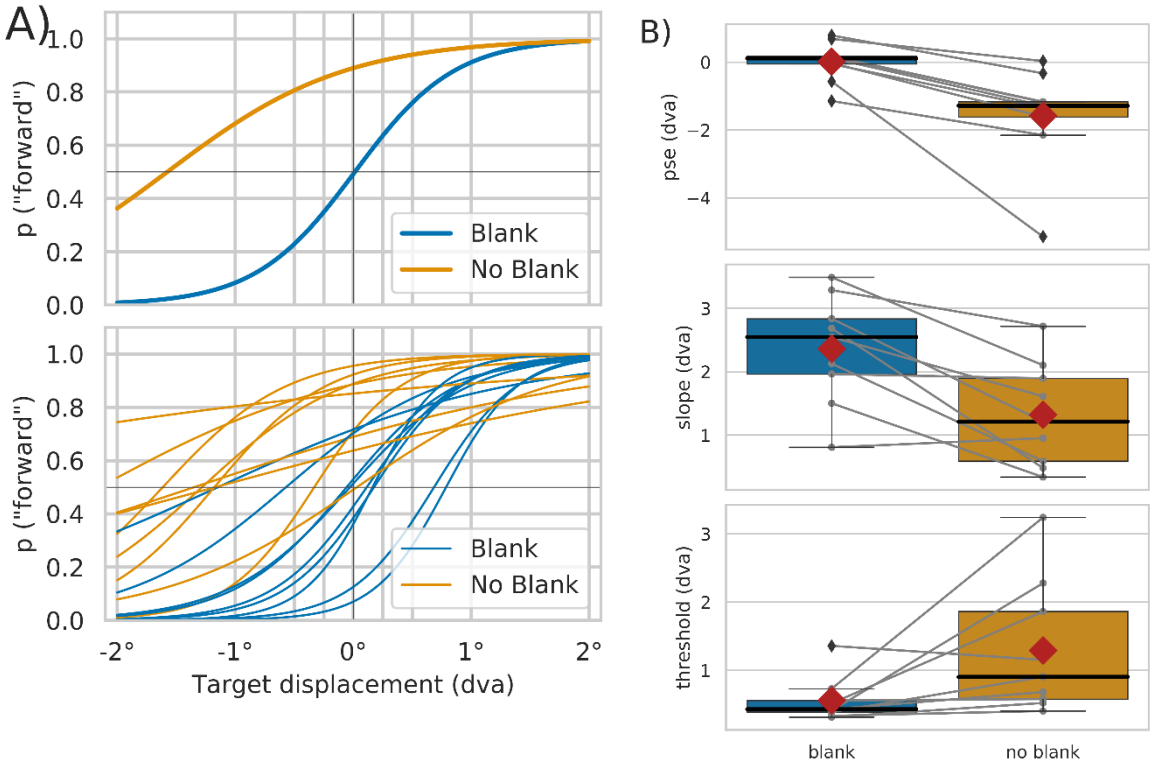

Figure 22 ***A)*** *Top, psychometric function for averaged PSE and slopes across subjects comparing blank (blue) and no-blank (orange) conditions. Ordinate axis shows the probability of a "forward" response given the displacement size (abscissa). Bottom, individual psychometric functions for all subjects comparing the blank (blue) and no-blank (orange) conditions.* ***B)*** *Mean PSE (top), mean slopes (centre) and mean thresholds (bottom) comparison for the blank (blue) and no-blank (orange) condition. Individual subjects are represented by dots under each condition and connected between conditions. Black thick lines are the medians; red diamonds represent the means.*

We also ran three different mixed-models on the experiment data (see MIXEDMODEL 1 {p. 94}). The no-blank condition serves as an intercept. For all three predicted variables (PSE, slope and threshold) the blank condition affected significantly subjects' sensitivity. PSEs (estimate=1.174; stderr=0.12; t-value=9.68; p-value=0) and slopes (estimate=0.689; stderr=0.25; t-value=2.76; p-value=0.026) showed an increment when blanking and thresholds decreased (estimate=-0.145; stderr=0.086; t-value=-1.68; p-value=0.099) in a non-significant manner.

pse	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	-1.1579	0.2139	-5.4135	0.002	**
blank	1.1737	0.1213	9.6768	0.000	***
slope	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	1.6708	0.2735	6.1088	0.001	**
blank	0.6895	0.2503	2.7553	0.026	*
threshold	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.7052	0.1035	6.8151	0.001	***
blank	-0.1448	0.0862	-1.6803	0.099	ns

MixedModel 1 Fixed effects results for three mixed-models. PSE (top), slope (centre) and threshold (bottom).

We then moved on to EXPERIMENT 1B, in which in addition to the intra-saccadic target displacement, we added an irrelevant orientation shift.

2.2 Experiments 1B & 1C

Displacement with orientation shift (report displacement)

2.2.1 Methods

Subjects

Seven healthy human adults (four men, three women; 35 ± 16 years old) participated in Experiment 1B, and seven in Experiment 1C (five women, two men; 27 ± 6 years old). All reported normal or corrected-to-normal vision and no oculomotor or neurological impairments, received payment of 10€/h for their time, and provided informed consent.

Stimuli and Procedure

EXPERIMENT 1B was identical to EXPERIMENT 1A, except the orientation of the Gabor patch at position F was not always vertical; instead, it was randomly selected from the range 45°–135°. The patch kept its orientation during the F to T displacement and underwent the orientation-shift simultaneously with the horizontal displacement (T to TD). Subjects were instructed to report the direction of the displacement (left/right) and to ignore any changes in orientation.

EXPERIMENT 1B was carried out in a randomized factorial design, with two saccade directions (left/right), two saccade amplitudes (6/8^{dva}), eight target displacements (±2^{dva}, ±1^{dva}, ±0.5^{dva}, ±0.25^{dva}), two onsets (no blank, blank), six orientation shifts (± 18°, ± 9°, ± 3°), and five repetitions, leading to a total of 1920 trials divided into 12 blocks. EXPERIMENT 1C was identical, except that there were only four orientations shifts (±45°, ±3°), leading to a total of 1280 trials, divided into 7 blocks.

The apparatus and eye movement acquisition were identical to those of EXPERIMENT 1A.

2.2.2 Hypothesis

We hypothesized that performance would be better in blank than in no-blank conditions, and that performance would also improve with the addition of the orientation shift. This way, PSEs would be closer to zero in the blank condition, and more so with greater orientation shifts.

2.2.3 Data Analysis

Performance was analysed as before (PSE, slope & threshold). We tested our hypotheses by performing mixed-models analyses on these measures using the `lme4` R package (Bates, Maechler, Bolker, & Walker, 2015). As predictors, we included blank (`blank` & `no-blank`) and orientation shift (±18°, ±9°, ±3° in EXPERIMENT 1B, and ±45°, ±3° in Experiments 1B & 1C). The random structure comprised a random intercept by subjects. Significant effects were defined as those having an absolute t-value ≥ 2.

We eliminated trials on the basis of the same criteria as previously: Trials in which the eye movement was initiated before the appearance of the target (EXPERIMENT 1B: 9%; EXPERIMENT 1C: 10%) and in which the saccade amplitude or latency was more than two standard deviations away from the average (< 1%).

2.2.4 Results for Experiment 1B

In EXPERIMENT 1B, the saccade characteristics were normal for visually guided saccades: The average saccade amplitude was 6.4^{dva} in both the blank and no-blank conditions (95% confidence intervals, [5.78–6.87] and [5.78–6.89], respectively), and the latencies were 175 [166–185] and 177 [168–187] ms in the blank and no-blank conditions, respectively.

As in EXPERIMENT 1 we ran three mixed-models using blank and orientation shift as predictors, predicting PSE, slopes and thresholds each. A Bayes Factor analysis was subsequently used to validate the models. PSE depended on blank condition (estimate=1.1; stderr=0.1689; t-value=6.58; p-value=0), but not on orientation shift (t-value=-1.69), with no interaction between the two factors (t-value=1.34; see MIXEDMODEL 2). PSEs tended to be negative in the no-blank condition, but not different from zero in the blank condition (FIGURE 23B {p. 97}). The Bayes Factor analysis favoured the model including only blank as a factor (BF > 10^2); there was weak or no evidence in favour of models including orientation shift, both blank and orientation shift, or both factors plus their interaction (BFs < 1.1). Slopes displayed the same behaviour as PSE, but the effect of the blank was smaller (see MIXEDMODEL 2), increasing performance (i.e. higher slopes; estimate=1.13; stderr=0.313; t-value=3.61; p-value=0.004). Orientation shift did not affect the slope (t-value=-0.005) and there was no interaction between the blanking and the orientation shift.

pse	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	-1.1674	0.2467	-4.7319	0.001	***
blank	1.1110	0.1689	6.5783	0.000	***
ori_shift	-0.0073	0.0043	-1.6908	0.098	ns
blank × ori_shift	0.0074	0.0055	1.3429	0.157	ns
slope	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	1.2865	0.2186	5.8852	0.000	***
blank	1.1298	0.3127	3.6132	0.004	**
ori_shift	0.0000	0.0087	-0.0046	0.390	ns
blank × ori_shift	0.0081	0.0113	0.7165	0.297	ns
threshold	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	1.0895	0.1508	7.2263	0.000	***
blank	-0.5328	0.1653	-3.2229	0.007	**
ori_shift	0.0075	0.0034	2.1946	0.044	*
blank × ori_shift	-0.0086	0.0044	-1.9629	0.064	ns

MixedModel 2 Fixed-effects for EXPERIMENT 1B, predicting slope (top) and PSE (centre) and threshold (bottom).

Thresholds as marker of performance was smaller in the blank than in the no-blank condition (estimate=-0.5328; stderr=0.1653; t-value=-3.223; p-value=0.007; MIXEDMODEL 2), and this did depend marginally on orientation shift (estimate=0.0075; stderr=0.003; t-value=2.195; p-value=0.044). There was no interaction between the two factors (t-value=-1.96). The Bayes Factor analysis favoured the model including blank as a factor (BF = 4.83); there was weak or no evidence in favour of models including orientation shift, both blank and orientation shift, or both factors plus their interaction (BFs < 1).

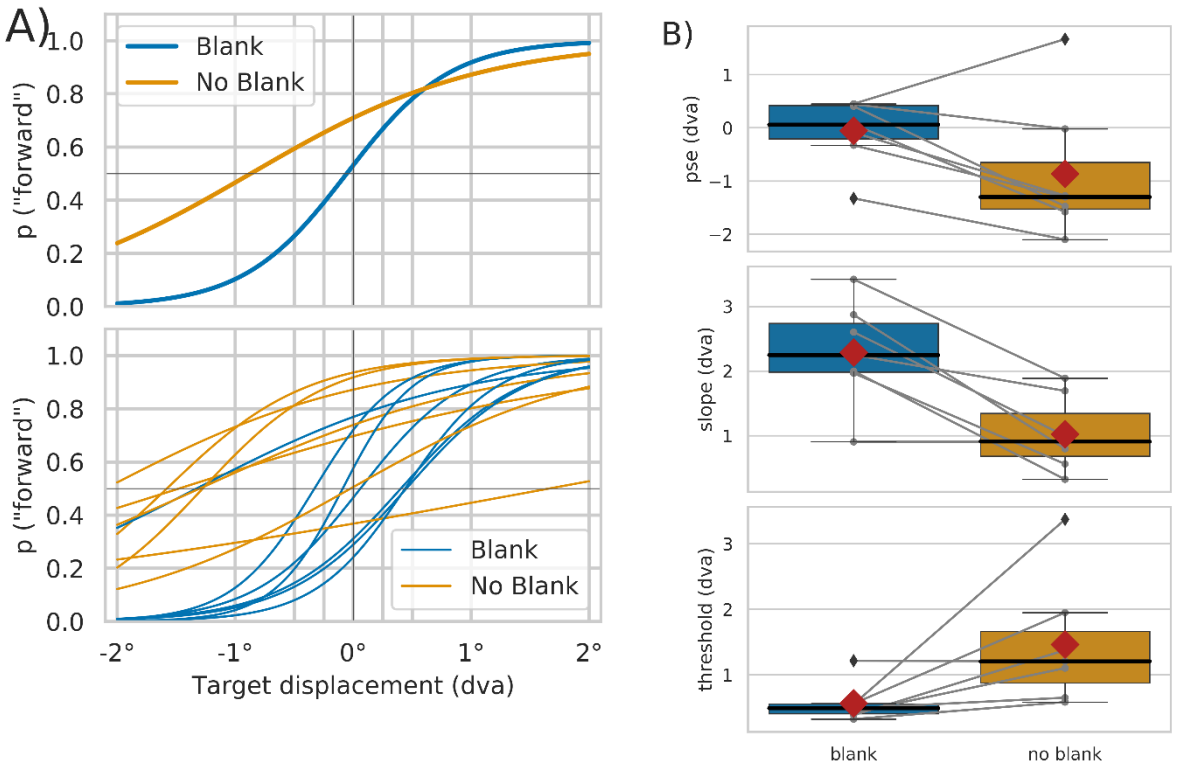

Figure 23 ***A)*** *Top, psychometric function for averaged PSE and slopes across subjects comparing blank (blue) and no-blank (orange) conditions. Ordinate axis shows the probability of a "forward" response given the displacement size (abscissa). Bottom, individual psychometric functions for all subjects comparing the blank (blue) and no-blank (orange) conditions.* ***B)*** *Mean PSE (top) and mean slopes (bottom) comparison for the blank (blue) and no-blank (orange) condition. Individual subjects are represented by dots under each condition and connected between conditions. Black thick lines are the medians; red diamonds represent the means.*

FIGURE 24 {p. 98} shows PSEs, slopes and thresholds for the different orientation shifts for EXPERIMENT 1B. Overall, mean values did not change across different values of orientation shift, regardless of their direction.

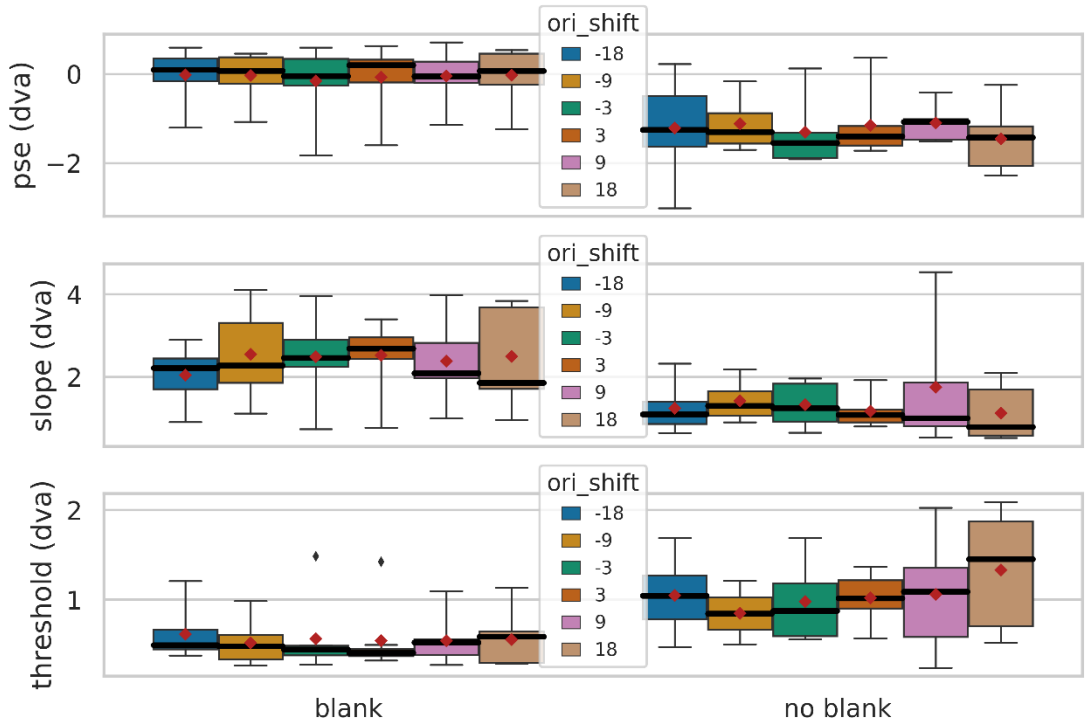

Figure 24 Mean PSE (top), mean slopes (centre) and thresholds (bottom) comparison for the blank (left) and no-blank (right) condition for all different task-irrelevant orientation shifts for EXPERIMENT 1B. Black thick lines are the medians; red diamonds represent the means.

Once we had observed this apparent null effect of orientation shift, we hypothesized that perhaps the shifts were simply too small to be seen or to break object continuity[18], and thus we reran the experiment with a large (45°) and a small (3°) clockwise and counter-clockwise orientation shifts.

2.2.5 *Results for Experiment 1C*

On average, in EXPERIMENT 1C the saccade amplitudes were 6.3 [6.1–6.5] and 6.3 [6.2–6.5] [dva], and the saccade latencies were 162 [154–173] and 162 [153–171] ms for the blank and no-blank conditions, respectively. FIGURE 25 {p. 99} shows comparisons between the blank and no-blank conditions. The difference between conditions follows the results from the two previous experiments: PSEs and slopes increase when the target is displaced with a blank time gap, and thresholds decrease.

[18] Although this was not reported by the subjects in post-experiment debriefings or supported by the performance in EXPERIMENT 2.

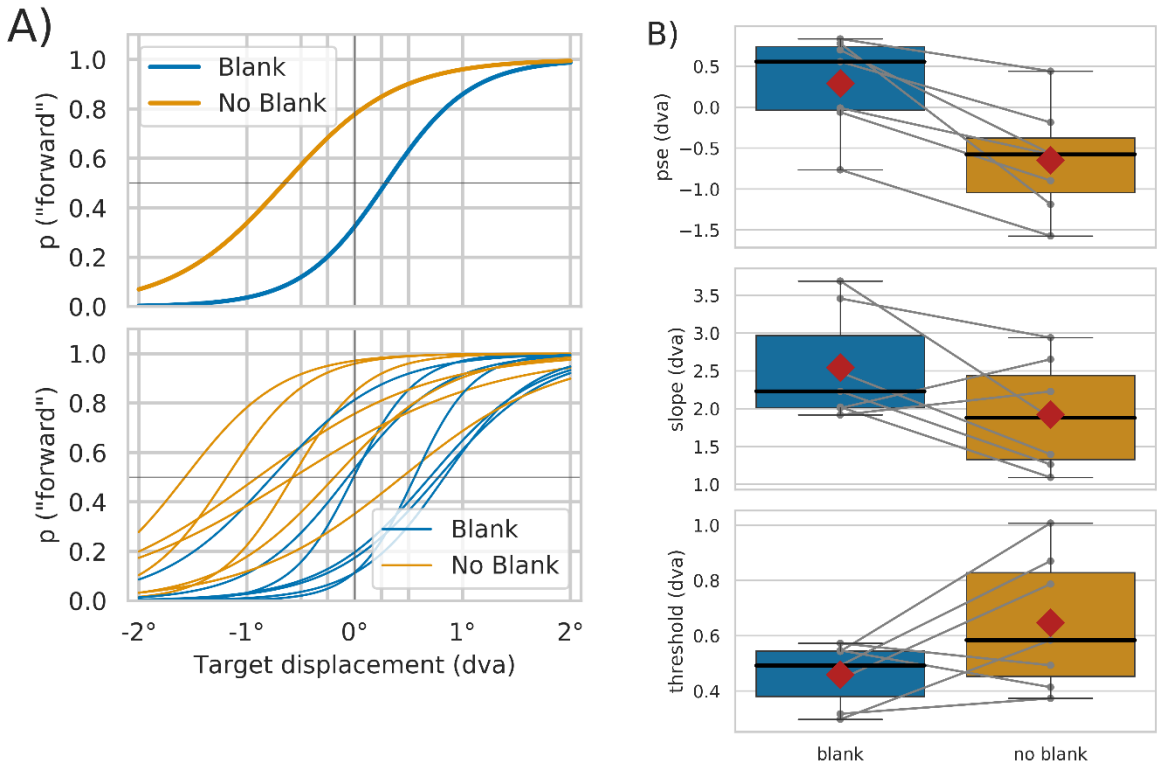

Figure 25 ***A)*** *Top, psychometric function for averaged PSE and slopes across subjects comparing blank (blue) and no-blank (orange) conditions. Ordinate axis shows the probability of a "forward" response given the displacement size (abscissa). Bottom, individual psychometric functions for all subjects comparing the blank (blue) and no-blank (orange) conditions.* ***B)*** *Mean PSE (top) and mean slopes (bottom) comparison for the blank (blue) and no-blank (orange) condition. Individual subjects are represented by dots under each condition and connected between conditions. Black thick lines are the medians; red diamonds represent the means.*

As in EXPERIMENT 1A, PSE increased under the blank condition (estimate=0.9325; stderr=0.1834; t-value=5.0834; p-value=0); with more negative PSEs in the no-blank condition than in the blank condition. However, there was no interaction between the two factors (t-value=-0.71). The Bayes Factor analysis for PSE provided evidence in favour of the model including blank as a factor (BF > 10^2); there was weak or no evidence in favour of models that included only orientation shift (BF < 1), both blank and orientation shift (BF = 1.8), or both factors plus their interaction (BF < 1).

