

<https://helda.helsinki.fi>

Eye movements in the wild : Oculomotor control, gaze behavior & frames of reference

Lappi, Otto

Elsevier Scientific Publ. Co

2016-10

Lappi , O 2016 , ' Eye movements in the wild : Oculomotor control, gaze behavior & frames
of reference ' , Neuroscience & Biobehavioral Reviews , vol. 69 , pp. 49-68 . <https://doi.org/10.1016/j.neubiorev.2016>

<http://hdl.handle.net/10138/224093>

[10.1016/j.neubiorev.2016.06.006](https://doi.org/10.1016/j.neubiorev.2016.06.006)

cc_by_nc_nd

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.


Please cite the original version.

Author's personal copy

Lappi, O. (2016). Eye movements in the wild: oculomotor events, gaze behavior and frames of reference. *Neuroscience & Biobehavioral Reviews*, 69, 49–68.
doi: 10.1016/j.neubiorev.2016.06.006

Academic references should be made to the copyedited published version:

Neuroscience and Biobehavioral Reviews 69 (2016) 49–68




ELSEVIER

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev



Review article

Eye movements in the wild: Oculomotor control, gaze behavior & frames of reference

Otto Lappi (Ph.D)

Cognitive Science, Institute of Behavioural Sciences, PO BOX 9, 00014 University of Helsinki, Finland

ARTICLE INFO

Article history:
Received 16 August 2015
Received in revised form 14 May 2016
Accepted 8 June 2016
Available online 25 July 2016

Keywords:
Eye movements
Oculomotor events
Gaze behavior
Naturalistic tasks
Sensorimotor transformations
Spatial representation
Frames of reference

ABSTRACT

Understanding the brain's capacity to encode complex visual information from a scene and to transform it into a coherent perception of 3D space and into well-coordinated motor commands are among the outstanding questions in the study of integrative brain function. Eye movement methodologies have allowed us to begin addressing these questions in increasingly naturalistic tasks, where eye and body movements are ubiquitous and, therefore, the applicability of most traditional neuroscience methods restricted. This review explores foundational issues in (1) how oculomotor and motor control in lab experiments extrapolates into more complex settings and (2) how real-world gaze behavior in turn decomposes into more elementary eye movement patterns. We review the received typology of oculomotor patterns in laboratory tasks, and how they map onto naturalistic gaze behavior (or not). We discuss the multiple coordinate systems needed to represent visual gaze strategies, how the choice of reference frame affects the description of eye movements, and the related but conceptually distinct issue of coordinate transformations between internal representations within the brain.

© 2016 Elsevier Ltd. All rights reserved.

Contents

1. Introduction.....	49
2. Eye movement control and received OE classification in laboratory tasks.....	51
2.1. Rapid eye movements (REM): the gaze shifting OE.....	52
2.2. Fixational eye movements (FEM): how to define "a fixation" in the laboratory?.....	54
2.3. Slow eye movements (SEM): multiple systems for image stabilization.....	55
2.4. Oculomotor event types as indices of "modular" control?.....	56
3. Eye movement behavior in the wild.....	57
3.1. Oculomotor events vs. gaze behavior: how to define "a fixation" in natural behavior?.....	58
3.2. Frames of reference.....	59
3.3. Internal representation of space: multisensory integration & sensorimotor coordinate transformations.....	61
4. Active gaze & feedback control as unifying themes.....	62
4.1. Active gaze (and optic flow).....	62
4.2. Feedback control (and internal representation revisited).....	63
5. Conclusions.....	65
References.....	65

1. Introduction

Understanding eye movements is essential for understanding all aspects of visual brain function, including perception, attention, memory, dynamic real-world decision making and motor control. For unlike untutored introspection would suggest, visual perception is not based on a continuous high-definition widescreen stream of input. Instead, the spatial world is sampled by eye movements that organize the brain's input into local discrete snapshots. It is from these that the brain must piece together a coherent and stable visual percept. What is more, in natural tasks the eye movements are always embedded in head movement and locomotor patterns. For the brain, this creates a complex control problem of

E-mail address: otto.lappi@helsinki.fi

<http://dx.doi.org/10.1016/j.neubiorev.2016.06.006>
0149-7634/© 2016 Elsevier Ltd. All rights reserved.