Slopes did not depend on blank condition (estimate=0.5679; stderr=0.4200; t-value=1.3524; p-value=0.155). There was no effect of orientation shift (t-value=-0.56) and no interaction between the two factors (t-value=0.57). The Bayes Factor analysis provided strong evidence in favour of the model including only blank as a factor (BF = 4883); weaker evidence in favour of the model including both blank and orientation shift (BF = 1,280) or both factors plus their interaction (BF = 407); and evidence for the null hypothesis that orientation alone did not influence the slopes (BF < 1).

Thresholds did not significantly change in the blank condition relative to the no-blank condition (t-value=-1.83), and did not depend on orientation shift (t-value=0.88). The interaction between the two factors was also insignificant (t-value=-0.98).

pse	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	-0.6463	0.2325	-2.7801	0.016	*
blank	0.9325	0.1834	5.0834	0.000	***
ori_shift	0.0021	0.0018	1.1623	0.195	ns
blank × ori_shift	-0.0018	0.0025	-0.7139	0.297	ns
slope	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	2.1955	0.2785	7.8836	0.000	***
blank	0.5679	0.4200	1.3524	0.155	ns
ori_shift	-0.0029	0.0051	-0.5574	0.330	ns
blank × ori_shift	0.0040	0.0072	0.5653	0.328	ns
threshold	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.5881	0.0737	7.9799	0.000	***
blank	-0.1365	0.0745	-1.8314	0.079	ns
ori_shift	0.0006	0.0007	0.8838	0.258	ns
blank × ori_shift	-0.0009	0.0009	-0.9838	0.235	ns

MixedModel 3 Fixed-effects for EXPERIMENT 1C.

FIGURE 26 {p. 101} shows the blank versus the no-blank condition within all four orientation shifts, and MIXEDMODEL 3 the fixed-effects for the mixed-models predicting PSE, slopes and thresholds.

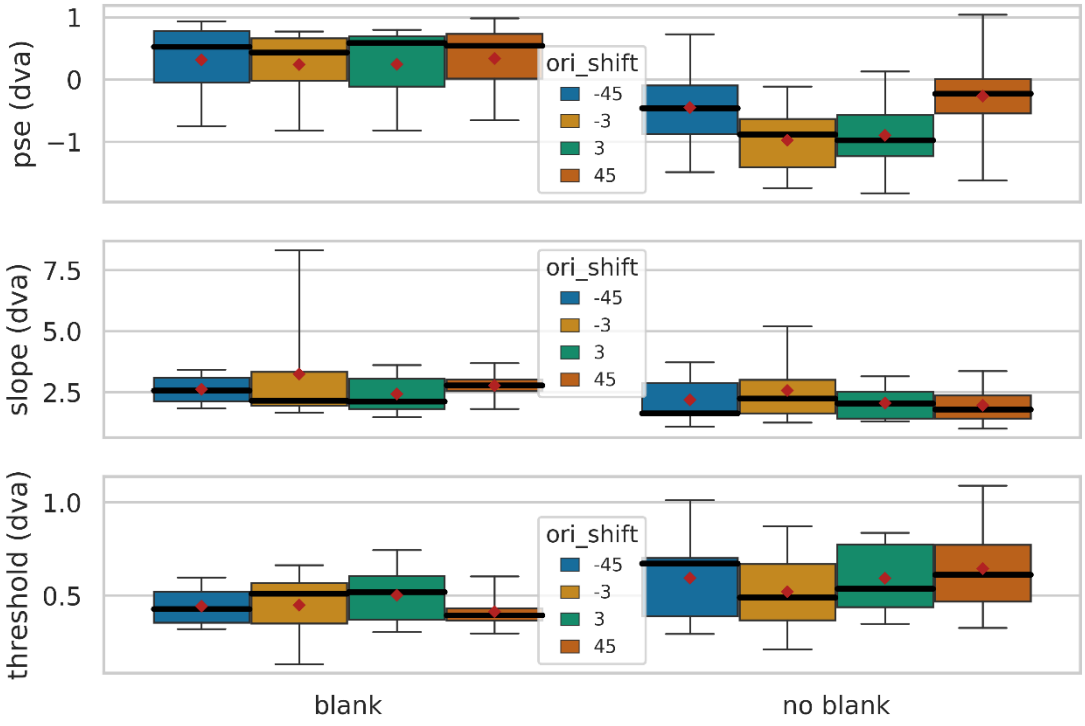

Figure 26 Mean PSE (top), mean slopes (middle) and thresholds (bottom) comparison for the blank (left) and no-blank (right) conditions for all different task-irrelevant orientation shifts for Experiment 1C. Black thick lines are the medians; red diamonds represent the means.

To compare conditions with and without an orientation shift, we averaged the thresholds across orientation shifts in EXPERIMENT 1B and ran an independent Bayesian t-test to compare performance between shift and no-shift conditions. The no-shift condition was taken from EXPERIMENT 1A. The relevant condition was the no-blank condition, since this was the condition in which we had expected an improvement. The Bayes Factor (BF = 0.36) favoured the null hypothesis that there was no difference in displacement discrimination performance between shift and no-shift conditions, over the alternative hypothesis that there was.

In sum, EXPERIMENT 1B and 1C confirmed the classic blanking effect but provided no evidence that orientation shifts improved displacement discrimination across saccades. To address our hypotheses about the role of visual features in visual stability, we wanted to probe not only displacement discrimination, but also orientation discrimination. In the second series of experiments, participants no longer reported the direction of displacement, but rather the direction of orientation shift (clockwise or counter-clockwise).

2.3 Experiment 2A

Orientation shift (report orientation shift)

The aim of this experiment was to determine to what extent the 250-ms blank influenced participants' performance in reporting the direction of changes in orientation *during* (no-blank condition) and *after* (blank condition) a saccade. Given that displacements which occur upon saccade landing are noticed by the observer (Deubel et al., 1996), we wanted to find out if changes to a visual feature (during and after a saccade) would make subjects to better —or also— detect feature-information discrepancies.

2.3.1 Methods

Subjects

Eight healthy human adults (four women; 32 ± 13 years old) participated. All reported normal or corrected-to-normal vision and no oculomotor or neurological impairments, received payment of 10€/h for their time, and provided informed consent.

Stimuli and Procedure

In EXPERIMENT 2A, the orientation of the Gabor patch was randomly initialized at 45°–135° when it appeared at position F. The orientation remained stable as the stimulus stepped from F to T. When the saccade was at 20% of the F–T distance, the patch underwent an orientation shift randomly selected from 18°, 9°, 6°, or 3° clockwise or counter-clockwise. The new patch appeared either immediately (no-blank condition) or after a 250 ms blank following saccade termination (blank condition). The patch never changed positions during the saccade. Participants were instructed to report whether the change of orientation had been clockwise or counter-clockwise[19], using the keyboard. Each subject performed 512 trials, divided into four blocks of 128 trials each.

The apparatus and eye movement acquisition were identical to those aspects of the previous experiments.

[19] Subjects were instructed to report the orientation shift of the *upper* part of the Gabor to avoid any confusions.

2.3.2 Hypothesis

We hypothesized that orientation shift discrimination would be better in blank than in no-blank conditions. We also hypothesized that the PSEs would be closer to zero under the blank condition, and biased under the no-blank condition.

2.3.3 Data Analysis

All data analyses were performed as in EXPERIMENT 1, except that the dependent variable was the proportion of "clockwise" responses. The elimination criteria were also the same as before: trials in which the eye movement was initiated before the appearance of the target (7%) and in which saccade amplitude or latency was more than two standard deviations away from the average (< 1%).

2.3.4 Results

On average, the saccade amplitude was 6.9 [dva] [6.7–7.1] and the latency was 162 ms [153–171] for both the blank and no-blank conditions.

FIGURE 27A {p. 104} shows the individual psychometric functions. Discriminating the orientation shift across the saccade did not depend on whether a blank occurred, with average thresholds of 8.0 [5.4–11.0] and 7.5 [5.1–11.0] degrees for the blank and no-blank conditions (FIGURE 27B). Mixed-models ran for PSEs, slopes and thresholds confirm this, as shown in MIXEDMODEL 4. The blank condition did not affect neither of the three predicted variables, with absolute t-values ≤ 2 (pse: -1.2; slope: 0.9; threshold: -0.2).

pse	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.6193	0.5199	1.1914	0.176	ns
blank	-0.9295	0.7610	-1.2214	0.170	ns
slope	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.1743	0.0345	5.0554	0.003	**
blank	0.0146	0.0165	0.8883	0.239	ns
threshold	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	8.0294	1.3940	5.7599	0.001	**
blank	-0.1313	0.6665	-0.1970	0.366	ns

MixedModel 4 Fixed-effects for EXPERIMENT 2A.

The Bayes Factor analysis comparing the hypothesis that performance was better in the blank than in the no-blank conditions to the null hypothesis of equal thresholds between the blank and no-blank conditions revealed evidence in favour of the null hypothesis (BF = 0.48).

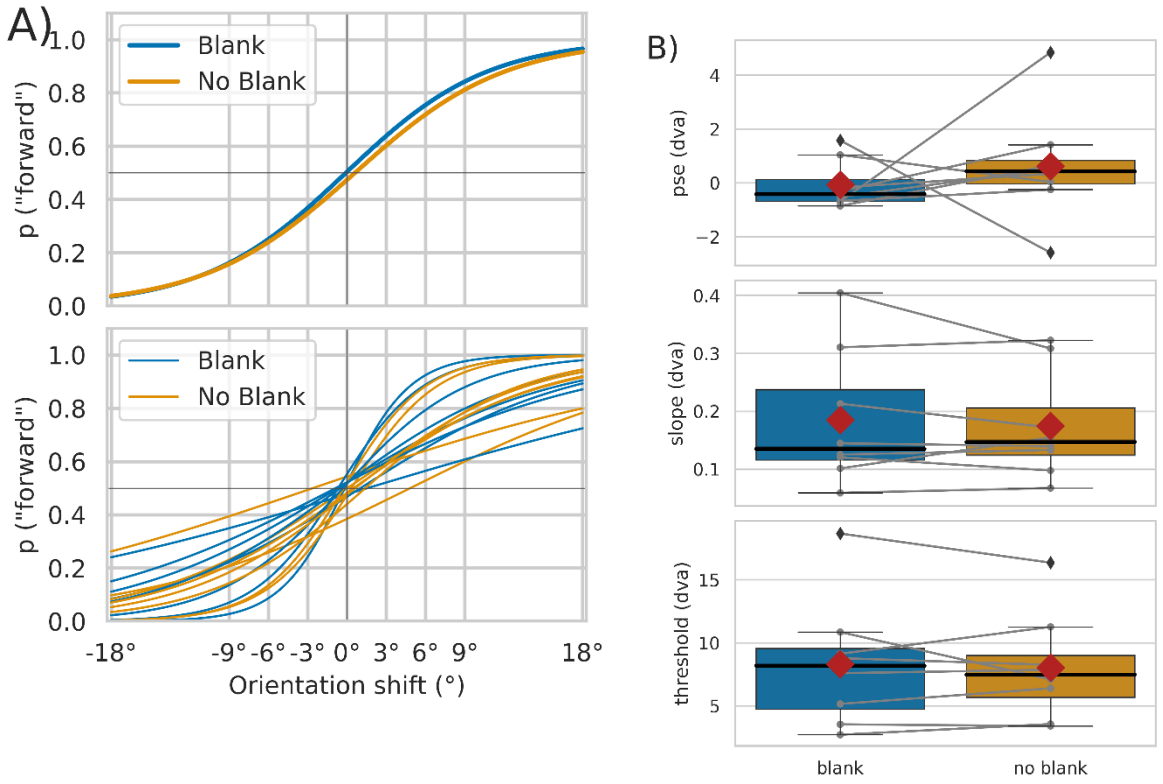

Figure 27 ***A)*** *Top, psychometric function for averaged PSE and slopes across subjects comparing blank (blue) and no-blank (orange) conditions. Ordinate axis shows the probability of a "forward" response given the displacement size (abscissa). Bottom, individual psychometric functions for all subjects comparing the blank (blue) and no-blank (orange) conditions.* ***B)*** *Mean PSE (top) and mean slopes (bottom) comparison for the blank (blue) and no-blank (orange) condition. Individual subjects are represented by dots under each condition and connected between conditions. Black thick lines are the medians; red diamonds represent the means.*

PSE did not depend on blank, with PSEs of – 0.20 [– 0.88 to 0.48] in the blank condition and 0.31 [– 0.10 to 0.74] in the no-blank condition. Again, the Bayes Factor analysis revealed evidence in favour of the null hypothesis of equal PSEs over the alternative hypothesis of different PSEs in the blank than in the no-blank condition (BF = 0.49).

In sum, in EXPERIMENT 2A we measured transsaccadic orientation discrimination and showed that the blank did not influence performance. However, this did not preclude an influence of an irrelevant displacement on orientation discrimination performance, which was necessary to examine in order

to test our hypothesis about the role of visual features and remapping in visual stability. In the next experiment, we thus added an irrelevant intra-saccadic displacement.

2.4 Experiment 2B

Displacement with orientation shift (report orientation shift)

2.4.1 Methods

Subjects

Seven healthy human adults (four women; 31 ± 14 years old) participated. All reported normal or corrected-to-normal vision and no oculomotor or neurological impairments, received payment of 10€/h for their time, and provided informed consent.

Stimuli and Procedure

This experiment was identical to EXPERIMENT 1B, except that participants reported the orientation shift and not the displacement. The apparatus and eye movement acquisition were identical to those aspects of the previous experiments.

2.4.2 Hypothesis

We hypothesized that performance would be better in conditions with an irrelevant target displacement.

2.4.3 Data Analysis

We analysed performance as in EXPERIMENT 2A. We tested our hypothesis by implementing a mixed-model using the `lme4` R package (Bates et al., 2015). As predictor variables, we included blank (blank, no blank) and target displacement ($\pm 2^{dva}$, $\pm 1^{dva}$, $\pm 0.25^{dva}$). The random structure comprised a random-intercept by subject. We eliminated trials as before: trials in which the eye movement was initiated before the appearance of the target (7%) and in which the saccade amplitude or latency was more than two standard deviations away from the average (< 1%).

2.4.4 Results

On average, the saccade amplitudes were 6.7 [6.6–6.9] and 6.7 [6.5–6.9] dva, and the latencies were 162 [152–172] and 162 [153–172] ms in the blank and no-blank conditions, respectively.

FIGURE 28 shows the fitted values and averages for blank versus no-blank conditions, and FIGURE 29 {p. 108} all six target displacements. The fixed-effects for the mixed-model are shown in MIXEDMODEL 5 {p. 107}.

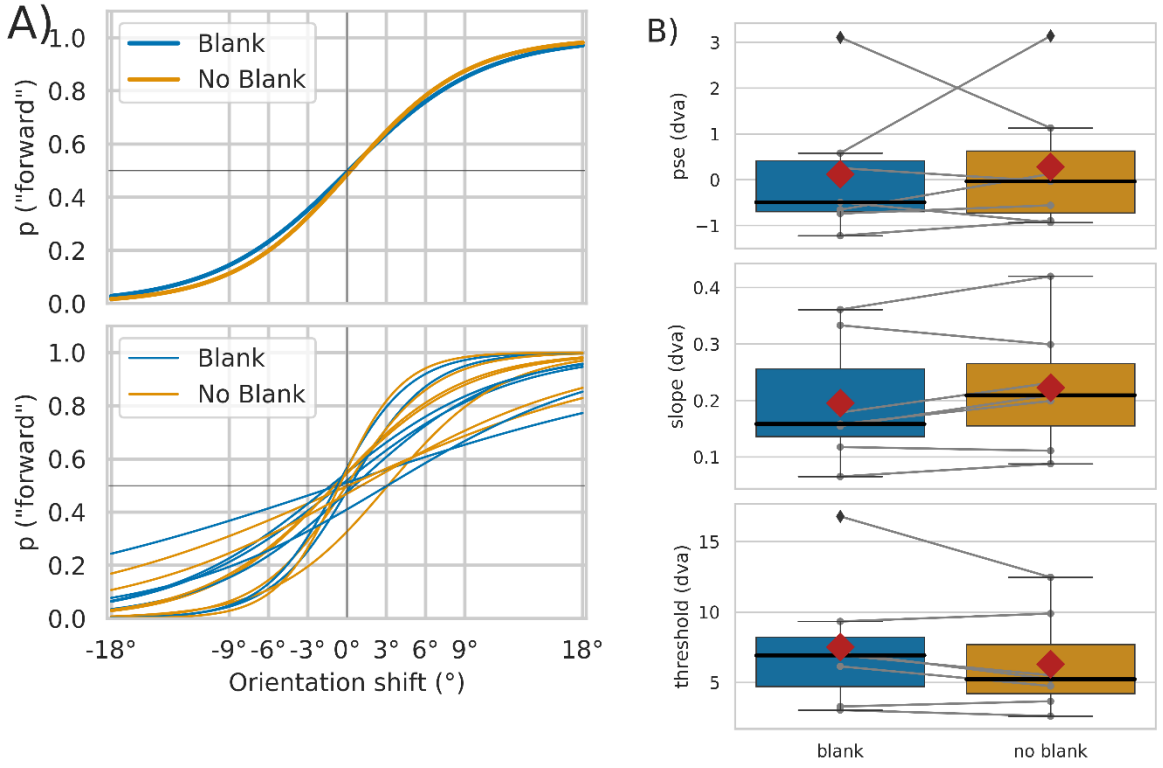

Figure 28 ***A)*** *Top, psychometric function for averaged PSE and slopes across subjects comparing blank (blue) and no-blank (orange) conditions. Ordinate axis shows the probability of a "forward" response given the displacement size (abscissa). Bottom, individual psychometric functions for all subjects comparing the blank (blue) and no-blank (orange) conditions.* ***B)*** *Mean PSE (top) and mean slopes (bottom) comparison for the blank (blue) and no-blank (orange) condition. Individual subjects are represented by dots under each condition and connected between conditions. Black thick lines are the medians; red diamonds represent the means.*

PSE did not depend on blank (t-value=-0.27) nor in step (t-value=-1.31). However, we found a relatively large interaction between blank and step (estimate=-1.1745; stderr=0.3562; t-value=-3.2973; p-value=0.006). Regardless of the interaction found, the Bayes Factor analysis revealed

evidence in favour of the null hypothesis (absence of a difference in PSE between blank and no blank) over models that included blank, step, and their interaction (all BFs < 1).

Performance as measured by slopes depended marginally on blank (estimate=-0.0368; stderr=0.016; t-value=-2.26; p-value=0.040), and not on displacement (t-value<1). There was no interaction between the two factors (t-value<1). The Bayes Factor analysis revealed evidence in favour of the null hypothesis (absence of a difference in thresholds) over models that included blank, displacement, and their interaction (all BFs < 1).

pse	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.2080	0.4948	0.4202	0.354	ns
blank	-0.1304	0.4795	-0.2719	0.375	ns
step	-0.4533	0.3458	-1.3107	0.163	ns
blank × step	-1.1745	0.3562	-3.2973	0.006	**
slope	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	0.2450	0.0459	5.3400	0.000	***
blank	-0.0368	0.0163	-2.2564	0.040	*
step	0.0062	0.0080	0.7708	0.284	ns
blank × step	0.0049	0.0113	0.4310	0.353	ns
threshold	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	6.5058	1.5469	4.2057	0.001	**
blank	0.9750	0.6355	1.5342	0.122	ns
step	-0.5019	0.3559	-1.4102	0.144	ns
blank × step	0.5647	0.5016	1.1258	0.203	ns

MixedModel 5 Fixed-effects for PSE, slope and threshold as predicted variables in EXPERIMENT 2B.

No dependence was found for thresholds (all t-values<1.6).

We compared performance between no displacement (EXPERIMENT 2A) and displacement (EXPERIMENT 2B) conditions in the no-blank version of the task. We averaged the values across displacements in EXPERIMENT 2B and ran an independent Bayesian t-test to compare performance between displacement and no-displacement conditions. The Bayes factor (BF = 0.41) favoured the null hypothesis of no difference in orientation discrimination performance between displacement and no displacement, over the alternative hypothesis of a difference.

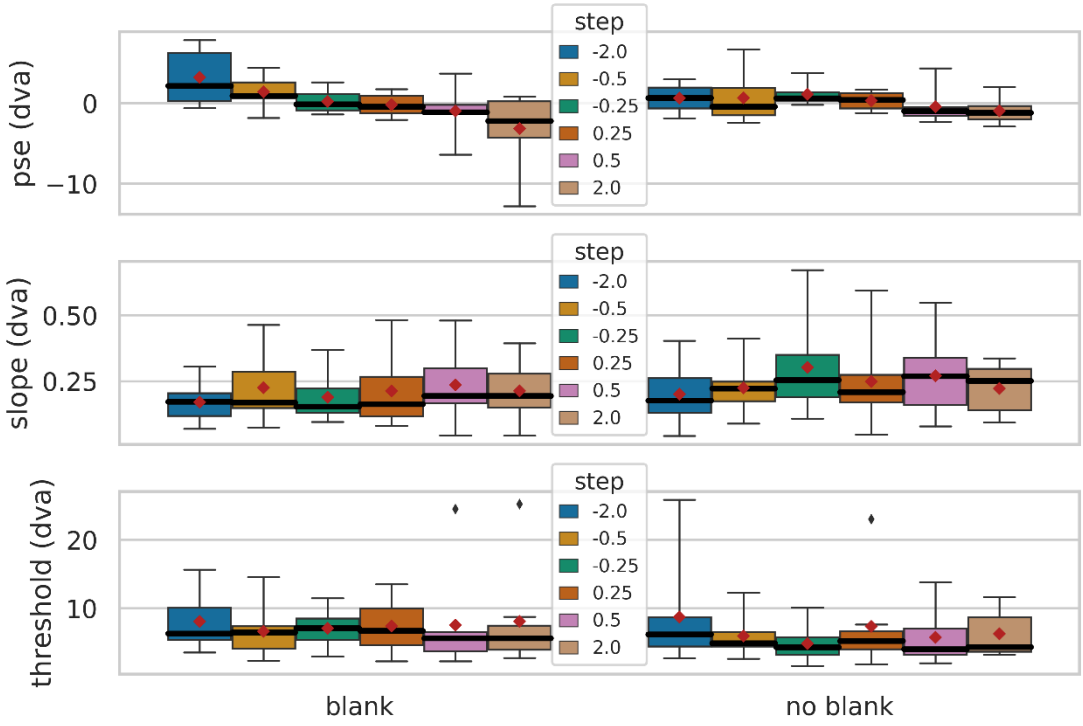

Figure 29 Mean PSE (top) and mean slopes (bottom) comparison for the blank (blue) and no-blank (orange) condition for all different task-irrelevant displacement sizes for Experiment 2B. Black thick lines are the medians; red diamonds represent the means.

2.5 A link between Feature and Spatial Information

For all three experiments in which the stimulus displacement and the orientation shift were variable (EXPERIMENTS 1B, 1C & 2B), we tested whether the direction of the orientation shift (clockwise vs counter-clockwise) and the direction of the displacement (left vs right or backwards vs forwards) influenced subjects' performance when these two variables were congruent. We did this by labelling trials as `congruent` whenever the orientation shift and the displacement were on the same direction, that is, clockwise & right or counter-clockwise & left, regardless of the side towards which the saccade had been made. Each trial, thus, was marked as congruent or incongruent according to this rule, and then a mixed-model was ran using `blank`, `step`, `ori_shift`, `saccade` & `congruency` as predictors for the correctness of the response.

The predictor `step` takes values based on different displacement sizes, and `ori_shift` based on different orientation shifts. Each of these values reflects a different magnitude of change. `congruency`, on the other hand, is not concerned by magnitudes, but by their *direction* (the signs

of the magnitudes). Therefore, it echoes if both magnitudes were changed in the same direction or not. Previous mixed-models measured the variance due to changes only in the magnitudes of these two predictors. We checked that `congruency` was worth considering a predictor by checking the correlation between its composing variables for each experiment. We found that neither correlated (1B: $\rho = -0.00052, p = 0.96$; 1C: $\rho = 0.0028, p = 0.81$; 2B: $\rho = 0.0031, p = 0.76$). Moreover, we verified the normalised mutual information (NMI) between the two variables and found that there was no mutual information (1B: $NMI = 0.0001$; 1C: $NMI = 0.0002$; 2B: $NMI = 0.0002$). Both measures guaranteed that `step` and `ori_shift` were independent, and that `congruency` in fact reflects additional information of each trial.

By analysing subjects' performance in this way, we hypothesised that we would be able to determine if the change in orientation —a visual feature of the stimulus— predisposed its perceived location —a spatial feature of the stimulus—, or vice versa. As we saw before (I4.4 VISUAL OBJECT RECOGNITION {p. 30}), both of these aspects are treated parallelly via different connections that exit the visual cortex: the ventral and the dorsal pathways. In principle, information carried into these two visual-processing streams must not necessarily be different, although once in their respective treating brain areas, it is decoded in different ways to achieve different tasks. Can we determine if there is information common to both pathways that affects perception of features *and* space? In other words, can orientation shifts be confounded by the dorsal system as spatial information? Or can horizontal displacements be "mistakenly" taken into account by the ventral system as visual features of the stimulus? In terms of the above-mentioned experiments the question we asked is, does a clockwise orientation shift give subjects the impression that a horizontal step was towards the right even though it might have been towards the left or just noticeable? Likewise, when reporting orientation shifts.

To respond to the last question, we thus created a mixed-model for each experiment with the added `congruent` predictor. The resulting fixed-effects are shown side-by-side in the next table (MIXEDMODEL 6 {p. 110}). Predictor variables names have been renamed as stated in the table's legend, to compact the its width.

	Experiment 1B					Experiment 1C					Experiment 2B				
resp_ok	**est.**	**err.**	**t-val**	**p-val**	**sign**	**est.**	**err.**	**t-val**	**p-val**	**sign**	**est.**	**err.**	**t-val**	**p-val**	**sign**
intrcp	0.551	0.012	45.15	0	***	0.659	0.013	50.52	0	***	0.756	0.034	22.05	0	***
B	0.180	0.025	7.26	0	***	0.072	0.022	3.26	0.0035	**	-0.047	0.017	-2.73	0.0126	*
S	0.010	0.029	0.35	0.3712	ns	0.052	0.026	1.98	0.0592	ns	0.011	0.010	1.06	0.2234	ns
O	-0.011	0.001	-10.06	0	***	0.000	0.000	-0.65	0.3189	ns	0.001	0.001	0.62	0.3249	ns
C	-0.009	0.019	-0.46	0.3545	ns	-0.039	0.023	-1.74	0.0892	ns	0.022	0.021	1.07	0.2228	ns
B×S	0.009	0.014	0.65	0.3198	ns	-0.031	0.015	-2.10	0.0473	*	-0.003	0.013	-0.20	0.3878	ns
B×O	0.012	0.001	9.10	0	***	0.003	0.001	6.27	0	***	-0.001	0.001	-0.69	0.3095	ns
S×O	**-0.007**	**0.001**	**-7.82**	**0**	***	**-0.003**	**0.0004**	**-9.19**	**0**	***	**-0.004**	**0.001**	**-4.69**	**0.0001**	***
B×C	-0.005	0.022	-0.24	0.3846	ns	0.040	0.025	1.62	0.1077	ns	0.033	0.021	1.55	0.1194	ns
S×C	0.005	0.015	0.32	0.3751	ns	0.011	0.016	0.68	0.3132	ns	0.002	0.014	0.12	0.3931	ns
O×C	0.022	0.002	14.69	0	***	0.001	0.001	1.04	0.2295	ns	-0.003	0.002	-1.67	0.1	ns
B×S×O	0.000	0.001	-0.12	0.3929	ns	0.001	0.001	1.60	0.112	ns	-0.001	0.001	-0.54	0.3406	ns
B×S×C	-0.010	0.019	-0.50	0.3482	ns	-0.018	0.021	-0.85	0.2731	ns	-0.003	0.018	-0.16	0.3906	ns
B×O×C	-0.025	0.002	-13.02	0	***	-0.007	0.001	-8.57	0	***	0.002	0.002	0.73	0.3021	ns
S×O×C	**0.016**	**0.001**	**12.13**	**0**	***	**0.007**	**0.001**	**13.32**	**0**	***	**0.009**	**0.001**	**6.75**	**0**	***
B×S×O×C	0.000	0.002	-0.10	0.3936	ns	-0.002	0.001	-2.64	0.0152	*	0.002	0.002	0.90	0.261	ns

MixedModel 6 Fixed-effects for the three mixed-models for experiments 1b, 1c & 2b. The intercept is the no-blank condition. Predictor names have been renamed to fit the three tables: B=`blank`, S=`step`, O=`ori_shift`, C=`congruency` with their respective interactions. Significant values have been highlighted in grey, and significant values common to all three experiments have been highlighted in different colours.

Changes in orientation shift affected minimally and negatively subjects' performance in EXPERIMENT 1B, showing an estimated decrease of 0.11 and a significant t-value of -10.06. This is in line with the estimated threshold increase previously found (MIXEDMODEL 2 {p. 96}), reflecting worse performance. In the remaining experiments, orientation shift either does not affect at all subjects' performance (EXPERIMENT 1C) or increases insignificantly their performance (EXPERIMENT 2B).

Congruency on its own was not a significant main effect for neither of the experiments. The `congruency` predictor decreases subjects' ability to discriminate displacements (Exp 1b: -0.009; Exp1c: -0.039) and improves their sensitivity to changes in orientation (Exp 2b: 0.022). These changes, nevertheless, were not significant probably because of their high error (Exp 1b: 0.019; Exp 1c: 0.023; Exp 2b: 0.021).

The interactions between predictor variables revealed information regarding subjects' performance. In EXPERIMENT 1C, in which only "extreme" orientation shifts were used (±45° and ±3°) and in which subjects reported direction of displacement, we found an interaction between `blank` and `step` (estimate=-0.031;stderr=0.015; t-value=-2.10; p-value=0.0473), which was not the case in its corresponding mixed-model (MIXEDMODEL 3 {p. 100}). That is, when tested with an extra 250-ms time-gap (blank condition) in combination with different displacement sizes, subjects performed worse.

Note that errors for this interaction are all of the same order of magnitude, which unveils the importance of the step-size when intervening orientation shifts are broader and less numerous.

Only in EXPERIMENT 1B did `ori_shift` interact with `congruency` (estimate=0.022; stderr=0.002; t-value=14.69; p-value=0), indicating that subjects were better in congruent trials as the orientation shift varied. This confirms the hypothesis that equal direction of change influenced the perception of direction. Note that in this experiment, `step` and `ori_shift` magnitudes vary proportionally, as opposed to EXPERIMENT 1C at which the interaction was not significant and magnitudes do not vary proportionally.

Next, for experiments 1B and 1C the mixed-model discovered a strong interaction between `blank`, `ori_shift` and `congruency` (1B: estimate=-0.025; stderr=0.002; t-value=-13.02; p-value=0. 1C: estimate=-0.007; stderr=0.001; t-value=-8.57; p-value=0), showing that different combinations of these two factors tended to reduce subjects' performance. Comparing the t-values of this interaction shows that a large set of orientation shifts with monotonically increasing values (Experiment 1B) renders subjects less sensitive to displacement discrimination, as opposed to a small and 'drastic' set of orientation shifts (Experiment 1C). Interestingly, when participants' task is to discriminate direction of orientation shift (Experiment 2B), this interaction no longer enhances performance.

The `blank` predictor interacted significantly with `ori_shift` in EXPERIMENTS 1B and 1C (1B: estimate=0.012; stderr=0.001; t-value=9.10; p-value=0. 1C: estimate=0.003; stderr=0.001; t-value=6.27; p-value=0), and, as expected, not in EXPERIMENT 2B. This reveals the aid that the blank time-gap gives in detecting displacements, but not in detecting orientation changes. This could mean that the vFM treats preferentially spatial information over feature-information, when the task at hand involves spatial attention rather than feature attention.

As expected from experiments 1B and 1C, `blank` significantly increased subjects' performance as compared to when there was no blank (the intercept). EXPERIMENT 2B, on the other hand, displays the opposite behaviour, showing that subjects perform *worse* when there is a blank than when there is not. This significance matches the decrease in performance given by the slope in MIXEDMODEL 5 for EXPERIMENT 2B {p. 107}. In both mixed-models, the blanking made subjects less performant.

Apart from the `blank` main effect which is common to all three experiments, there are two significant interactions present in all three experiments. These are: `step` × `ori_shift` and `step` × `ori_shift` × `congruency`. Both interactions reveal relevant information.

For all three experiments, `step` showed a significant interaction with `ori_shift` (1B: estimate =-0.007; stderr=0.001; t-value=-7.82; p-value=0. 1C: estimate=-0.003; stderr=0.000; t-value=-9.19; p-value=0. 2B: estimate =-0.004; stderr=0.001; t-value=-4.69; p-value=0.0001). Such interactions showed a decrease in performance regardless of the task, with a more important decrease in performance when the orientation shifts are gradually changed as in EXPERIMENT 1B; nevertheless, for the three experiments the decrease remains stable at the same degree of magnitude and with small errors.

Most importantly, for all three experiments we also found a strong interaction between `step`, `ori_shift` and `congruency` (1B: estimate=0.016; stderr=0.001; t-value=12.13; p-value=0. 1C: estimate =0.007; stderr=0.001; t-value=13.32; p-value=0. 2B: estimate =0.009; stderr=0.001; t-value=6.75; p-value=0). The increase of the estimates suggests that with different combinations of these three factors, subjects tended slightly —but significantly— to better detect the probed visual change. This can be interpreted as we stated before: when the changes are aligned in the same direction, participants' performance is enhanced.

The `step×ori_shift` interaction illustrates a *decline* in performance that depends on the *change of magnitudes* of the horizontal step and the orientation shift. Additionally, when such changes of magnitude take the *same direction*, as in the `step×ori_shift×congruency`, subjects *increase* their performance. These interactions suggest that, regardless of the task that subjects are performing (reporting orientation shifts or reporting displacements), both types of information interact in the visual system, which in turn could suggest that both ventral and dorsal streams are used in both tasks, as the two types of information —spatial and features— co-exist in the task.

2.6 Discussion

In a modified version of the in-flight displacement task, we replicated classic SSD findings. Performance in discriminating direction of an intra-saccadic target displacement was higher when a ~250 ms blank separated the end of the saccade and the appearance of the displaced target. This was not the case when the displaced target was changed immediately after the saccade. We also found that when reporting displacements, the no-blank condition displayed negative PSEs; positive PSEs were found when reporting orientation shifts. A stable target was perceived to have stepped forward, and stability was perceived for a target stepping backward. In the blank condition, this bias disappeared.

These findings reflect the importance of the spatiotemporal correspondence between the predicted post-saccadic location of the target and its actual location (Deubel, 2004). Points of subjective equality are a measure of the gain of the CD signal. Imagine a saccade target at 8^{dva}. In the case of a perfect 8^{dva} saccade with a CD with a gain greater than 1, the predicted location of the post-saccadic target would be negative; that is, according to the internal copy of the saccade, the eyes would land beyond the target. However, because the saccade actually landed exactly on the target, there was a mismatch between the predicted and actual locations of the post-saccadic target, and subjects reported that the target had stepped backwards. Stability would be perceived if the target stepped backwards to match the overestimated gain. Thus, in our data, the negative PSEs suggest a CD gain greater than 1. In the blank condition, PSEs shifted back to zero, suggesting that the gain of the CD signal no longer played a role in determining the response. Rather, performance was determined by accurate location signals.

Contrary to our hypotheses, spatiotopic performance (displacement discrimination thresholds and slopes) did not improve with the addition of an irrelevant orientation shift (although the effect of the blank was still present with these shifts). We checked that this was not due to the fact that our orientation shifts were too small to be perceived, in three ways: We reran the experiment with very large shifts (45°) and replicated the null result (EXPERIMENT 1C); we asked participants in a post-experiment debriefing whether they had perceived the shifts (they did, and easily); and we measured perception in a second series of experiments (our large shift levels were well above threshold). Our result was in contrast to several previous studies showing that irrelevant feature changes across saccades can improve displacement discrimination. Demeyer et al. (2010) switched shape (e.g., square to triangle) and found that spatiotopic performance was better than in the no-switch control. Tas et al. (2012) switched contrast polarity (white to black) and object identity (e.g., button to clock) and also reported that performance was improved in conditions with an irrelevant feature change. In these studies, however, blanking and feature changes were manipulated in separate trials, not simultaneously as in the present experiments. By manipulating them separately, the hypothesised vFM received either spatial- or feature-information, but not both. Although it is an empirical question whether such a small methodological difference could cause the large difference in the pattern of results, the only hypothesis we imagine as to why this would gate the influence of feature changes on spatiotopic performance is by means of the vFM's input. Zimmermann, Born, Fink, and Cavanagh (2014) switched a single bar from left-oriented to right-oriented, and this shape/orientation shift also improved performance. However, orientation shifts of a single bar also modify location signals. We had therefore expected orientation shifts to impede SSD; our negative results suggest that not all visual features contribute equally to object continuity across saccades.

In the second series of experiments, we asked participants to report whether the feature underwent a change across the saccade, by asking them to report the orientation shift and ignore any displacements. Subjects were able to perform the task, but none of our experimental manipulations influenced their performance: neither the blank nor the displacement. Again, this contrasts with previous reports suggesting that a blank improved transsaccadic detection of form (Deubel et al., 2002), colour (Tas et al., 2014), or spatial frequency (Weiß, Schneider, & Herwig, 2015).