Abstract

Understanding the brain's capacity to encode complex visual information from a scene and to transform it into a coherent perception of 3D space and into well-coordinated motor commands are among the outstanding questions in the study of integrative brain function. Eye movement methodologies have allowed us to begin addressing these questions in increasingly naturalistic tasks, where eye and body movements are ubiquitous and, therefore, the applicability of most traditional neuroscience methods restricted. This review explores foundational issues in (1) how oculomotor and motor control in lab experiments extrapolates into more complex settings and (2) how real-world gaze behavior in turn decomposes into more elementary eye movement patterns. We review the received typology of oculomotor patterns in laboratory tasks, and how they map onto naturalistic gaze behavior (or not). We discuss the multiple coordinate systems needed to represent visual gaze strategies, how the choice of reference frame affects the description of eye movements, and the related but conceptually distinct issue of coordinate transformations between internal representations within the brain.

Keywords: *eye movements, oculomotor events, gaze behavior, naturalistic tasks, sensorimotor transformations, spatial representation, frames of reference*

1. Introduction

Understanding eye movements is essential for understanding all aspects of visual brain function, including perception, attention, memory, dynamic real-world decision making and motor control. For unlike untutored introspection would suggest, visual perception is not based on a continuous high-definition widescreen stream of input. Instead, the spatial world is sampled by eye movements that organize the brain's input into local discrete snapshots. It is from these that the brain must piece together a coherent and stable visual percept. What is more, in natural tasks the eye movements are always embedded in head movement and locomotor patterns. For the brain, this creates a complex control problem of coordinating multiple sensory systems and multiple effectors in a coherent and goal-driven way. For the visual scientist, all this creates methodological and conceptual challenges - but also an opportunity to analyze visual function in its natural ecological context.

Our understanding of eye movement patterns and their underlying neurobiology has advanced considerably in the last 20 years; this is partly due to several methodological advances. First, eye tracking and single-unit measurement methods have been complemented by the application of brain imaging methods and eye movements have proved to be a useful tool for neuroscientist interested in attention, memory and motor control in the human and in the monkey. Here, the eye movements per se are not the primary interest, but a means to access integrative visual and cognitive function. The relatively few degrees of freedom in the oculomotor plant and the phylogenetically conserved control circuitry have made the eye movement system an ideal model for studying the programming and execution of discrete motor action sequences. Second, the past two decades have seen a flurry of development in the measurement equipment and algorithms useful for eye tracking in realistically complex and even fully naturalistic task environments. Inexpensive and easy-to-use video-based eye tracking equipment in particular has become widely available. With the spread of these emerging methodologies, there has been a proliferation of studies investigating human visual behavior in its natural ecological context "in the wild" (we shall use this term to refer to complex sequential behavior in naturalistic tasks with unrestrained physical - or virtual - locomotion and free head-movement, typically in a richly structured visual environment). This line of research has revealed recurring qualitative patterns of gaze behavior in many real-world tasks (see e.g. Land, 2006; Regan & Gray, 2000; Tatler et al., 2011).

Yet much of what we think is actually happening in the visuomotor system during gaze-control is extrapolated from experiments that isolate specific oculomotor event types, and then characterize the neural circuits involved. What is more, real world gaze behavior is typically classified into elementary patterns corresponding to this "received" typology. But just how well can we expect the traditional classification of oculomotor patterns in laboratory tasks to fit the need for describing gaze behavior in the wild? How do the need to consider multiple 3D coordinate systems in locomotion and the physical relativity of movement to the choice of reference frame impact on the description of eye movement behavior, and on the hypotheses concerning underlying mechanisms?

There are many excellent recent reviews on the emerging literature on naturalistic gaze behavior (Hayhoe and Ballard, 2005; Kowler, 2011; Land, 2006; Regan and Gray, 1999; Steinman, Kowler & Collewyn, 1990; Tatler et al., 2011) - and of course many decades' worth of textbooks and reviews on the details of oculomotor control in traditional lab tasks. Nevertheless, it seems the literature could do with a systematic integrative review that would cover both the traditional lab-based approach to characterizing of oculomotor behavior, *and* the particular problems raised by studying eye movement behavior in fully naturalistic locomotor contexts. Because the two approaches still tend to get treated separately, there are conceptual and methodological gaps between the domains; without a clear road map researchers taking eye tracking equipment and experimental procedures outside the lab are having to tackle with the same foundational issues (and come up with solutions each on their own).