In a related set of studies, Poth and Schneider showed that changing the surface features of the target object across the saccade deteriorated post-saccadic object recognition (measured by the identification of a letter embedded within the post-saccadic object), relative to a condition in which the surface features of the object remained the same across the saccade (Poth and Schneider, 2016). Poth, Herwig, and Schneider (2015) showed that a blank improved displacement discrimination (the classic blanking effect), but also found that it deteriorated letter identification, similar to our results. These results were interpreted as suggesting that when pre- and post-saccadic objects are considered two distinct objects, they may compete for attention, leaving fewer resources available for the letter identification task (Schneider, 2013). In our study, we did not ask participants to report the orientation of the post-saccadic object independently from the pre-saccadic orientation, instead we asked them to report the direction of *change* between the two, so we cannot directly address this idea of competing representations between the pre- and post-saccadic objects. However, competition between representations may degrade performance in the comparison between pre- and post-saccadic objects, leading to higher thresholds (worse performance) in the blank condition. We did not replicate this finding with our orientation task, since the blank had no effect on orientation discrimination of the post-saccadic object. There are several possible sources for this difference. First, Poth et al. (2015) did not actually test post-saccadic object perception, but perception of a letter within the object. It is not mutually exclusive to have deteriorated letter identification and normal object identification. Second, letter identification taps into processes that are probably quite different from those involved in orientation discrimination.

Our results do not provide unequivocal support for either account of visual stability: remapping or VSTM. The VSTM account of visual stability predicts that changing the orientation should unmask target displacements, but we found no effect of orientation shifts on displacement discrimination. The remapping account of visual stability predicts that displacing the target should unmask feature changes, but we found no effect of displacements on orientation discrimination. It also predicts an improvement in performance with a blank, which again we did not find or found it to decrease performance instead. We do not believe that this is due to the displacements being too small to be perceived, because we asked participants in post-experiment debriefings whether they had perceived

any displacements (which they did), and we measured that perception in the first series of experiments (our large displacements were well above threshold). Perhaps our orientation shifts were too easy to discriminate, leading to a ceiling effect. We think this is unlikely: First, because the discrimination thresholds measured in the no-displacement condition (Experiment 2A) were higher than some of the tested shifts. Also, even in the subjects with high thresholds, in which there may have been room for improvement, the blank barely changed performance.

We are thus left with evidence in the literature and in the present results, of the importance of spatiotemporal information about the predicted post-saccadic location of the target, arguing in favour of remapping as a crucial mechanism for visual stability when spatial information is critical. We also have evidence in the literature, but not in the present results, of the importance of retaining feature information across saccades for visual stability. We conclude that both mechanisms must play a role for visual stability— that is, spatiotemporal knowledge both about the saccade in the form of a CD signal, and the visual attributes of the target. These two aspects are treated by the remapping account and by the VSTM account. The former manages spatial information, while the latter features information. Our results suggest that not all visual features are equally taken into account to establish object continuity and visual stability. In particular, orientation does not play the same role as shape or polarity. A fascinating question for future research will be to map out which visual features do serve visual stability and which do not—and why.

3 DISCUSSION

In the studies presented in the experimental part, we used psychophysics to test the way in which Forward Models make predictions about the location and the features of visual stimuli. This was done with the help from the Ideomotor Theory and its implications. While the majority of current studies has mainly investigated how visual features of stimuli are predicted, we tested the way in which spatial predictions occur. Interpreting results based on Forward Models facilitated the incorporation of the Ideomotor Theory to oculomotor control. In this section I will *briefly* expose topics that can be further discussed, improved or analysed to complement the experimental discussions of each experiment.

A Divided Attention Approach

Taking both Action-Effect experiments results together, there appears to exist a relation between Action-Effect association predictions and attention, which has been studied before (Cao and Gross, 2015; Schröger et al., 2015; Timm et al., 2013), with the possibility of attention being part of the FM's input as discussed on the next topic.

In Action-Effect experiments we are safe to assume subjects' ability to attend to at least two locations. In this section, I present a complementary way to understand the results, without discarding the interpretations discussed for each experiment. The point of view exposed bellow is not rigorous, and would certainly need further analysis to integrate divided attention as part of the experiments' interpretation.

In their 2008 article, Herwig & Waszak tested the functional differences of action–effect learning between stimulus based (exogenous) and intention based (endogenous) stimuli (Herwig and Waszak, 2008). In one experiment, attention was driven towards the effect, while in the other experiment it was driven away from the effect. They found that ideomotor learning depends on whether or not the action is selected with an intention. Most importantly, they showed that the "amount" of attention dedicated to the action's effect in ideomotor learning was not as important as the intention.

Awh and Pashler (2000) showed, that horizontally, vertically or diagonally oriented stimuli pairs can form two separate "streams" of attention, or *attentional streams*. Awh and Pashler's result could shed some light in explaining the results observed in the COHERENT MOTION DETECTION EXPERIMENT.

Under the divided attention perspective, the usage of features and space makes sense, and the cost of making a saccade is visible through the usage of available attentional streams. To simplify, I will take the saccade condition as an example. Before saccade execution, attentional resources are split into four possible regions (or zoomed-out, depending on the theory) to encompass the four possible locations. The action's acquired effects are pre-activated, to which the Forward Model outputs the expected consequences: an RDK at the acquisition position with a specific motion. It is uncertain if such information is passed to the FM before or after the saccade, because as I have mentioned before, the saccade is a second action. This pre-activation is inhibited as Sensory Attenuation phenomenon, neglected, translating into a reduced sensitivity; Attentional resources are updated (before or after sensory attenuation) in all stimulus dimensions (space and motion) to account for the attenuation. If the zoom lens hypothesis takes place, attention's resolution comprises primarily non-acquisition positions. If divided attention splits into contiguous regions, attention is split into the blue and the yellow rectangles of FIGURE 30, the green square being the intersection of both rectangles; thus, two

large attentional areas are created. Then, when the RDK appears, if it is at the neglected (acquisition) position, detection is very hard because no attentional stream is there and because of attenuation; if it is at any other location, its detectability will depend on the attentional stream it falls in: when the RDK appears at the spatiotopic test or the retinotopic control, subjects will be able to detect the coherent RDK because at both locations an attentional stream is allocated; finally, if the test occurs at the spatiotopic control (green square) they are very good at detecting the target stimulus because both areas —the horizontal and the vertical— overlap and encompass that position.

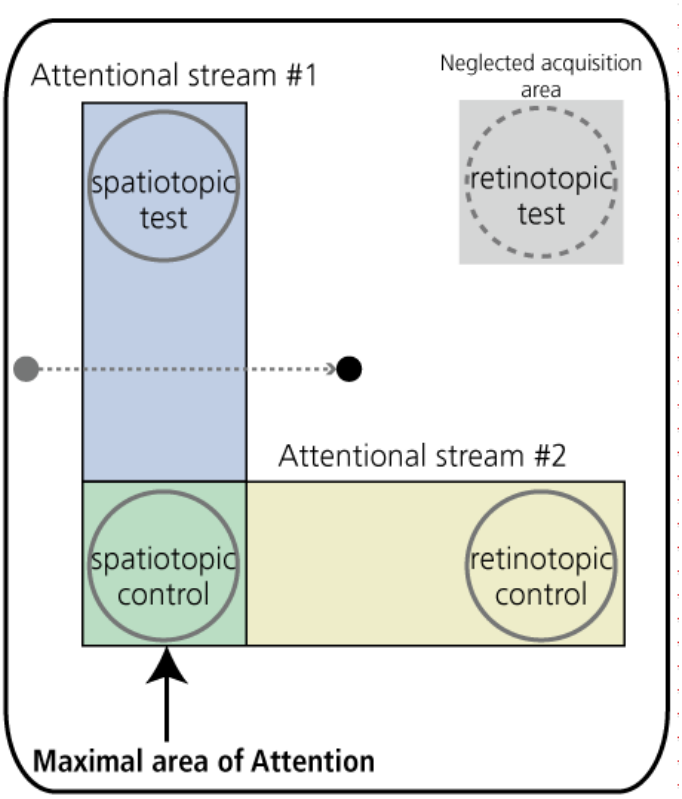

Figure 30 Divided attention hypothesis to the Action-Effect experiments. An attentional stream is created horizontally encompassing lower locations (yellow + green); another attentional stream is created vertically encompassing right locations (blue + green). The intersection between the yellow and the blue rectangles creates a square that represents the maximal attention region. The retinotopic test location is neglected after the repetition of the acquisition phases across the experiment, such that no attentional stream comprises that area, or it does, but very weakly.

In the MOTION DISCRIMINATION EXPERIMENT, the spatiotopic and retinotopic positions have similar scores because only one attentional stream is required to encompass them (to the left and to the right of fixation dot). The control position may have an attentional stream, but being so far away from the fixation dot, it remains difficult to discriminate motion with that eccentricity.

This hypothesis leaves the potential extension of this experiments to learn more about spatial divided attention and its links to action control. {from p. 87}

On Attention

Attention is a complex and almost "universal" cognitive process that deserves a whole chapter in almost any thesis on psychology or cognitive neuroscience. However, for brevity and clarity, I have only treated attention superficially.

One thing that was not mentioned before, is the fact that attention is not normally considered as part of the Forward Model's input. First, because it is difficult to conceive attention as an *input parameter*, and next because in principle attention should not be required to make the comparison of expected against actual inputs. Attention and sensory attenuation seem to have opposite effects, therefore there must exist a mechanism to privilege one over the other. There are situations in which attenuation could be undesired, like when walking on thin ice (Cao and Gross, 2015).

We can imagine that the Forward and Inverse Models displayed in FIGURE 3 {p. 15} receive input information *modified* by attention, in the sense that the "accuracy" or the "precision" of such information may be enhanced by attention, or diminished by the lack of it. This means that spatial- and feature-information may vary according to the allocated attentional resources of the ongoing action. If this were the case, this would imply that the prediction error may be different for two identical actions, consequently the sensory attenuation effect too. This notion can further be extended to determine if the Forward Model receives information concerning the reference frame under which the prediction should be made. Such information could be in the state or in the efference copy.

A way to further investigate this would be by means of the Remapping Attention Pointers theory (Cavanagh et al., 2010) or by considering visual attention as an *integration process* (Schneider, 1995; Stewart and Schütz, 2018). That would allow us to determine to what extent attention enhances action-effect predictions.

On multiple actions

I mentioned before that the experiments which tested Action-Effect associations contained two actions. This is relevant because in both experiments the acquisition phases involved one action: moving the gamepad stick either to the left or to the right. The test phases, on the other hand, could have a simple action in fixation, or an additional action: the saccade. The fixation condition "replicates" the acquisition phase because each action is followed by the same sensory effects, with the only difference that the target stimulus may appear at different locations. In the saccade

condition, however, there is an asymmetry: the gamepad action must be followed by a saccade[20] before perceiving the acquired effects.

We tried to fix this asymmetry by making the effects delay equal in all three stages (acquisition, fixation and saccade; see ACTION-EFFECT DELAY {p. 54}). This fixed only the *temporal asymmetry* but not the *effect asymmetry*. In saccade trials, the sensory effects of the first action are possibly mixed with the effects of the second. This can be seen in the overall difference between both conditions.

The current experimental design does not allow us to determine if under saccade trials —where two actions were involved— *two* Forward Models are activated at different stages, or if a single time-spanning Forward Model can integrate both actions. Yet another possibility would be that two different FM be activated, one in charge of manual predictions, while the other is in charge of visual predictions. Therefore, the possibility exists that spatial- and feature-information are *not* predicted by the same forward model, but by two separate, independent ones.

A way in which we could shed more light into this topic would be to make an experiment with a saccade acquisition phase and have only saccade test trials. Another experimental dataset could replicate the current design excluding saccade trials. Just then a comparison could be made.

On spatial manipulations

Displacements are manipulations of spatial information that we hypothesised would influence the vFM. Within Saccadic Suppression of Displacement, it is well known that when the displacement is too short, it goes unnoticed; also, that as displacement size increases, the likelihood of detecting it also increases. Concerning SSD experiments, we saw (in A LINK BETWEEN FEATURE AND SPATIAL INFORMATION {p. 108}) that spatial and feature information can indeed be confounded, giving the observer the impression that a change in a stimulus's feature (orientation in our case) is a change in its spatial properties. Previous findings in which Gabor patches were not used have shown that this confound did not take place (Demeyer et al., 2010; Tas et al., 2014, 2012). However, a change in the orientation of an object may involve a change of its position depending on its shape, for example, when rotating an asymmetrical object. Further investigation into this subject should be conducted in order to elucidate how orientation—and features in general—interact with space.

[20] Additionally, it had to end at the most 300 ms after the gamepad action.

On foveal and peripheral stimuli

Post-saccadic features are much richer than pre-saccadic ones. First, because post-saccadic features about a target are gathered foveally, while pre-saccadic are collected peripherally. Additionally, post-saccadic data represents *present* information, while pre-saccadic is past and discarded some time after the saccade.

Concerning the experiments in which we used RDKs as stimuli (1.1 COHERENT MOTION DETECTION EXPERIMENT {p. 49} and 1.2 MOTION DISCRIMINATION EXPERIMENT {p. 69}), we based our hypotheses on a similar experiment (Desantis et al., 2014). Nonetheless, as it was mentioned, the temporal and spatial properties of stimuli on both experiments were very different, rendering the results incomparable. Therefore, we cannot infer the temporal dynamics of the predictions carried out by the forward model in our experiments following Desantis's results (2014).

Implementing the action-effects experiments using neuroimaging methods would allow us to improve our knowledge on the peripheral processing of motion under the action-effects paradigm. Hypothetically, the valid time-window found by Desantis and colleagues would be different for a detection task like the present one in which action-effect associations are acquired peripherally; this, because of the differences mentioned in the experiments' discussion.

By analysing BOLD responses to motion in V5, one could test the reference frame under which motion action-effect associations are created and the way in which they are *accessed*, similar to what d'Avossa and colleagues did (2007), who found invariant signals in retinotopic, but varying signals in spatiotopic reference frame when subjects fixated at different screen locations.

Finally, following previous results (Melcher and Morrone, 2003) which showed that motion can be integrated across fixations, our experiments could be adapted to determine how Action-Effect associations can be affected with saccadic movements. For sure, presenting motion foveally would also help clarify how the visual system processes trans-saccadic motion correspondence, and its reference frame.

On the 40%–60% ratio

The 40%–60% ratio is of extreme relevance, as it might be a crucial piece of information to determine to what extent object recognition utilizes retinotopic and non-retinotopic information. As stated in THE PRIMARY VISUAL PATHWAY {p. 22}, information from both visual hemi-fields is distributed

across hemispheres according to the same 40%–60% proportion at the optic chiasm. The distribution of retinal information into both brain hemispheres is asymmetrical, because a region of the left (peripheral) visual field, for instance, is available only monocularly. Such asymmetry of information could help a high-level network to "locate retinal information in space", or in other words, to update an alleged spatiotopic map. {from p. on page 32}

REFERENCES

Adamo, M., Pun, C., Pratt, J., Ferber, S., 2008. Your divided attention, please! The maintenance of multiple attentional control sets over distinct regions in space. Cognition 107, 295–303. https://doi.org/10.1016/j.cognition.2007.07.003

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus Predictability Reduces Responses in Primary Visual Cortex. J. Neurosci. 30, 2960–2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

Andersen, R.A., 1989. Visual and Eye Movement Functions of the Posterior Parietal Cortex. Annual Review of Neuroscience 12, 377–403. https://doi.org/10.1146/annurev.ne.12.030189.002113

Andersen, R.A., Mountcastle, V.B., 1983. The influence of the angle of gaze upon the excitability of the light- sensitive neurons of the posterior parietal cortex. J. Neurosci. 3, 532–548.

Anzai, A., Peng, X., Van Essen, D.C., 2007. Neurons in monkey visual area V2 encode combinations of orientations. Nature Neuroscience 10, 1313–1321. https://doi.org/10.1038/nn1975

Atick, J.J., 2011. Could information theory provide an ecological theory of sensory processing? Network: Computation in Neural Systems 22, 4–44. https://doi.org/10.3109/0954898X.2011.638888

Atkeson, C.G., 1989. Learning arm kinematics and dynamics. Annu. Rev. Neurosci. 12, 157–183. https://doi.org/10.1146/annurev.ne.12.030189.001105

Atkinson, J., 1992. Early visual development: Differential functioning of parvocellular and magnocellular pathways. Eye 6, 129–135. https://doi.org/10.1038/eye.1992.28

Awh, E., Pashler, H., 2000. Evidence for split attentional foci. J Exp Psychol Hum Percept Perform 26, 834–846.

Barlow, H.B., 1961a. The coding of sensory messages, in: and, W.H.T. (Ed.), Current Problems in Animal Behaviour. Cambridge University Press, pp. 331–360.

Barlow, H.B., 1961b. Three points about lateral inhibition, in: Rosenblith, W.A. (Ed.), Sensory Communication. M.I.T. Press, Cambridge, Massachusetts, pp. 782–786.

Barlow, H.B., Levick, W.R., 1976. Threshold setting by the surround of cat retinal ganglion cells. The Journal of Physiology 259, 737–757. https://doi.org/10.1113/jphysiol.1976.sp011492

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language 68, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Basole, A., White, L.E., Fitzpatrick, D., 2003. Mapping multiple features in the population response of visual cortex. Nature 423, 986–990. https://doi.org/10.1038/nature01721

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M.C., Tassani, S., Piva, F., Perez-Amodio, S., Strippoli, P., Canaider, S., 2013. An estimation of the number of cells in the human body. Annals of Human Biology 40, 463–471. https://doi.org/10.3109/03014460.2013.807878

Blakemore, S.-J., Frith, C.D., Wolpert, D.M., 2001. The cerebellum is involved in predicting the sensory consequences of action. NeuroReport 12, 1879.

Blakemore, S.J., Wolpert, D., Frith, C., 2000. Why can't you tickle yourself? Neuroreport 11, R11-16.

Born, R.T., Bradley, D.C., 2005. Structure and function of visual area MT. Annu. Rev. Neurosci. 28, 157–189. https://doi.org/10.1146/annurev.neuro.26.041002.131052

Braddick, O.J., O'Brien, J.M.D., Wattam-Bell, J., Atkinson, J., Hartley, T., Turner, R., 2001. Brain Areas Sensitive to Coherent Visual Motion. Perception 30, 61–72. https://doi.org/10.1068/p3048

Bridgeman, B., 1973. Receptive fields in single cells of monkey visual cortex during visual tracking. Int. J. Neurosci. 6, 141–152.

Bridgeman, B., Hendry, D., Stark, L., 1975. Failure to detect displacement of the visual world during saccadic eye movements. Vision Res. 15, 719–722.

Brincat, S.L., Connor, C.E., 2004. Underlying principles of visual shape selectivity in posterior inferotemporal cortex. Nat Neurosci 7, 880–886. https://doi.org/10.1038/nn1278

Bruce, C.J., Goldberg, M.E., Bushnell, M.C., Stanton, G.B., 1985. Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. Journal of Neurophysiology 54, 714–734. https://doi.org/10.1152/jn.1985.54.3.714

Burr, D., Tozzi, A., Morrone, M.C., 2007. Neural mechanisms for timing visual events are spatially selective in real-world coordinates. Nature Neuroscience 10, 423–425. https://doi.org/10.1038/nn1874

Callaway, E.M., 2005. Structure and function of parallel pathways in the primate early visual system. The Journal of Physiology 566, 13–19. https://doi.org/10.1113/jphysiol.2005.088047

Cao, L., Gross, J., 2015. Attention Wins over Sensory Attenuation in a Sound Detection Task. PLOS ONE 10, e0136585. https://doi.org/10.1371/journal.pone.0136585

Cardoso-Leite, P., Mamassian, P., Schütz-Bosbach, S., Waszak, F., 2010. A New Look at Sensory Attenuation: Action-Effect Anticipation Affects Sensitivity, Not Response Bias. Psychol Sci 21, 1740–1745. https://doi.org/10.1177/0956797610389187

Carpenter, W.B., 1876. Principles of mental physiology: with their applications to the training and discipline of the mind, and the study of its morbid conditions. London: H.S. King & Co.

Cavanagh, P., Hunt, A.R., Afraz, A., Rolfs, M., 2010. Visual stability based on remapping of attention pointers. Trends Cogn. Sci. (Regul. Ed.) 14, 147–153. https://doi.org/10.1016/j.tics.2010.01.007

Colby, C.L., Duhamel, J.-R., 1996. Spatial representations for action in parietal cortex. Cognitive Brain Research 5, 105–115. https://doi.org/10.1016/S0926-6410(96)00046-8

Collins, T., 2014. Trade-off between spatiotopy and saccadic plasticity. Journal of Vision 14, 28–28. https://doi.org/10.1167/14.12.28

Collins, T., Rolfs, M., Deubel, H., Cavanagh, P., 2009. Post-saccadic location judgments reveal remapping of saccade targets to non-foveal locations. Journal of Vision 9, 29–29. https://doi.org/10.1167/9.5.29

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience 3, 201. https://doi.org/10.1038/nrn755

Curtis, C.E., Connolly, J.D., 2008. Saccade Preparation Signals in the Human Frontal and Parietal Cortices. Journal of Neurophysiology 99, 133–145. https://doi.org/10.1152/jn.00899.2007

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. Nat Neurosci 10, 249–255. https://doi.org/10.1038/nn1824

Demeyer, M., De Graef, P., Wagemans, J., Verfaillie, K., 2010. Object form discontinuity facilitates displacement discrimination across saccades. Journal of Vision 10, 17–17. https://doi.org/10.1167/10.6.17

den Ouden, H.E.M., Friston, K.J., Daw, N.D., McIntosh, A.R., Stephan, K.E., 2009. A Dual Role for Prediction Error in Associative Learning. Cereb Cortex 19, 1175–1185. https://doi.org/10.1093/cercor/bhn161

Desantis, A., Roussel, C., Waszak, F., 2014. The temporal dynamics of the perceptual consequences of action-effect prediction. Cognition 132, 243–250. https://doi.org/10.1016/j.cognition.2014.04.010

Desimone, R., 1991. Face-Selective Cells in the Temporal Cortex of Monkeys. Journal of Cognitive Neuroscience 3, 1–8. https://doi.org/10.1162/jocn.1991.3.1.1

Desimone, R., Albright, T.D., Gross, C.G., Bruce, C., 1984. Stimulus-selective properties of inferior temporal neurons in the macaque. J. Neurosci. 4, 2051–2062. https://doi.org/10.1523/JNEUROSCI.04-08-02051.1984

Deubel, H., Bridgeman, B., Schneider, W.X., 2004. Different effects of eyelid blinks and target blanking on saccadic suppression of displacement. Perception & Psychophysics 66, 772–778. https://doi.org/10.3758/BF03194971

Deubel, H., Bridgeman, B., Schneider, W.X., 1998. Immediate post-saccadic information mediates space constancy. Vision Research 38, 3147–3159. https://doi.org/10.1016/S0042-6989(98)00048-0

Deubel, H., Schneider, W.X., Bridgeman, B., 2002. Transsaccadic memory of position and form, in: Hyona, J., Munoz, D.P., Heide, W., Radach, R. (Eds.), Progress in Brain Research, The Brain's Eye: Neurobiological and Clinical Aspects of Oculomotor Research. Elsevier, pp. 165–180. https://doi.org/10.1016/S0079-6123(02)40049-0

Deubel, H., Schneider, W.X., Bridgeman, B., 1996. Postsaccadic target blanking prevents saccadic suppression of image displacement. Vision Research 36, 985–996. https://doi.org/10.1016/0042-6989(95)00203-0

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events: a neurophysiological study. Exp Brain Res 91, 176–180. https://doi.org/10.1007/BF00230027

Doré-Mazars, K., Pouget, P., Beauvillain, C., 2004. Attentional selection during preparation of eye movements. Psychological Research 69, 67–76. https://doi.org/10.1007/s00426-003-0166-1

Drew, T., Marigold, D.S., 2015. Taking the next step: cortical contributions to the control of locomotion. Current Opinion in Neurobiology, Motor circuits and action 33, 25–33. https://doi.org/10.1016/j.conb.2015.01.011

Dubner, R., Zeki, S.M., 1971. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. Brain Research 35, 528–532. https://doi.org/10.1016/0006-8993(71)90494-X

Duhamel, Colby, C., Goldberg, M., 1992. The updating of the representation of visual space in parietal cortex by intended eye movements. Science 255, 90–92. https://doi.org/10.1126/science.1553535

Duhamel, J.-R., Bremmer, F., BenHamed, S., Graf, W., 1997. Spatial invariance of visual receptive fields in parietal cortex neurons. Nature 389, 845–848. https://doi.org/10.1038/39865

Egner, T., Monti, J.M., Summerfield, C., 2010. Expectation and Surprise Determine Neural Population Responses in the Ventral Visual Stream. J. Neurosci. 30, 16601–16608. https://doi.org/10.1523/JNEUROSCI.2770-10.2010

Egner, T., Summerfield, C., 2013. Grounding predictive coding models in empirical neuroscience research. Behavioral and Brain Sciences 36, 210–211. https://doi.org/10.1017/S0140525X1200218X

Elsner, B., Hommel, B., 2001. Effect anticipation and action control. Journal of Experimental Psychology Human Perception & Performance 27, 229–240. https://doi.org/10.1037/0096-1523.27.1.229

Engbert, K., Wohlschläger, A., Haggard, P., 2008. Who is causing what? The sense of agency is relational and efferent-triggered. Cognition 107, 693–704. https://doi.org/10.1016/j.cognition.2007.07.021

Eriksen, C.W., St. James, J.D., 1986. Visual attention within and around the field of focal attention: A zoom lens model. Perception & Psychophysics 40, 225–240. https://doi.org/10.3758/BF03211502

Evarts, E.V., 1986. Motor Cortex Output in Primates, in: Jones, E.G., Peters, A. (Eds.), Sensory-Motor Areas and Aspects of Cortical Connectivity, Cerebral Cortex. Springer US, Boston, MA, pp. 217–241. https://doi.org/10.1007/978-1-4613-2149-1_6

Eymond, C., Cavanagh, P., Collins, T., 2016. Feature-based attention across saccades and immediate postsaccadic selection. Atten Percept Psychophys 78, 1293–1301. https://doi.org/10.3758/s13414-016-1110-y

Ezzati, A., Golzar, A., Afraz, A.S.R., 2008. Topography of the motion aftereffect with and without eye movements. Journal of Vision 8, 23–23. https://doi.org/10.1167/8.14.23

Feldman, H., Friston, K., 2010. Attention, Uncertainty, and Free-Energy. Front. Hum. Neurosci. 4. https://doi.org/10.3389/fnhum.2010.00215

Felleman, D.J., Van Essen, D.C., 1991. Distributed Hierarchical Processing in the Primate Cerebral Cortex. Cereb Cortex 1, 1–47. https://doi.org/10.1093/cercor/1.1.1-a

Folk, C., Remington, R., 2006. Top-down modulation of preattentive processing: Testing the recovery account of contingent capture. Visual Cognition 14. https://doi.org/10.1080/13506280500193545

Foulsham, T., 2015. Eye movements and their functions in everyday tasks. Eye (Lond) 29, 196–199. https://doi.org/10.1038/eye.2014.275

Freud, E., Plaut, D.C., Behrmann, M., 2016. 'What' Is Happening in the Dorsal Visual Pathway. Trends in Cognitive Sciences 20, 773–784. https://doi.org/10.1016/j.tics.2016.08.003

Friston, K., 2005. A theory of cortical responses. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 815–836. https://doi.org/10.1098/rstb.2005.1622

Fuchs, A.F., Luschei, E.S., 1970. Firing patterns of abducens neurons of alert monkeys in relationship to horizontal eye movement. Journal of Neurophysiology 33, 382–392. https://doi.org/10.1152/jn.1970.33.3.382

Galletti, C., Battaglini, P.P., Fattori, P., 1995. Eye Position Influence on the Parieto-occipital Area PO (V6) of the Macaque Monkey. European Journal of Neuroscience 7, 2486–2501. https://doi.org/10.1111/j.1460-9568.1995.tb01047.x

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. J. Neurosci. 28, 3988–3999. https://doi.org/10.1523/JNEUROSCI.5476-07.2008

Gazzaniga, M., Ivry, R.B., 2013. Cognitive Neuroscience: The Biology of the Mind. WW Norton.

Golomb, J.D., Chun, M.M., Mazer, J.A., 2008. The Native Coordinate System of Spatial Attention Is Retinotopic. J. Neurosci. 28, 10654–10662. https://doi.org/10.1523/JNEUROSCI.2525-08.2008

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

Green, D.M., Dai, H., 1991. Probability of being correct with 1 of m orthogonal signals. Perception & Psychophysics 49, 100–101. https://doi.org/10.3758/BF03211621

Green, D.M., Swets, J.A., 1966. Signal detection theory and psychophysics, Signal detection theory and psychophysics. John Wiley, Oxford, England.

Greenwald, A.G., 1970. Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. Psychological Review 77, 73–99. https://doi.org/10.1037/h0028689

Gross, C.G., Rocha-Miranda, C.E., Bender, D.B., 1972. Visual properties of neurons in inferotemporal cortex of the Macaque. Journal of Neurophysiology 35, 96–111. https://doi.org/10.1152/jn.1972.35.1.96

Grüsser, O.-J., Krizic¸, A., Weiss, L.-R., 1987. Afterimage movement during saccades in the dark. Vision Research 27, 215–226. https://doi.org/10.1016/0042-6989(87)90184-2

Habak, C., Wilkinson, F., Zakher, B., Wilson, H.R., 2004. Curvature population coding for complex shapes in human vision. Vision Research 44, 2815–2823. https://doi.org/10.1016/j.visres.2004.06.019

Haggard, P., Clark, S., Kalogeras, J., 2002. Voluntary action and conscious awareness. Nature Neuroscience 5, 382–385. https://doi.org/10.1038/nn827

Haggard, P., Miall, R.C., Wade, D., Fowler, S., Richardson, A., Anslow, P., Stein, J., 1995. Damage to cerebellocortical pathways after closed head injury: a behavioural and magnetic resonance imaging study. Journal of Neurology, Neurosurgery & Psychiatry 58, 433–438. https://doi.org/10.1136/jnnp.58.4.433

Hamker, F.H., 2003. The reentry hypothesis: linking eye movements to visual perception. Journal of Vision 3, 14–14. https://doi.org/10.1167/3.11.14

Harleß, E., 1861. Der apparat des willens. Zeitschrift für Philosophie und philosophische Kritik 38, 50–73.

Harrison, N.R., Ziessler, M., 2016. Effect Anticipation Affects Perceptual, Cognitive, and Motor Phases of Response Preparation: Evidence from an Event-Related Potential (ERP) Study. Front Hum Neurosci 10. https://doi.org/10.3389/fnhum.2016.00005

Hayhoe, M., Lachter, J., Feldman, J., 1991. Integration of Form across Saccadic Eye Movements. Perception 20, 393–402. https://doi.org/10.1068/p200393

Hegdé, J., Essen, D.C.V., 2000. Selectivity for Complex Shapes in Primate Visual Area V2. J. Neurosci. 20, RC61–RC61. https://doi.org/10.1523/JNEUROSCI.20-05-j0001.2000

Helmholtz, H. von, 1867. Handbuch der physiologischen Optik. Leipzig: Leopold Voss.

Henderson, J.M., 1994. Two representational systems in dynamic visual identification. Journal of Experimental Psychology: General 123, 410–426. https://doi.org/10.1037/0096-3445.123.4.410

Herbart, J.F., 1826. A text-book in psychology; an attempt to found the science of psychology on experience, metaphysics, and mathematics. New York, D. Appleton and company.

Herwig, A., Prinz, W., Waszak, F., 2007. Two modes of sensorimotor integration in intention-based and stimulus-based actions. The Quarterly Journal of Experimental Psychology 60, 1540–1554. https://doi.org/10.1080/17470210601119134

Herwig, A., Schneider, W.X., 2014. Predicting object features across saccades: evidence from object recognition and visual search. J Exp Psychol Gen 143, 1903–1922. https://doi.org/10.1037/a0036781

Herwig, A., Waszak, F., 2008. Intention and attention in ideomotor learning. Quarterly journal of experimental psychology (2006) 62, 219–27. https://doi.org/10.1080/17470210802373290

Hollingworth, A., Richard, A.M., Luck, S.J., 2008. Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. Journal of Experimental Psychology: General 137, 163–181. https://doi.org/10.1037/0096-3445.137.1.163

Hommel, B., Müsseler, J., Aschersleben, G., Prinz, W., 2001. The Theory of Event Coding (TEC): a framework for perception and action planning. Behav Brain Sci 24, 849–878; discussion 878-937.

Honda, H., 1991. The time courses of visual mislocalization and of extraretinal eye position signals at the time of vertical saccades. Vision Research 31, 1915–1921. https://doi.org/10.1016/0042-6989(91)90186-9

Honda, H., 1989. Perceptual localization of visual stimuli flashed during saccades. Percept Psychophys 45, 162–174.

Howarth, M., Walmsley, L., Brown, T.M., 2014. Binocular Integration in the Mouse Lateral Geniculate Nuclei. Current Biology 24, 1241–1247. https://doi.org/10.1016/j.cub.2014.04.014

Hubel, D.H., Wiesel, T.N., 1977. Functional Architecture of Macaque Monkey Visual Cortex. Proceedings of the Royal Society of London. Series B, Biological Sciences 198, 1–59.

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, 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Irwin, D., Gordon, R., 1998. Eye Movements, Attention and Trans-saccadic Memory. Visual Cognition 5, 127–155. https://doi.org/10.1080/713756783

Irwin, D.E., 1996. Integrating Information Across Saccadic Eye Movements. Curr Dir Psychol Sci 5, 94–100. https://doi.org/10.1111/1467-8721.ep10772833

Irwin, D.E., Andrews, R.V., 1996. Integration and accumulation of information across saccadic eye movements, in: Attention and Performance 16: Information Integration in Perception and Communication, Attention and Performance. The MIT Press, Cambridge, MA, US, pp. 125–155.

Ishikawa, T., Tomatsu, S., Izawa, J., Kakei, S., 2016. The cerebro-cerebellum: Could it be loci of forward models? Neuroscience Research, Body representation in the brain 104, 72–79. https://doi.org/10.1016/j.neures.2015.12.003

Ito, M., Tamura, H., Fujita, I., Tanaka, K., 1995. Size and position invariance of neuronal responses in monkey inferotemporal cortex. Journal of Neurophysiology 73, 218–226. https://doi.org/10.1152/jn.1995.73.1.218

Itthipuripat, S., Garcia, J.O., Serences, J.T., 2013. Temporal dynamics of divided spatial attention. Journal of Neurophysiology 109, 2364–2373. https://doi.org/10.1152/jn.01051.2012

James, W., 1890. The principles of psychology, Vol I., The principles of psychology, Vol I. Henry Holt and Co, New York, NY, US. https://doi.org/10.1037/10538-000

Janczyk, M., Heinemann, A., Pfister, R., 2012. Instant Attraction: Immediate Action-Effect Bindings Occur for Both, Stimulus- and Goal-Driven Actions. Front Psychol 3. https://doi.org/10.3389/fpsyg.2012.00446

Jans, B., Peters, J.C., De Weerd, P., 2010. Visual spatial attention to multiple locations at once: The jury is still out. Psychological Review 117, 637–682. https://doi.org/10.1037/a0019082

Jiang, J., Summerfield, C., Egner, T., 2013. Attention sharpens the distinction between expected and unexpected percepts in the visual brain. J. Neurosci. 33, 18438–18447. https://doi.org/10.1523/JNEUROSCI.3308-13.2013

Joiner, W.M., Cavanaugh, J., Wurtz, R.H., 2011. Modulation of shifting receptive field activity in frontal eye field by visual salience. Journal of Neurophysiology 106, 1179–1190. https://doi.org/10.1152/jn.01054.2010

Jonikaitis, D., Theeuwes, J., 2013. Dissociating oculomotor contributions to spatial and feature-based selection. J. Neurophysiol. 110, 1525–1534. https://doi.org/10.1152/jn.00275.2013

Jordan, M.I., Rumelhart, D.E., 1992. Forward Models: Supervised Learning with a Distal Teacher. Cognitive Science 16, 307–354. https://doi.org/10.1207/s15516709cog1603_1

Kandel, E.R., Schwartz, J.H., Jessell, T.M., Department of Biochemistry and Molecular Biophysics Thomas Jessell, Siegelbaum, S., Hudspeth, A., 2000. Principles of neural science. McGraw-hill New York.

Kanwisher, N., Woods, R.P., Iacoboni, M., Mazziotta, J.C., 1997. A Locus in Human Extrastriate Cortex for Visual Shape Analysis. Journal of Cognitive Neuroscience 9, 133–142. https://doi.org/10.1162/jocn.1997.9.1.133

Kapoula, Z., Robinson, D.A., 1986. Saccadic undershoot is not inevitable: Saccades can be accurate. Vision Research 26, 735–743. https://doi.org/10.1016/0042-6989(86)90087-8

Kastner, S., Weerd, P.D., Desimone, R., Ungerleider, L.G., 1998. Mechanisms of Directed Attention in the Human Extrastriate Cortex as Revealed by Functional MRI. Science 282, 108–111. https://doi.org/10.1126/science.282.5386.108

Klier, E.M., Wang, H., Crawford, J.D., 2001. The superior colliculus encodes gaze commands in retinal coordinates. Nat Neurosci 4, 627–632. https://doi.org/10.1038/88450

Knapen, T., Rolfs, M., Wexler, M., Cavanagh, P., 2010. The reference frame of the tilt aftereffect. Journal of Vision 10, 8–8. https://doi.org/10.1167/10.1.8

Knapen, T., Rolfs, M., Wexler, M., Cavanagh, P., 2009. The reference frame of the tilt aftereffect. Journal of Vision 10, 8–8. https://doi.org/10.1167/10.1.8

Kok, P., Rahnev, D., Jehee, J.F.M., Lau, H.C., de Lange, F.P., 2012. Attention Reverses the Effect of Prediction in Silencing Sensory Signals. Cereb Cortex 22, 2197–2206. https://doi.org/10.1093/cercor/bhr310

Komatsu, H., Wurtz, R.H., 1989. Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. Journal of Neurophysiology 62, 31–47. https://doi.org/10.1152/jn.1989.62.1.31

Kuffler, S.W., 1953. Discharge patterns and functional organization of mammalian retina. Journal of Neurophysiology 16, 37–68. https://doi.org/10.1152/jn.1953.16.1.37

Kühn, S., Seurinck, R., Fias, W., Waszak, F., 2010. The Internal Anticipation of Sensory Action Effects: When Action Induces FFA and PPA Activity. Front Hum Neurosci 4. https://doi.org/10.3389/fnhum.2010.00054

Kunde, W., 2003. Temporal response-effect compatibility. Psychological Research 67, 153–159. https://doi.org/10.1007/s00426-002-0114-5

Laycock, T., 1860. Mind and brain, or The correlations of consciousness and organisation, with their applications to philosophy, zoology, physiology, mental pathology, and the practice of medicine, Vol 1., Mind and brain, or The correlations of consciousness and organisation, with their applications to philosophy, zoology, physiology, mental pathology, and the practice of medicine, Vol 1. Sutherland and Knox, Edinburgh, Scotland. https://doi.org/10.1037/12161-000

Lee, D., Quessy, S., 2003. Activity in the Supplementary Motor Area Related to Learning and Performance During a Sequential Visuomotor Task. Journal of Neurophysiology 89, 1039–1056. https://doi.org/10.1152/jn.00638.2002

Lee, T.S., Mumford, D., 2003. Hierarchical Bayesian inference in the visual cortex. J. Opt. Soc. Am. A, JOSAA 20, 1434–1448. https://doi.org/10.1364/JOSAA.20.001434

Libet, B., Gleason, C.A., Wright, E.W., Pearl, D.K., 1983. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. Brain 106 (Pt 3), 623–642.