How do the received eye movement research in the laboratory and the emerging eye movement research in the wild relate to one another - theoretically, empirically and methodologically? When such foundational issues as these are tackled, the arguments must draw on a relatively broad array of

literature, and the overall plot can be challenging to follow. Figure 1 gives the reader a “bird’s eye view” of the main themes in this review and where the issues sit in the relevant literature. The breadth of scope means that not all the topics can be covered in full technical detail. So instead of focusing, say, on the physiological and computational models of saccade generating circuits, or on mathematical modeling of 3D geometry of the eye and head, or on recent algorithmic solutions to the compensation of head movement in outdoors eye tracking, this paper attempts to outline the major conceptual issues in the relatively less charted middle ground.

--- INSERT FIGURE 1 HERE ---

We begin with a review of the received classification of major eye movement patterns in laboratory tasks: fixation (and fixational eye movements), saccades, smooth pursuit, optokinetic nystagmus and the vestibulo-ocular response. We then consider the foundational question of how they might contribute to the complex streams of motor activity typical of real-world behavior. In particular, we evaluate the idea of classical *eye movement types* as suite of oculomotor building blocks combined to control *gaze* (location and orientation of the line of sight in the world). We look at how the brain’s control task needs to be analyzed in terms of multiple sensory systems, multiple effectors and multiple 3D frames of reference, and how these undermine attempts to directly relate classical oculomotor events to naturalistic gaze behavior. Finally, we look at active gaze and feedback control as potentially unifying ideas that might help to bridge the existing gap between laboratory-based and naturalistic eye movement research. Please note that the many abbreviations used are collected in Table 3 at the end of the paper for convenient reference.

2. Eye movement control and received OE classification in laboratory tasks

This section gives a brief overview of the “received” typological classification of *oculomotor events* (OE). The major physiological characteristics of eye movement patterns in laboratory tasks are fairly well-established in the eye tracking literature (Tables 1 & 2). Eye-in-head velocity profiles, typical eliciting tasks or stimuli, reaction times and the procedures for identifying and reporting them in specific experimental conditions have been codified over the years. There is also a fairly wide agreement on their *functional* contributions to visual behavior and perception.

Moreover, the underlying oculomotor circuitry and their microanatomical and computational details have in many cases been modeled in exquisite detail. Figure 2 outlines some of the major anatomical pathways (based on previous reviews by Munoz, 2004; Krauzlis, 2004; Pierrot-Deseilligny et al., 2004; Goodale, 2011). There are many excellent reviews that cover in depth the physiology and functional characteristics of circuit behavior underlying saccades (Girard & Berthoz, 2005; Ibbotson & Krekelberg, 2011; Munoz, 2004; Scudder, Kaneko & Fuchs, 2002; Sparks, 2002), microsaccades (Collewijn & Kowler, 2008; Engbert, 2006; Martinez-Conde, Macknik & Hubel, 2004; Martinez-Conde et al., 2009; Rolfs, 2009), vestibulo-ocular and optokinetic compensatory eye movements (Angelaki & Hess, 2005; Miles, 1997) and smooth pursuit movements (Barnes, 2008; Ilg, 1997; Krauzlis, 2004; Thier & Ilg, 2005).

--- INSERT FIGURE 2 HERE ---

--- INSERT TABLE 1 & TABLE 2 HERE---

The main dichotomies in oculomotor event classification are, first, between fixation and eye movements, and, second, between rapid eye movements (REM) and slow eye movements (SEM). Indeed, looking at eye-position as a function of time in a typical eye movement experiment (e.g. viewing naturalistic images, doing visual search or reading) immediately suggests a natural way to analyze signal into discrete “events”. Periods of relative stability (fixation) are interspersed with rapid jerky movements

(saccades). Thus, most eye movement results are based on identifying fixations and/or saccades and reporting event parameters, such as fixation dispersion, individual fixation durations, total fixation dwell time on a target area of interest or saccade amplitudes and velocities¹.