Livingstone, M., Hubel, D., 1988. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 240, 740–749. https://doi.org/10.1126/science.3283936

López-Aranda, M.F., López-Téllez, J.F., Navarro-Lobato, I., Masmudi-Martín, M., Gutiérrez, A., Khan, Z.U., 2009. Role of Layer 6 of V2 Visual Cortex in Object-Recognition Memory. Science 325, 87–89. https://doi.org/10.1126/science.1170869

Lotze, H., 1852. Medicinische psychologie; oder, Physiologie der seele. Weidmann, Leipzig.

Lui, L.L., Bourne, J.A., Rosa, M.G.P., 2006. Functional Response Properties of Neurons in the Dorsomedial Visual Area of New World Monkeys (Callithrix jacchus). Cereb Cortex 16, 162–177. https://doi.org/10.1093/cercor/bhi094

Lynch, J.C., Hoover, J.E., Strick, P.L., 1994. Input to the primate frontal eye field from the substantia nigra, superior colliculus, and dentate nucleus demonstrated by transneuronal transport. Exp Brain Res 100, 181–186. https://doi.org/10.1007/BF00227293

Makris, N., Kennedy, D.N., McInerney, S., Sorensen, A.G., Wang, R., Caviness, V.S., Pandya, D.N., 2005. Segmentation of Subcomponents within the Superior Longitudinal Fascicle in Humans: A Quantitative, In Vivo, DT-MRI Study. Cereb Cortex 15, 854–869. https://doi.org/10.1093/cercor/bhh186

Martin, A., 2007. The Representation of Object Concepts in the Brain. Annual Review of Psychology 58, 25–45. https://doi.org/10.1146/annurev.psych.57.102904.190143

Masland, R.H., 2012. The Neuronal Organization of the Retina. Neuron 76, 266–280. https://doi.org/10.1016/j.neuron.2012.10.002

Mathôt, S., Theeuwes, J., 2009. Evidence for the predictive remapping of visual attention. Exp Brain Res 200, 117. https://doi.org/10.1007/s00221-009-2055-3

Matin, L., 1982. Visual Localization and Eye Movements, in: Wertheim, A.H., Wagenaar, W.A., Leibowitz, H.W. (Eds.), Tutorials on Motion Perception, NATO Conference Series. Springer US, Boston, MA, pp. 101–156. https://doi.org/10.1007/978-1-4613-3569-6_5

Maunsell, J.H., Nealey, T.A., DePriest, D.D., 1990. Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. J. Neurosci. 10, 3323–3334. https://doi.org/10.1523/JNEUROSCI.10-10-03323.1990

Maunsell, J.H., Van Essen, D.C., 1983. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. Journal of Neurophysiology 49, 1127–1147. https://doi.org/10.1152/jn.1983.49.5.1127

McConkie, G.W., Currie, C.B., 1996. Visual stability across saccades while viewing complex pictures. J Exp Psychol Hum Percept Perform 22, 563–581.

McKyton, A., Zohary, E., 2007. Beyond Retinotopic Mapping: The Spatial Representation of Objects in the Human Lateral Occipital Complex. Cereb Cortex 17, 1164–1172. https://doi.org/10.1093/cercor/bhl027

McMains, S.A., Somers, D.C., 2004. Multiple Spotlights of Attentional Selection in Human Visual Cortex. Neuron 42, 677–686. https://doi.org/10.1016/S0896-6273(04)00263-6

McSorley, E., Gilchrist, I.D., McCloy, R., 2019. The programming of sequences of saccades. Exp Brain Res 237, 1009–1018. https://doi.org/10.1007/s00221-019-05481-7

Melcher, D., 2007. Predictive remapping of visual features precedes saccadic eye movements. Nat Neurosci 10, 903–907. https://doi.org/10.1038/nn1917

Melcher, D., 2005. Spatiotopic Transfer of Visual-Form Adaptation across Saccadic Eye Movements. Current Biology 15, 1745–1748. https://doi.org/10.1016/j.cub.2005.08.044

Melcher, D., Morrone, M.C., 2003. Spatiotopic temporal integration of visual motion across saccadic eye movements. Nat Neurosci 6, 877–881. https://doi.org/10.1038/nn1098

Merriam, E.P., Colby, C.L., 2005. Active Vision in Parietal and Extrastriate Cortex. Neuroscientist 11, 484–493. https://doi.org/10.1177/1073858405276871

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

Miall, R.C., Wolpert, D.M., 1995. The Cerebellum as a Predictive Model of the Motor System: A Smith Predictor Hypothesis, in: Ferrell, W.R., Proske, U. (Eds.), Neural Control of Movement. Springer US, Boston, MA, pp. 215–223. https://doi.org/10.1007/978-1-4615-1985-0_27

Michel, M.M., Chen, Y., Geisler, W.S., Seidemann, E., 2013. An illusion predicted by V1 population activity implicates cortical topography in shape perception. Nature Neuroscience 16, 1477–1483. https://doi.org/10.1038/nn.3517

Milner, A.D., 2017. How do the two visual streams interact with each other? Exp Brain Res 235, 1297–1308. https://doi.org/10.1007/s00221-017-4917-4

Mishkin, M., Ungerleider, L.G., Macko, K.A., 1983. Object vision and spatial vision: two cortical pathways. Trends in Neurosciences 6, 414–417. https://doi.org/10.1016/0166-2236(83)90190-X

Moore, T., Armstrong, K.M., 2003. Selective gating of visual signals by microstimulation of frontal cortex. Nature 421, 370–373. https://doi.org/10.1038/nature01341

Morin, R.E., Grant, D.A., 1955. Learning and performance on a key-pressing task as function of the degree of spatial stimulus-response correspondence. J Exp Psychol 49, 39–47.

Mumford, D., 1992. On the computational architecture of the neocortex. Biol. Cybern. 66, 241–251. https://doi.org/10.1007/BF00198477

Mustari, M.J., Ono, S., Das, V.E., 2009. Signal Processing and Distribution in Cortical-Brainstem Pathways for Smooth Pursuit Eye Movements. Annals of the New York Academy of Sciences 1164, 147–154. https://doi.org/10.1111/j.1749-6632.2009.03859.x

Nassi, J.J., Callaway, E.M., 2009. Parallel processing strategies of the primate visual system. Nature Reviews Neuroscience 10, 360–372. https://doi.org/10.1038/nrn2619

Nattkemper, D., Prinz, W., 1997. Stimulus and response anticipation in a serial reaction task. Psychol. Res 60, 98–112. https://doi.org/10.1007/BF00419683

Norman, J., 2002. Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. Behavioral and Brain Sciences 25, 73–96. https://doi.org/10.1017/S0140525X0200002X

O'Regan, J.K., 1992. Solving the "real" mysteries of visual perception: the world as an outside memory. Can J Psychol 46, 461–488.

O'Regan, J.K., Regan, Noë, A., 2001. A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24, 939–973. https://doi.org/10.1017/S0140525X01000115

Passingham, R.E., 1993. The frontal lobes and voluntary action, The frontal lobes and voluntary action. Oxford University Press, New York, NY, US.

Penfield, W., Welch, K., 1951. The Supplementary Motor Area of the Cerebral Cortex: A Clinical and Experimental Study. AMA Arch NeurPsych 66, 289–317. https://doi.org/10.1001/archneurpsyc.1951.02320090038004

Perry, V.H., Oehler, R., Cowey, A., 1984. Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. Neuroscience 12, 1101–1123. https://doi.org/10.1016/0306-4522(84)90006-X

Pohl, W., 1973. Dissociation of spatial discrimination deficits following frontal and parietal lesions in monkeys. Journal of Comparative and Physiological Psychology 82, 227–239. https://doi.org/10.1037/h0033922

Pollatsek, A., Rayner, K., Collins, W.E., 1984. Integrating pictorial information across eye movements. J Exp Psychol Gen 113, 426–442.

Posner, M., Cohen, Y., 1984. Components of visual orienting. Attention and performance X: Control of language processes 32, 531.

Posner, M.I., Petersen, S.E., 1990. The attention system of the human brain. Annu. Rev. Neurosci. 13, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Poth, C.H., Herwig, A., Schneider, W.X., 2015. Breaking Object Correspondence Across Saccadic Eye Movements Deteriorates Object Recognition. Front. Syst. Neurosci. 9. https://doi.org/10.3389/fnsys.2015.00176

Poth, C.H., Schneider, W.X., 2016. Breaking object correspondence across saccades impairs object recognition: The role of color and luminance. Journal of Vision 16, 1–1. https://doi.org/10.1167/16.11.1

Pratt, J., Hommel, B., 2003. Symbolic Control of Visual Attention: The Role of Working Memory and Attentional Control Settings. Journal of experimental psychology. Human perception and performance 29, 835–45. https://doi.org/10.1037/0096-1523.29.5.835

Prime, S.L., Tsotsos, L., Keith, G.P., Crawford, J.D., 2007. Visual memory capacity in transsaccadic integration. Exp Brain Res 180, 609–628. https://doi.org/10.1007/s00221-007-0885-4

Prinz, W., 1990. A Common Coding Approach to Perception and Action, in: Neumann, O., Prinz, Wolfgang (Eds.), Relationships Between Perception and Action. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 167–201.

Prinz, W., 1984. Modes of Linkage Between Perception and Action, in: Prinz, W., Sanders, A.F. (Eds.), Cognition and Motor Processes. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 185–193. https://doi.org/10.1007/978-3-642-69382-3_13

Purves, D., 2007. Neuroscience, Fourth Edition, 4th edition. ed. Sinauer Associates, Inc., Sunderland, Mass.

Rao, R.P.N., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience 2, 79–87. https://doi.org/10.1038/4580

Reeves, A., Sperling, G., 1986. Attention gating in short-term visual memory. Psychological Review 93, 180–206. https://doi.org/10.1037/0033-295X.93.2.180

Riechelmann, E., Pieczykolan, A., Horstmann, G., Herwig, A., Huestegge, L., 2017. Spatio-temporal dynamics of action-effect associations in oculomotor control. Acta Psychologica 180, 130–136. https://doi.org/10.1016/j.actpsy.2017.09.003

Riesenhuber, M., Poggio, T., 2000. Models of object recognition. Nat Neurosci 3, 1199–1204. https://doi.org/10.1038/81479

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L., 1996. Premotor cortex and the recognition of motor actions. Cognitive Brain Research, Mental representations of motor acts 3, 131–141. https://doi.org/10.1016/0926-6410(95)00038-0

Rizzolatti, G., Fogassi, L., Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews Neuroscience 2, 661. https://doi.org/10.1038/35090060

Robinson, D.L., Goldberg, M.E., Stanton, G.B., 1978. Parietal association cortex in the primate: sensory mechanisms and behavioral modulations. Journal of Neurophysiology 41, 910–932. https://doi.org/10.1152/jn.1978.41.4.910

Rock, I., Ebenholtz, S., 1962. Stroboscopic Movement Based on Change of Phenomenal Rather than Retinal Location. The American Journal of Psychology 75, 193–207. https://doi.org/10.2307/1419603

Rodieck, R.W., Watanabe, M., 1993. Survey of the morphology of macaque retinal ganglion cells that project to the pretectum, superior colliculus, and parvicellular laminae of the lateral geniculate nucleus. Journal of Comparative Neurology 338, 289–303. https://doi.org/10.1002/cne.903380211

Roland, P.E., Larsen, B., Lassen, N.A., Skinhoj, E., 1980. Supplementary motor area and other cortical areas in organization of voluntary movements in man. Journal of Neurophysiology 43, 118–136. https://doi.org/10.1152/jn.1980.43.1.118

Rolfs, M., Jonikaitis, D., Deubel, H., Cavanagh, P., 2011. Predictive remapping of attention across eye movements. Nature Neuroscience 14, 252. https://doi.org/10.1038/nn.2711

Roussel, C., Hughes, G., Waszak, F., 2013. A preactivation account of sensory attenuation. Neuropsychologia 51, 922–929. https://doi.org/10.1016/j.neuropsychologia.2013.02.005

Sakmann, B., Creutzfeldt, O.D., 1969. Scotopic and mesopic light adaptation in the cat's retina. Pflugers Arch. 313, 168–185. https://doi.org/10.1007/BF00586245

Sanderson, K.J., Bishop, P.O., Darian-Smith, I., 1971. The properties of the binocular receptive fields of lateral geniculate neurons. Exp Brain Res 13, 178–207. https://doi.org/10.1007/BF00234085

Schafer, E.W.P., Marcus, M.M., 1973. Self-Stimulation Alters Human Sensory Brain Responses. Science 181, 175–177. https://doi.org/10.1126/science.181.4095.175

Schall, J.D., 2004. On the role of frontal eye field in guiding attention and saccades. Vision Research, Visual Attention 44, 1453–1467. https://doi.org/10.1016/j.visres.2003.10.025

Schiller, P.H., Logothetis, N.K., 1990. The color-opponent and broad-band channels of the primate visual system. Trends in Neurosciences 13, 392–398. https://doi.org/10.1016/0166-2236(90)90117-S

Schiller, P.H., Stryker, M., 1972. Single-unit recording and stimulation in superior colliculus of the alert rhesus monkey. Journal of Neurophysiology 35, 915–924. https://doi.org/10.1152/jn.1972.35.6.915

Schiller, P.H., True, S.D., Conway, J.L., 1980. Deficits in eye movements following Frontal Eye-Field and Superior Colliculus ablations. Journal of Neurophysiology 44, 1175–1189. https://doi.org/10.1152/jn.1980.44.6.1175

Schmid, M.C., Schmiedt, J.T., Peters, A.J., Saunders, R.C., Maier, A., Leopold, D.A., 2013. Motion-Sensitive Responses in Visual Area V4 in the Absence of Primary Visual Cortex. J. Neurosci. 33, 18740–18745. https://doi.org/10.1523/JNEUROSCI.3923-13.2013

Schneider, W.X., 2013. Selective visual processing across competition episodes: a theory of task-driven visual attention and working memory. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 20130060. https://doi.org/10.1098/rstb.2013.0060

Schneider, W.X., 1995. VAM: A neuro-cognitive model for visual attention control of segmentation, object recognition, and space-based motor action. Visual Cognition 2, 331–376. https://doi.org/10.1080/13506289508401737

Schröger, E., Marzecová, A., SanMiguel, I., 2015. Attention and prediction in human audition: a lesson from cognitive psychophysiology. European Journal of Neuroscience 41, 641–664. https://doi.org/10.1111/ejn.12816

Serrien, D.J., Strens, L.H.A., Oliviero, A., Brown, P., 2002. Repetitive transcranial magnetic stimulation of the supplementary motor area (SMA) degrades bimanual movement control in humans. Neuroscience Letters 328, 89–92. https://doi.org/10.1016/S0304-3940(02)00499-8

Sheinberg, D.L., Logothetis, N.K., 1997. The role of temporal cortical areas in perceptual organization. PNAS 94, 3408–3413. https://doi.org/10.1073/pnas.94.7.3408

Sherrington, C.S., 1906. The integrative action of the nervous system, The integrative action of the nervous system. Yale University Press, New Haven, CT, US. https://doi.org/10.1037/13798-000

Shima, K., Tanji, J., 1998. Both Supplementary and Presupplementary Motor Areas Are Crucial for the Temporal Organization of Multiple Movements. Journal of Neurophysiology 80, 3247–3260. https://doi.org/10.1152/jn.1998.80.6.3247

Shin, Y.K., Proctor, R.W., Capaldi, E.J., 2010. A review of contemporary ideomotor theory. Psychological Bulletin 136, 943–974. https://doi.org/10.1037/a0020541

Shulman, G.L., Remington, R.W., McLean, J.P., 1979. Moving attention through visual space. Journal of Experimental Psychology: Human Perception and Performance 5, 522–526. https://doi.org/10.1037/0096-1523.5.3.522

Sincich, L.C., Park, K.F., Wohlgemuth, M.J., Horton, J.C., 2004. Bypassing V1: a direct geniculate input to area MT. Nat Neurosci 7, 1123–1128. https://doi.org/10.1038/nn1318

Sommer, M.A., Wurtz, R.H., 2006. Influence of the thalamus on spatial visual processing in frontal cortex. Nature 444, 374–377. https://doi.org/10.1038/nature05279

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, 45–63. https://doi.org/10.1152/jn.1983.49.1.45

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, 300–318.

Sperry, R.W., 1950. Neural basis of the spontaneous optokinetic response produced by visual inversion. Journal of Comparative and Physiological Psychology 43, 482–489. https://doi.org/10.1037/h0055479

Sprague, J.M., 1996. Neural mechanisms of visual orienting responses, in: Norita, M., Bando, T., Stein, B.E. (Eds.), Progress in Brain Research, Extrageniculostriate Mechanisms Underlying Visually-Guided Orientation Behavior. Elsevier, pp. 1–15. https://doi.org/10.1016/S0079-6123(08)63317-8

Srinivasan, M.V., Laughlin, S.B., Dubs, A., 1982. Predictive coding: a fresh view of inhibition in the retina. Proc. R. Soc. Lond., B, Biol. Sci. 216, 427–459. https://doi.org/10.1098/rspb.1982.0085

Stein, J., 2009. Cerebellar forward models to control movement. The Journal of Physiology 587, 299. https://doi.org/10.1113/jphysiol.2008.167627

Stewart, E.E.M., Schütz, A.C., 2018. Attention modulates trans-saccadic integration. Vision Research 142, 1–10. https://doi.org/10.1016/j.visres.2017.11.006

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.-M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nature Neuroscience 11, 1004–1006. https://doi.org/10.1038/nn.2163

Summerfield, C., Wyart, V., Mareike Johnen, V., de Gardelle, V., 2011. Human Scalp Electroencephalography Reveals that Repetition Suppression Varies with Expectation. Front. Hum. Neurosci. 5. https://doi.org/10.3389/fnhum.2011.00067

Sun, L.D., Goldberg, M.E., 2016. Corollary Discharge and Oculomotor Proprioception: Cortical Mechanisms for Spatially Accurate Vision. Annu. Rev. Vis. Sci. 2, 61–84. https://doi.org/10.1146/annurev-vision-082114-035407

Tas, A.C., Moore, C.M., Hollingworth, A., 2014. The representation of the saccade target object depends on visual stability. Vis cogn 22, 1042–1046. https://doi.org/10.1080/13506285.2014.960671

Tas, A.C., Moore, C.M., Hollingworth, A., 2012. An object-mediated updating account of insensitivity to transsaccadic change. Journal of Vision 12, 18–18. https://doi.org/10.1167/12.11.18

Tian, J.R., Lynch, J.C., 1996. Corticocortical input to the smooth and saccadic eye movement subregions of the frontal eye field in Cebus monkeys. Journal of Neurophysiology 76, 2754–2771. https://doi.org/10.1152/jn.1996.76.4.2754

Timm, J., SanMiguel, I., Saupe, K., Schröger, E., 2013. The N1-suppression effect for self-initiated sounds is independent of attention. BMC Neurosci 14, 2. https://doi.org/10.1186/1471-2202-14-2

Todorovic, A., Ede, F. van, Maris, E., Lange, F.P. de, 2011. Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. J. Neurosci. 31, 9118–9123. https://doi.org/10.1523/JNEUROSCI.1425-11.2011

Toyomura, A., 2003. Saccadic undershoot can be explained as a trade-off between accuracy and flight-time.

Turi, M., Burr, D., 2012. Spatiotopic perceptual maps in humans: evidence from motion adaptation. Proceedings of the Royal Society B: Biological Sciences 279, 3091–3097. https://doi.org/10.1098/rspb.2012.0637

Umeno, M.M., Goldberg, M.E., 1997. Spatial Processing in the Monkey Frontal Eye Field. I. Predictive Visual Responses. Journal of Neurophysiology 78, 1373–1383.

Ungerleider, L.G., Desimone, R., 1986. Cortical connections of visual area MT in the macaque. Journal of Comparative Neurology 248, 190–222. https://doi.org/10.1002/cne.902480204

Vercher, J.-L., Gauthier, G.M., 1988. Cerebellar involvement in the coordination control of the oculo-manual tracking system: effects of cerebellar dentate nucleus lesion. Exp Brain Res 73, 155–166. https://doi.org/10.1007/BF00279669

Vercher, J.L., Gauthier, G.M., Guedon, O., Blouin, J., Cole, J., Lamarre, Y., 1996. Self-moved target eye tracking in control and deafferented subjects: roles of arm motor command and proprioception in arm-eye coordination. Journal of Neurophysiology 76, 1133–1144. https://doi.org/10.1152/jn.1996.76.2.1133

Verstraten, F.A.J., Cavanagh, P., Labianca, A.T., 2000. Limits of attentive tracking reveal temporal properties of attention. Vision Research 40, 3651–3664. https://doi.org/10.1016/S0042-6989(00)00213-3

Vogel, E.K., Woodman, G.F., Luck, S.J., 2001. Storage of features, conjunctions, and objects in visual working memory. Journal of Experimental Psychology: Human Perception and Performance 27, 92–114. https://doi.org/10.1037/0096-1523.27.1.92

von der Heydt, R., Zhou, H., Friedman, H.S., 2000. Representation of stereoscopic edges in monkey visual cortex. Vision Research 40, 1955–1967. https://doi.org/10.1016/S0042-6989(00)00044-4

von Holst, E., Mittelstaedt, H., 1950. Das Reafferenzprinzip: Wechselwirkungen zwischen Zentralnervensystem und Peripherie. Naturwissenschaften 37, 464–476. https://doi.org/10.1007/BF00622503

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, 1988–2003. https://doi.org/10.1152/jn.1995.73.5.1988

Wallace, D.J., Fitzpatrick, D., Kerr, J.N.D., 2016. Primate Thalamus: More Than Meets an Eye. Current Biology 26, R60–R61. https://doi.org/10.1016/j.cub.2015.11.025

Waszak, F., Cardoso-Leite, P., Hughes, G., 2012. Action effect anticipation: Neurophysiological basis and functional consequences. Neuroscience & Biobehavioral Reviews 36, 943–959. https://doi.org/10.1016/j.neubiorev.2011.11.004

Wegner, D.M., Wheatley, T., 1999. Apparent mental causation: Sources of the experience of will. American Psychologist 54, 480–492. https://doi.org/10.1037/0003-066X.54.7.480

Wenderoth, P., Wiese, M., 2008. Retinotopic encoding of the direction aftereffect. Vision Research 48, 1949–1954. https://doi.org/10.1016/j.visres.2008.06.013

Wexler, M., Collins, T., 2014. Orthogonal steps relieve saccadic suppression. Journal of Vision 14, 13–13. https://doi.org/10.1167/14.2.13

Wurtz, R.H., 2008. Neuronal mechanisms of visual stability. Vision Research 48, 2070–2089. https://doi.org/10.1016/j.visres.2008.03.021

Wyart, V., Nobre, A.C., Summerfield, C., 2012. Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. PNAS 109, 3593–3598. https://doi.org/10.1073/pnas.1120118109

Zeater, N., Cheong, S.K., Solomon, S.G., Dreher, B., Martin, P.R., 2015. Binocular Visual Responses in the Primate Lateral Geniculate Nucleus. Current Biology 25, 3190–3195. https://doi.org/10.1016/j.cub.2015.10.033

Zeki, S., 1980. The representation of colours in the cerebral cortex. Nature 284, 412–418. https://doi.org/10.1038/284412a0

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, 3066–3076. https://doi.org/10.1152/jn.00156.2014

APPENDIX

ADDITIONAL FIGURES

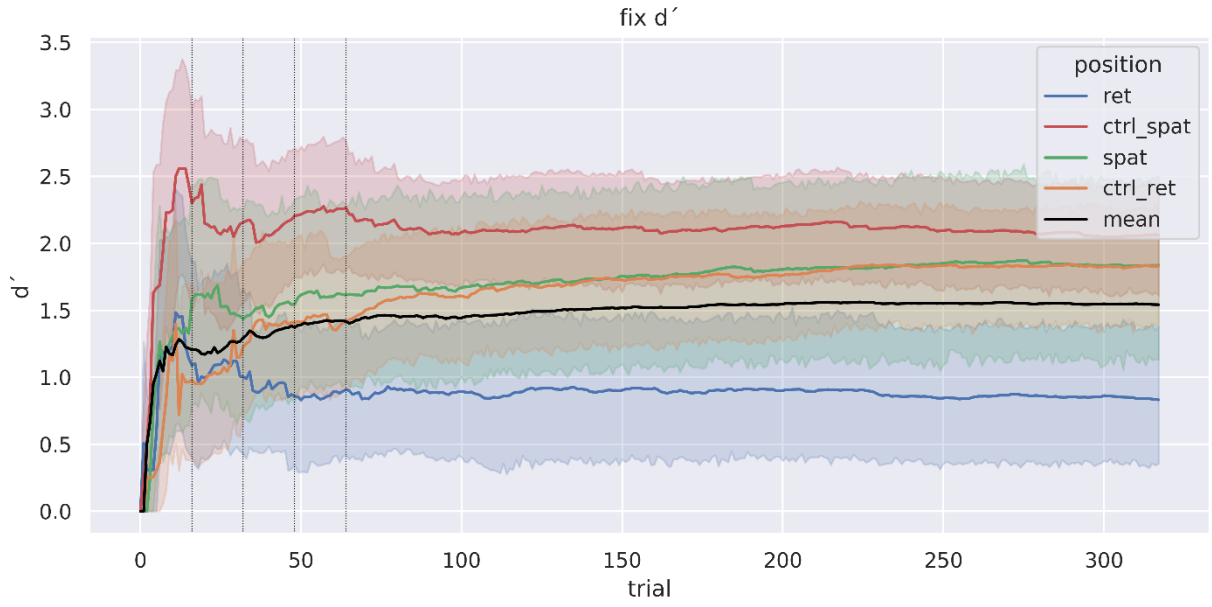

Figure 31 Performance by position across fixation trials.

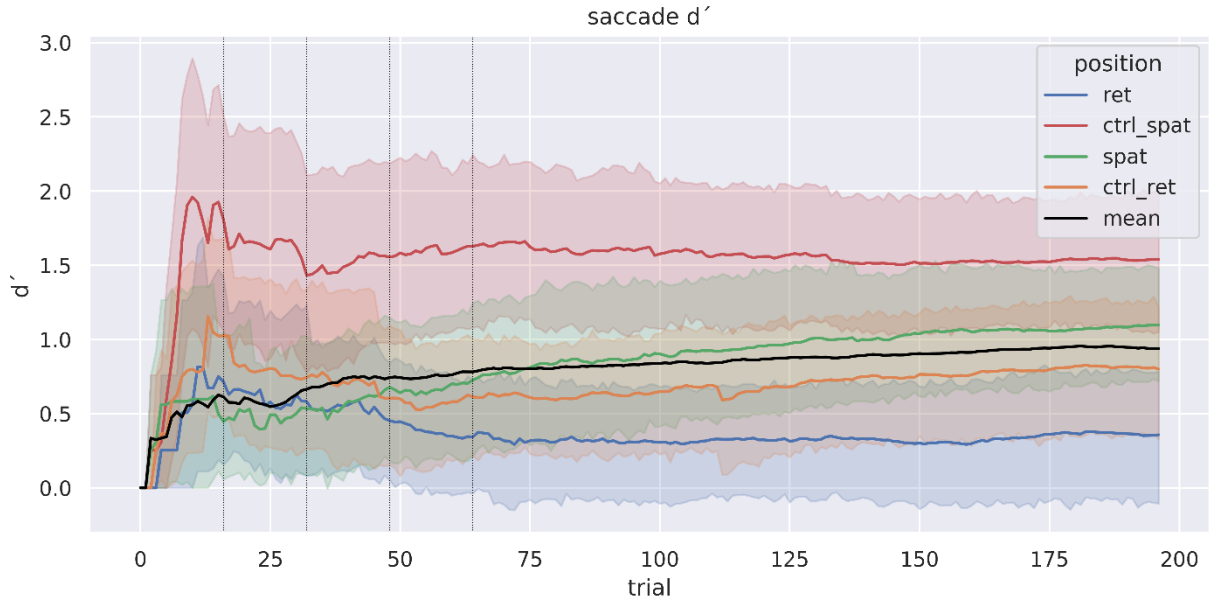

Figure 32 Performance by position across saccade trials.

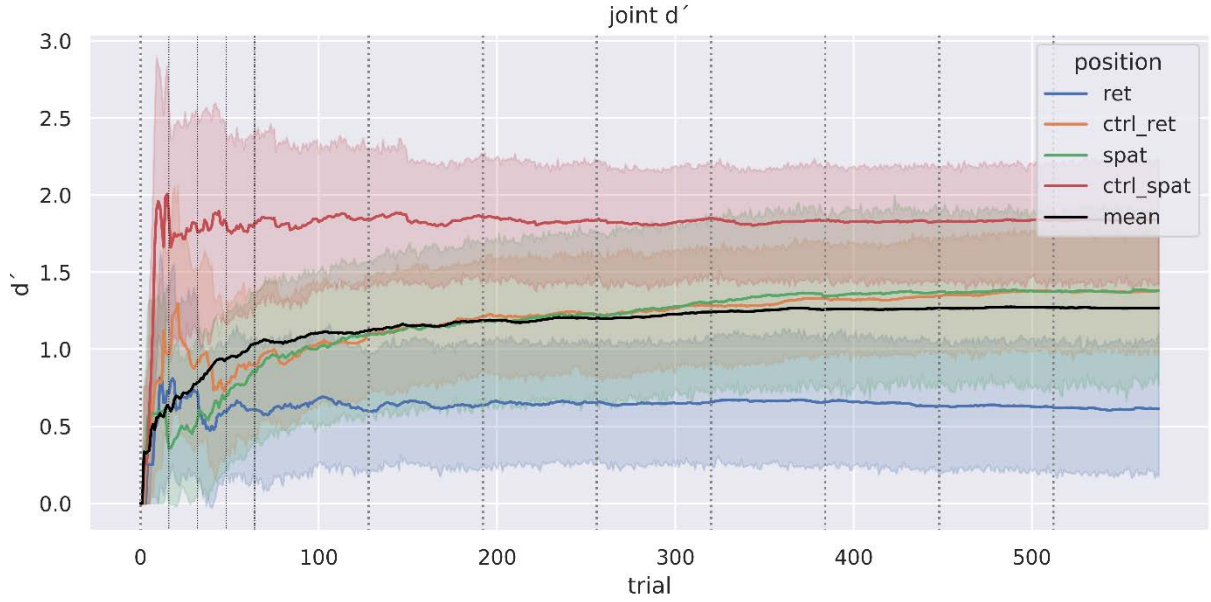

Figure 33 Performance by position across the experiment.

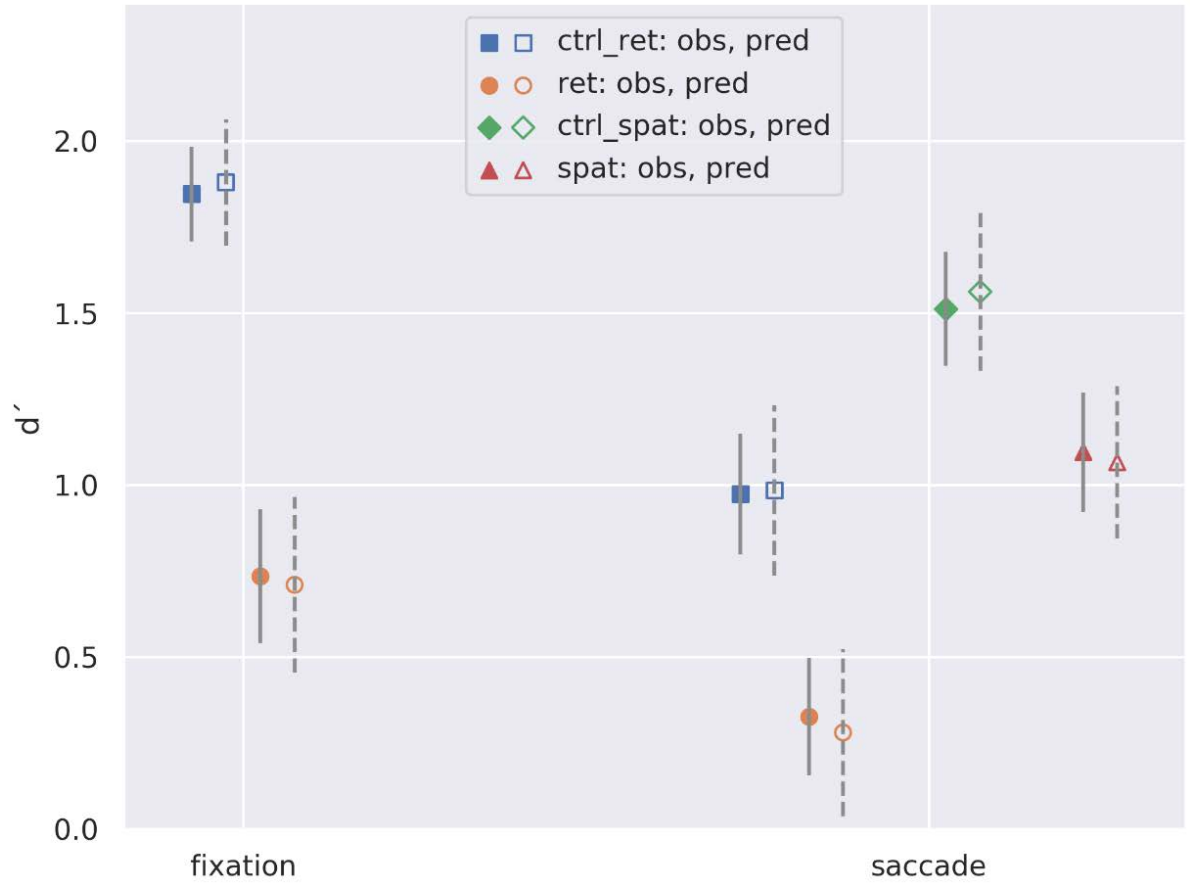

Figure 34 Plot that compares actual d' results and estimated results given by the LMEM. On the left, fixation trials display basically the same values. On the right, saccade trials show a slight decrease of the spatiotopic test.

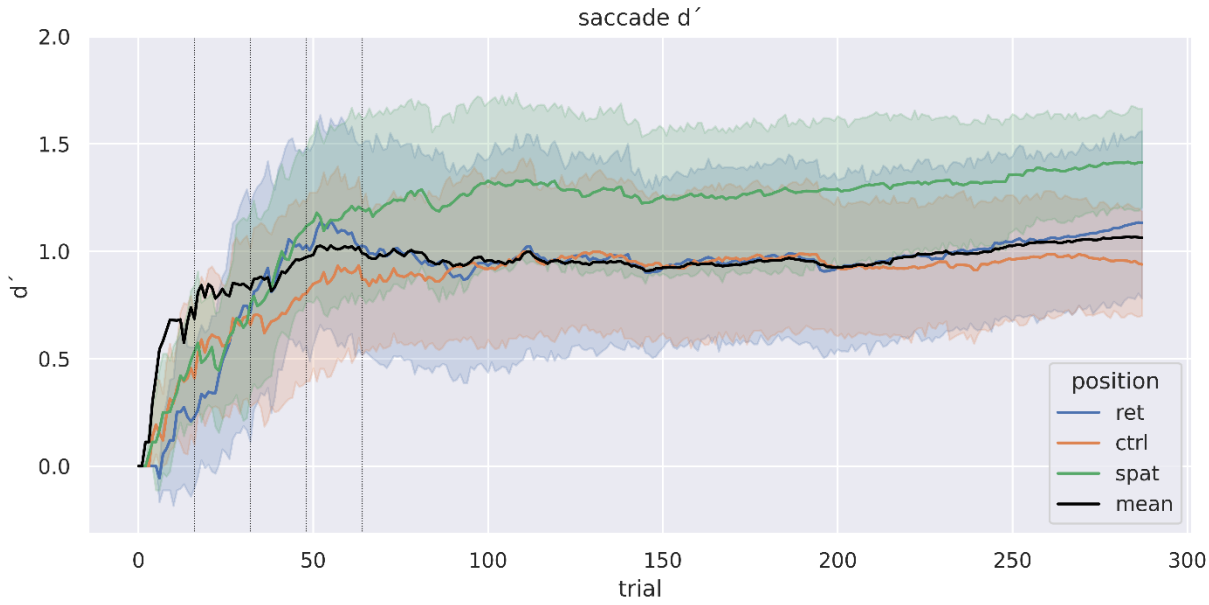

Figure 35 Performance by position across saccade trials.

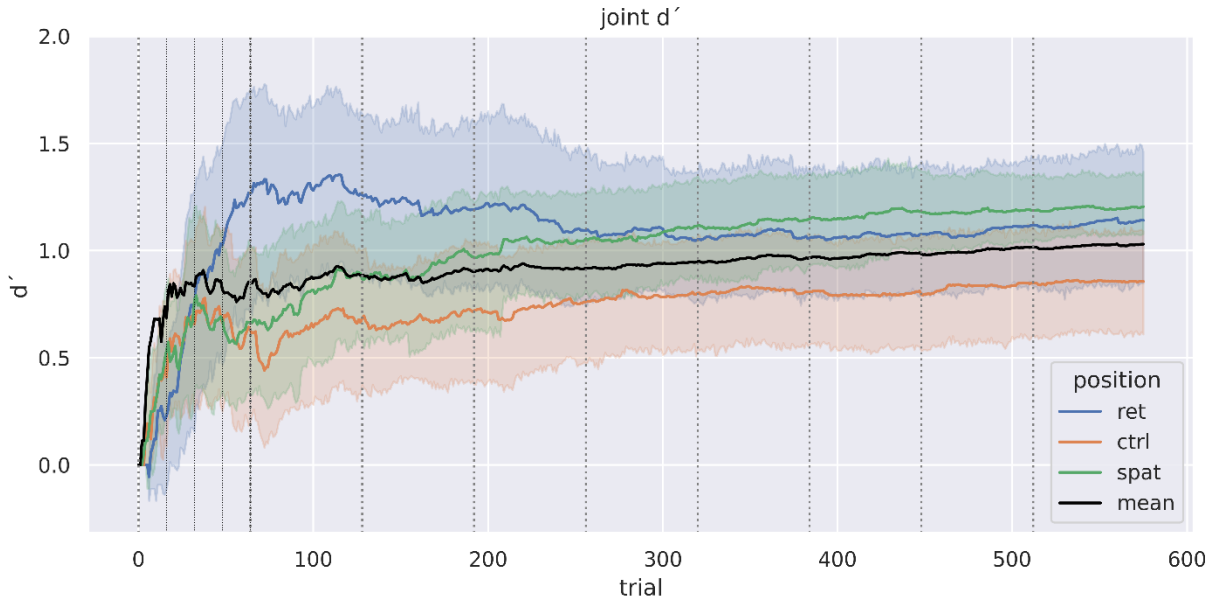

Figure 36 Performance by position across the experiment. The first thin vertical lines denote each test sub-block, wider dashed lines are markers of previous acquisition phases.