2.1. Rapid Eye movements (REM): the gaze shifting OE

Saccades are the most readily identified eye movements, easily seen in other people with the naked eye, or felt through the eyelid with your own fingertip. The launch of a saccade (and to a slightly lesser extent also the landing) involves a very rapid increase in eye-acceleration (jerk). This discontinuous or “jerky” movement is where saccades derive their name (*saccade* is French for a jerking or yanking motion).

Saccades are also the easiest eye movement type to classify objectively, using eye-position velocity and amplitude parameters, as the angular velocity and duration depend systematically on saccade amplitude (Becker & Fuchs, 1969; Robinson, 1964). The saccade main sequence (Bahill, Clark & Stark, 1975) refers to a nonlinear relation between peak velocity and amplitude of a saccade. This relation is well described by a power function, over orders of magnitude of saccade amplitude (0.1° to 15°) and peak eye velocity (10°/s to over 500°/s), until eye velocity begins to saturate somewhere beyond 500°/s. Because large saccades are among the fastest movements generated by the human body, reaching angular velocities of up 600–800 deg/s, saccades are referred to as rapid eye movements (REM). It is the main sequence, however, that provides the *de facto* standard physiological criteria for differentiating saccades from so-called slow eye movements (Ilg, 1997).

There are two main reasons for why the distribution and duration of fixations is considered to be of particular interest in investigating cognitive processes: (1) When there is no head or body motion (such as when the head is stabilized with a bite bar), maintaining oculomotor fixation stabilizes gaze relative to the stimulus, and thus stabilizes the retinal image. This is required for resolving high-resolution detail (commonly considered to be important for higher level perceptual processing). (2) During saccades, information flow into the visual system is reduced. This saccadic suppression is thought to be partly due to smearing of the retinal image, partly due to backward masking by the higher definition image with the initiation of fixation, and also active suppression of processing (of visual motion) in the brain seems to be involved (Ibbotson & Krekelberg, 2011; Thiele et al., 2002). Fixation is therefore considered functionally as the “window” when new visual information is available to the brain, punctuated by saccades during which relatively little information is received.

The control of vergence in natural viewing is also largely achieved with the eyes converging or diverging during saccades (except when tracking a target in depth, when it is considered part of smooth pursuit, or in response to linear forward or backward acceleration of the head, when it is considered part of the translational vestibulo-ocular reflex). In what follows, we shall, however, discuss saccades and other oculomotor events mostly as if they were conjunctive – i.e. both eyes rotating by an approximately equal amount, in a yoked manner, enabling eye movements to be described in terms of monocular or cyclopean eye-in-head rotations. This is not out of thinking that control of vergence would constitute a separate oculomotor system, but simply to keep the length of the paper manageable. How the binocular control of vergence is combined to the binocular/cyclopean eye movement patterns, how these are related to the extraction and representation of depth information and stereoscopy, and the overall role of depth information in oculomotor and locomotor control would deserve a separate review unto itself ... and anyway, the general points to be raised can be raised with reference to conjugate control.

¹ Strictly, because the percentage of total looking time in a trial spent in the area of interest (AOI) around some putative visual target – “dwell time” or “AOI catch” – does not require parsing the gaze signal into fixations and saccades and other events, it is more properly termed *glance behavior*, or *look behavior*, not fixation behavior.

2.2. Fixational eye movements (FEM): how to define “a fixation” in the laboratory?

In idealized terms, a fixation would be an oculomotor event where zero movement of the eye in its orbit occurs. This, of course, is a pure abstraction: in physiological conditions the eyes are never completely fixed. Fixation can only be physiologically defined as a period of *relative* stability between saccades (operationally defined by manual inspection, or criteria such as dispersion or velocity thresholding, minimum duration etc., when using eye tracking algorithms).

Fixational eye movements (FEM) are small-amplitude movements that require highly sensitive instruments to observe reliably, and are usually studied during sustained fixation – i.e. instructing the subject to keep looking at a small target, often with the head restrained. FEM consists of alternating *microsaccades* and *drift*. A potentially third form, physiological *tremor* will not be discussed further here because its extremely small amplitude and high frequency (40-100Hz) make it unobservable in naturalistic tasks (and unclear whether it plays a functional role in them). Some authors discuss tremor as a feature of drift motion, rather than a separate FEM (Rucci & Victor, 2015).

FEM are sometimes considered a third major class of oculomotor events, alongside rapid eye movements (REM) and slow eye movements (SEM). It is an important issue whether this is just due to experimental tradition, or whether the classification has a physiological basis – i.e. whether the function and oculomotor control of FEM are substantially different from their larger counterparts, REM and SEM (see Rucci & Victor, 2015 and Sinn & Engbert, 2016; Ahissar et al. 2016).

In naturalistic tasks, the body and the head are never really quite still. (Nor are many parts of the visual environment, especially one's own limbs). Maintaining fixation therefore requires compensatory eye-rotation. Even when the sedentary subject attempts to sit or stand completely still (without head restraint), the head and body sway slightly; and as far as preventing the point of fixation from moving is concerned, compensation is therefore not perfect. Skavenski et al. (1979) were the first to make highly accurate recordings of eye and head (gaze) rotation without rigid head support, but with subjects attempting to keep the head and gaze as stable as possible, and found that even at its most stable, gaze during natural fixation moves more than had been assumed to be the case (from experiments with bite bars or chin rests).

But actually keeping the point of fixation *completely still* cannot be considered as the goal of the oculomotor control systems responsible for “maintaining fixation”. *Perfect* retinal image stabilization would lead to image fading due to neural adaptation (Ditchburn & Ginsborg, 1952; see also the discussion on Troxler's effect in Martinez-Conde et al., 2004), and Skavenski et al. (ibid.) suggested that oculomotor compensation is actually adjusted to maintain a level of retinal image motion that is sufficiently small to prevent blurring, but at the same time large enough to prevent fading through retinal adaptation (see also Steinman, Cushman & Martins, 1982). Along these lines, Aytekin, Victor & Rucci (2014) analyzed high-quality data from Epelboim et al. (1995; other results reported from this dataset of four subjects are found in Epelboim et al., 1997; Epelboim, 1998; Malinov et al., 2000; Herst, Epelboim & Steinman, 2001), and found that fixational head-rotation (about 2°/s), head translation (10.5mm/s) and eye-rotation (3.2°/s) were all present. The magnitude of eye-rotation was much larger than observed in head-fixed conditions but, crucially, gaze accuracy was not compromised to the extent one would predict if head and eye movement were independent. Indeed, the drift in gaze was *less* than would have been predicted from either head or eye movement alone. This indicates that the movement of the eye compensated for the movement of the head and, the “imperfect” compensation actually appears to be active motor control serving vision (i.e. the “slow control” of Steinman et al., 1973; microsaccades were rare in this data). Thus, an intriguing hypothesis is that slow control is an *active gaze strategy* for temporal encoding.

Temporal structure has long been considered relevant for processing of visual motion of moving targets, but the instability of fixation mean that the proximal stimulus is in motion even for stationary targets. Ahissar & Arieli (2001; 2012) present a physiologically detailed theory where high spatial frequency information is encoded in the relative timing of ganglion cell firing, an analogous mechanism to the way fine texture is encoded in relative timing of touch receptors in *active touch* when we move our fingers over an object. Another active gaze idea is that temporal modulation of receptor responses by

drift is used to enhance spatial frequencies that are underrepresented in natural images (in the spatiotemporal domain, by creating energy in non-zero temporal frequencies). This produces a more equal power distribution, “whitening” in statistical image processing terms - but through *efferent* motor activity rather than *afferent* signal processing (Kuang et al., 2012; Aytekin, Victor & Rucci, 2014; Rucci & Victor, 2015; see also Barlow, 2001). The common view of FEM has been simply that too little or too much movement is bad for visual accuracy (leading to receptor adaptation or smearing, respectively). The temporal encoding hypotheses reassess the role of drift in vision, as *active gaze* strategies, means of creating temporal structure in signal in the optic nerve essential for accurate vision at the finest spatial scale (on active gaze in locomotion, see Section 4.1.).