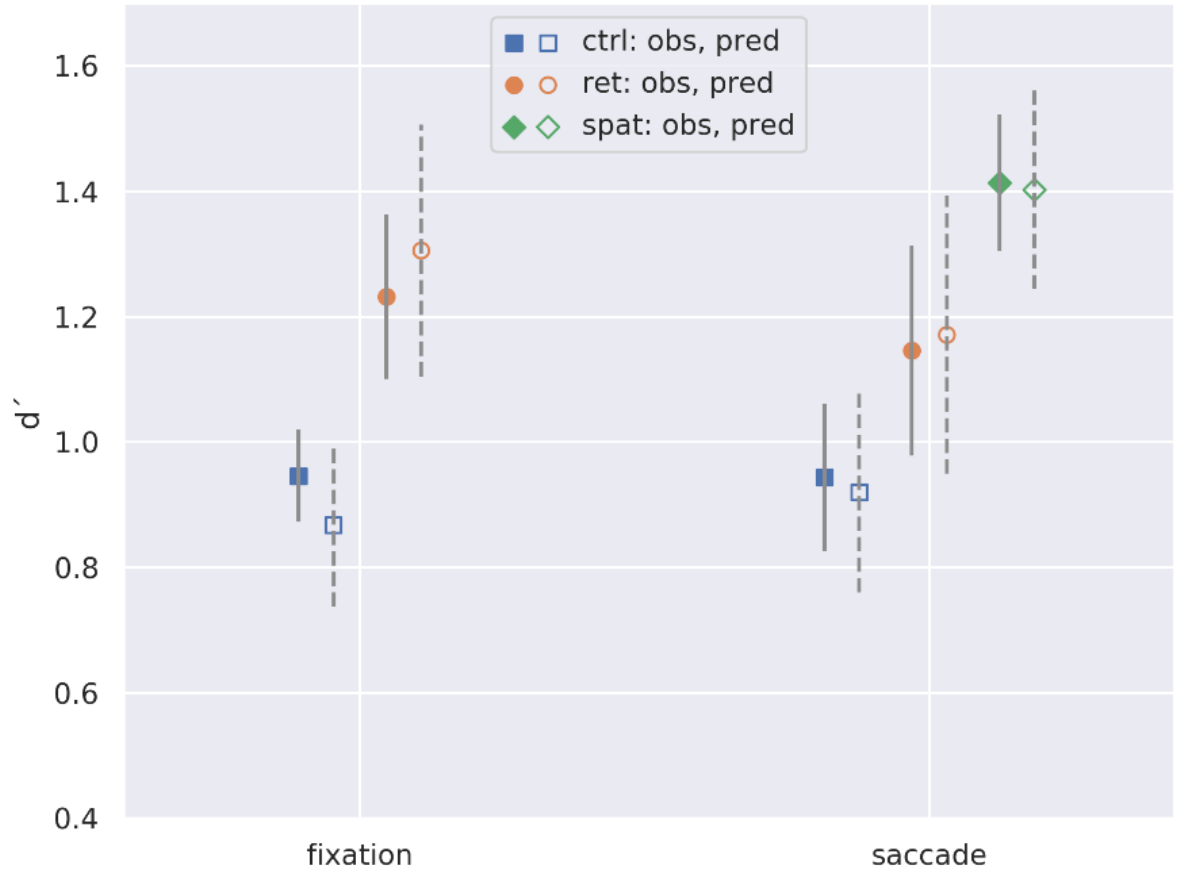

Figure 37 Comparison of observed subject sensitivity and sensitivity as predicted by the LMEM. Fixation (left) and Saccade (right) trials. Filled shapes are observed averaged values, unfilled shapes are values as predicted by the mixed-model.

Additional Mixed Models

Coherent Motion Detection Experiment

Random Effects:					
Groups	Name	Variance	Std.Dev.	Corr	
subject	(Intercept)	0.4755	0.6895		
	ret	0.9835	0.9906	-0.37	
	congruent	0.0222	0.4900	0.37	0.35
Residual		0.0309	0.7570		

LMEM 13 Random effects for fixation trials in the motion detection experiment.

Random Effects:							
Groups	Name	Variance	Std.Dev.	Corr			
subject	(Intercept)	0.6246	0.7903				
	ctrl_spat	0.7842	0.8476	-0.53			
	ret	0.6407	0.7836	-0.23	-0.30		
	spat	0.4503	0.2043	-0.38	0.84	-0.05	
	congruent	0.0722	0.3200	0.80	0.29	0.49	0.09
Residual		0.0570	0.2274				

LMEM 14 Random-effects for the expanded fixation trials in the motion detection experiment.

Random Effects:							
Groups	Name	Variance	Std.Dev.	Corr			
subject	(Intercept)	0.8303	0.9200				
	ctrl_spat	0.7243	0.8500	-0.54			
	ret	0.4979	0.7056	-0.40	-0.24		
	spat	0.9931	0.9965	-0.64	0.70	0.00	
	congruent	0.0977	0.4060	-0.27	0.37	0.02	0.87
Residual		0.0845	0.2907				

LMEM 15 Random effects for saccade trials in the motion detection experiment.

Random Effects:								
Groups	Name	Variance	Std.Dev.	Corr				
subject	(Intercept)	0.4783	0.6960					
	ctrl_spat	0.5223	0.7227	-0.80				
	ret	0.5008	0.7077	-0.24	-0.60			
	spat	0.6393	0.7996	0.00	0.59	-0.57		
	congruent	0.0086	0.0926	0.60	0.20	-0.54	0.67	
	saccade	0.2768	0.4666	-0.24	-0.33	0.27	-0.73	-0.36
Residual		0.2846	0.3584					

LMEM 16 Random effects of the joint mixed model for the motion detection experiment. Fixation positions are not expanded.

Random Effects:								
Groups	Name	Variance	Std.Dev.	Corr				
subject	(Intercept)	0.6709	0.8908					
	ctrl_spat	0.6876	0.8292	-0.55				
	ret	0.4059	0.6408	-0.20	-0.27			
	spat	0.6067	0.0773	-0.39	0.77	-0.20		
	congruent	0.0088	0.0939	0.50	0.62	-0.04	0.72	
	saccade	0.9907	0.4467	-0.22	-0.06	-0.09	-0.40	-0.47
Residual		0.0485	0.3238					

LMEM 17 Random effects of the joint mixed model for the motion detection experiment. Fixation positions are expanded.

Motion Discrimination Experiment

Random Effects:					
Groups	Name	Variance	Std.Dev.	Corr	
subject	(Intercept)	0.4570	0.3870		
	ret	0.0741	0.2722	0.00	
	congruent	0.0700	0.3890	-0.82	-0.86
Residual		0.0556	0.2358		

LMEM 18 Random effects for fixation trials in the motion discrimination experiment.

Random Effects:						
Groups	Name	Variance	Std.Dev.	Corr		
subject	(Intercept)	0.9722	0.4440			
	ret	0.5043	0.3879	0.39		
	spat	0.0576	0.2400	-0.29	-0.79	
	congruent	0.2099	0.4494	-0.34	-0.05	-0.53
Residual		0.0899	0.3300			

LMEM 19 Random effects for saccade trials in the motion discrimination experiment.

Random Effects:							
Groups	Name	Variance	Std.Dev.	Corr			
subject	(Intercept)	0.0965	0.3300				
	ret	0.0884	0.2973	0.70			
	spat	0.0735	0.2700	-0.28	-0.85		
	congruent	0.0000	0.3370	-0.65	-0.30	-0.90	
	saccade	0.0472	0.2730	0.40	0.60	-0.48	0.00
Residual		0.9800	0.3346				

LMEM 20 Random effects for "joint" fixation and saccade trials. Fixation trials still have only two positions (control and test).

Random Effects:						
Groups	Name	Variance	Std.Dev.	Corr		
subject	(Intercept)	0.2240	0.4705			
	ret	0.0247	0.5700	0.78		
	spat	0.2388	0.3520	-0.66	0.00	
	congruent	0.0484	0.2990	-0.73	-0.99	0.06
Residual		0.0962	0.3020			

LMEM 21 Random effects for the expanded fixation trials of the discrimination experiment.

Random Effects:							
Groups	Name	Variance	Std.Dev.	Corr			
subject	(Intercept)	0.5960	0.3995				
	ret	0.0547	0.2338	0.82			
	spat	0.0541	0.2325	-0.64	-0.67		
	congruent	0.0754	0.2740	-0.50	-0.42	-0.25	
	saccade	0.0256	0.5990	0.25	0.20	0.20	-0.60
Residual		0.3595	0.3687				

LMEM 22 Random effects for the joint mixed model in the discrimination experiment. Fixation trials are expanded.

Fixation	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	566.3115	48.4897	11.679	0	***
ret	5.2188	21.788	0.2395	0.378	ns
congruent	-3.4712	13.7182	-0.253	0.3766	ns
ret × congruent	16.2625	22.665	0.7175	0.2964	ns
Saccade	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	634.7243	51.2934	12.3744	0	***
ret	17.1662	23.3699	0.7345	0.2968	ns
spat	-27.6929	24.6899	-1.1216	0.2068	ns
congruent	-32.4618	20.6143	-1.5747	0.1154	ns
ret × congruent	-2.3076	27.6411	-0.0835	0.3917	ns
spat × congruent	21.4396	27.635	0.7758	0.2874	ns

LMEM 23 RT mixed-model fixed effects for fixation (top) and saccade (bottom) trials in the discrimination experiment.

Fixation	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	571.5787	49.1132	11.638	0	***
ret	-0.0462	27.4916	-0.0017	0.3931	ns
spat	-10.4526	22.2171	-0.4705	0.3499	ns
congruent	-3.7106	19.3357	-0.1919	0.3855	ns
ret × congruent	16.5009	26.1183	0.6318	0.319	ns
spat × congruent	0.3624	26.1354	0.0139	0.3931	ns
Saccade	**estimate**	**stderr**	**t-value**	**p-value**	**sign**
(Intercept)	634.7243	51.2934	12.3744	0	***
ret	17.1662	23.3699	0.7345	0.2968	ns
spat	-27.6929	24.6899	-1.1216	0.2068	ns
congruent	-32.4618	20.6143	-1.5747	0.1154	ns
ret × congruent	-2.3076	27.6411	-0.0835	0.3917	ns
spat × congruent	21.4396	27.635	0.7758	0.2874	ns

LMEM 24 RT fixed effects of mixed model for expanded fixation positions in the motion discrimination experiment.

Random Effects:					
Groups	Name	Variance	Std.Dev.	Corr	
subject	(Intercept)	2783.8	64.88		
	ret	265.7	5.4	-0.50	
	congruent	200.4	4.6	-0.40	0.67
Residual		9823.8	33.39		

LMEM 25 Random-effects for fixation trials in the motion discrimination experiment.

Random Effects:						
Groups	Name	Variance	Std.Dev.	Corr		
subject	(Intercept)	29287	7.3			
	ret	979.7	44.49	-0.60		
	spat	2745	52.39	-0.02	0.96	
	congruent	509.2	22.57	-0.40	0.73	0.82
Residual		949.2	330.89			

LMEM 26 Random-effects for saccade trials in the motion discrimination experiment.

Random Effects:						
Groups	Name	Variance	Std.Dev.	Corr		
subject	(Intercept)	26897	64			
	ret	4980	70.57	-0.45		
	spat	85.8	42.6	0.00	0.88	
	congruent	389.8	9.74	-0.43	0.69	0.69
Residual		97802.3	32.73			

LMEM 27 Random-effects for the expanded fixation trials in the motion discrimination experiment.

Random Effects:							
Groups	Name	Variance	Std.Dev.	Corr			
subject	(Intercept)	26527	62.87				
	ret	254.4	50.4	-0.37			
	spat	3733.9	6	-0.20	0.96		
	saccade	635	40.43	0.25	0.00	-0.20	
	congruent	270.8	6.46	-0.45	0.87	0.85	-0.57
Residual		3792.7	322.7				

LMEM 28 Random-effects for the joint mixed-model with fixation controls averaged.

Random Effects:							
Groups	Name	Variance	Std.Dev.	Corr			
subject	(Intercept)	26586.7	63.05				
	ret	379.9	6.58	-0.39			
	spat	2742	52.36	0.00	0.96		
	saccade	867.5	43.2	0.40	0.26	0.32	
	congruent	233.3	5.27	-0.58	0.90	0.80	-0.40
Residual		3675.2	32.99				

LMEM 29 Random-effects for the joint mixed-model with expanded fixation positions.

Random Effects:						
Groups	Name	Variance	Std.Dev.	Corr		
subject	(Intercept)	26663.6	63.29			
	test	2889.7	53.76	-0.23		
	ref_frame	844.4	29.06	0.32	0.36	
	congruent	8.5	3.47	-0.58	0.87	-0.09
Residual		4336	323			

LMEM 30 Random-effects for the relabelled mixed-model.

POSITION COMPARISON TABLES

Coherent Motion Detection Experiment

dprime		**ctrl_ret**	**ctrl_spat**	**ret**	**spat**
diff	**ctrl_ret**	0	-0.5775	0.7027	-0.0823
	ctrl_spat	0.5775	0	1.2802	0.4951
	ret	-0.7027	-1.2802	0	-0.7851
	spat	0.0823	-0.4951	0.7851	0
t-value	**ctrl_ret**	3.986	-2.3662	3.3325	0.2958
	ctrl_spat	-2.3662	6.7851	-3.8218	-2.3087
	ret	3.3325	-3.8218	1.1623	2.357
	spat	0.2958	-2.3087	2.357	4.811
p-value	**ctrl_ret**	0.0007	0.0287	0.0034	0.3772
	ctrl_spat	0.0287	0	0.001	0.0322
	ret	0.0034	0.001	0.1991	0.0292
	spat	0.3772	0.0322	0.0292	0.0001

Table 1 Difference in d' estimates (sub-table 1), corresponding comparison t-values (sub-table 2) and p-values (sub-table 3) between the different positions for saccade trials. These results refer to saccade trials and are referenced from the MIXED-MODELS {p. 58} section of the COHERENT MOTION DETECTION EXPERIMENT.

dprime		**ctrl_ret**	**ctrl_spat**	**ret**	**spat**
diff	**ctrl_ret**	0	-0.0872	1.2225	0.0341
	ctrl_spat	0.0872	0	1.3097	0.1212
	ret	-1.2225	-1.3097	0	-1.1885
	spat	-0.0341	-0.1212	1.1885	0
t-value	**ctrl_ret**	9.104	-0.3724	5.5898	-0.1058
	ctrl_spat	-0.3724	9.4824	-3.9973	-0.6288
	ret	5.5898	-3.9973	2.7388	3.056
	spat	-0.1058	-0.6288	3.056	6.1994
p-value	**ctrl_ret**	0	0.3674	0	0.3925
	ctrl_spat	0.3674	0	0.0007	0.3219
	ret	0	0.0007	0.0132	0.0065
	spat	0.3925	0.3219	0.0065	0

Table 2 Difference in d' estimates (sub-table 1), corresponding comparison t-values (sub-table 2) and p-values (sub-table 3) for the comparison between the different positions for the expanded fixation trials. These results are referenced from the EXPANDED MIXED-MODEL {p. 60} section of the COHERENT MOTION DETECTION EXPERIMENT.

dprime		**ctrl_ret**	**ctrl_spat**	**ret**	**spat**
diff	**ctrl_ret**	0	-0.0872	1.2225	0.0341
	ctrl_spat	0.0872	0	1.3097	0.1212
	ret	-1.2225	-1.3097	0	-1.1885
	spat	-0.0341	-0.1212	1.1885	0
t-value	**ctrl_ret**	8.5002	-0.3564	6.0118	-0.1127
	ctrl_spat	-0.3564	9.2576	-4.0161	-0.5705
	ret	6.0118	-4.0161	2.7819	3.2966
	spat	-0.1127	-0.5705	3.2966	6.5998
p-value	**ctrl_ret**	0	0.3714	0	0.3938
	ctrl_spat	0.3714	0	0.0004	0.3356
	ret	0	0.0004	0.0107	0.0029
	spat	0.3938	0.3356	0.0029	0

Table 3 Difference in d' estimates (sub-table 1), corresponding comparison t-values (sub-table 2) and p-values (sub-table 3) for the comparison between different positions for the joint trials (fixation trials are expanded). These results refer to joint mixed-model and are referenced from the JOINT MIXED-MODEL {p. 62} of section of the COHERENT MOTION DETECTION EXPERIMENT.

Motion Discrimination Experiment

dprime		ctrl	ret	spat
diff	ctrl	0	-0.2511	-0.4822
	ret	0.2511	0	-0.2311
	spat	0.4822	0.2311	0
t-value	ctrl	5.7599	1.4329	3.1818
	ret	1.4329	5.2849	-1.057
	spat	3.1818	-1.057	8.8407
p-value	ctrl	0	0.1409	0.0059
	ret	0.1409	0.0001	0.2217
	spat	0.0059	0.2217	0

Table 4 Difference in estimates (sub-table 1), corresponding comparison t-values (sub-table 2) and p-values (sub-table 3) for the comparison between different positions for saccade trials. These results are referenced from the MIXED-MODELS section {p. 74} of the MOTION DISCRIMINATION EXPERIMENT.

dprime		ctrl	ret	spat
diff	ctrl	0	-0.652	-0.4247
	ret	0.652	0	0.2273
	spat	0.4247	-0.2273	0
t-value	ctrl	4.0176	4.847	2.616
	ret	4.847	6.6825	1.3257
	spat	2.616	1.3257	7.9205
p-value	ctrl	0.001	0.0002	0.0187
	ret	0.0002	0	0.1622
	spat	0.0187	0.1622	0

Table 5 Difference in estimates (sub-table 1), corresponding comparison t-values (sub-table 2) and p-values (sub-table 3) for the comparison between different expanded positions for the fixation trials. These results are referenced from the EXPANDED MIXED-MODEL section {p. 76} of the MOTION DISCRIMINATION EXPERIMENT.

dprime		ctrl	ret	spat
diff	ctrl	0	-0.652	-0.4247
	ret	0.652	0	0.2273
	spat	0.4247	-0.2273	0
t-value	ctrl	4.1649	3.9525	2.5771
	ret	3.9525	6.3796	1.1692
	spat	2.5771	1.1692	7.7874
p-value	ctrl	0.0004	0.0006	0.0181
	ret	0.0006	0	0.1981
	spat	0.0181	0.1981	0

Table 6 Difference in estimates (sub-table 1), corresponding comparison t-values (sub-table 2) and p-values (sub-table 3) for the comparison between the different positions for joint fixation (expanded) and saccade trials. These results are referenced from JOINT EXPANDED MIXED-MODEL {p. 78} of the MOTION DISCRIMINATION EXPERIMENT.

TABLE OF FIGURES

TABLE OF LINEAR MIXED-EFFECTS MODELS

TABLE OF MIXED MODELS