Microsaccades are the most extensively studied FEM. They are small conjugate movements (less than $\frac{1}{2}^\circ$), jerky (REM-type), and occur at a rate of 1-2 per second (i.e. not necessarily observed in each fixation, typical fixation frequency being much higher). The frequency of microsaccades depends on the visual acuity and gaze stability instructions of the task: they are most readily observed in tasks where the subject is asked to sustain fixation on a small stationary target, but can be also voluntarily suppressed in these conditions (Steinman et al., 1973), they are rare in more natural active tasks (Aytekin, Victor & Rucci, 2014), and in highest-acuity tasks such as threading a needle their occurrence depends on the task phase. Namely, when the thread approaches the needle there are microsaccades with decreasing amplitude between the thread and the needle (Ko, Poletti & Rucci, 2010; see also Valsecchi & Gegenfurtner, 2015), which then fall away at the final phase when the thread must make contact with the needle. This dependence on task phase, incidentally, explains the previous result where microsaccades were reported to be absent from high-acuity tasks in Winterson & Collewijn, (1976; cf. Kowler & Collewijn, 2010).

Microsaccades fall on the saccade main sequence, which makes the “rapid” type oculomotor events, although because of their small amplitude the actual velocity may be well below some “slow” eye movements (a point we will return to below). There has been some debate on the function – and proper definition – of microsaccades (Collewijn & Kowler, 2008, Ditchburn, 1980, Kowler & Steinman, 1980; Martinez-Conde et al., 2004, Martinez-Conde et al., 2009; Martinez-Conde, Otero-Millan & Macknick, 2013; Steinman et al., 1973; Poletti & Rucci, 2016). The most common verbal definition is that they are *small, involuntary, and occur during intended fixation*. In practice, an amplitude definition (“small”) is used as the operational criterion. “Involuntary” is a concept that cannot be given a clear physiological or experimental interpretation, as practically *all* saccades, we make – irrespective of size – are automatic and occur without declarative *awareness*. The same goes for “intended” fixation. Only when “intended” means *instructed* by the experimenter (a “fixation task”), can we tell whether the subject “intends” to maintain fixation. In naturalistic tasks “intention” cannot be decoded from oculomotor data. (Note that the intention is a necessary part of the definition here: defining microsaccades simply as saccades that occur during fixation would be circular, because it presupposes a classification of saccades into saccades that break fixation and those that do not).

Any careful discussion of microsaccades needs therefore to keep in mind the difference between the “classical” and “modern” era (see discussion in Collewijn & Kowler, 2008; Martinez-Conde, Otero-Millan & Macknick, 2013; Nyström et al., 2016; Poletti & Rucci, 2016). The “classical” studies (1950s–1980s) were conducted with high-precision but invasive techniques used on highly experienced subjects (typically the experimenters themselves). The “modern” research (since the beginning of the 21st century) uses less precise video-oculography, and typically naïve subjects. In the “classical” literature reported microsaccade amplitudes are much smaller (12’ is often mentioned as a typical value), and there the debate tended to revolve around whether these *very small* microsaccades serve some *special* function, such as preventing visual fading by refreshing the retinal image, as opposed to directing gaze in the visual scene for detailed visual analysis of new targets, the primary function always attributed to “normal” saccades. In the modern literature, saccades up to 2° amplitude have been interpreted as “microsaccades”. the most commonly used conventional value to operationally define microsaccades seems to be 1° , however. Based on anatomical grounds (the size of the rod-free retinal bouquet) and

behavioral grounds (the distribution of saccade amplitudes during sustained fixation) Poletti & Rucci (2016) advocate a conventional value of 0.5° as the *de facto* operational criterion. This also seems to be in line with much of the previous research (see Martinez-Conde et al., 2004, their Table 3).

After some wrangling, the field also seems to have converged on a consensual view that most saccades, right down to the classical “micro” range serve an exploratory function; that is accurately (re)align the line of sight from the preferred retinal locus of fixation to the target of interest (Engbert & Kliegl, 2004; Ko, Poletti & Rucci, 2010; Martinez-Conde, Otero-Millan & Macknik, 2013; Poletti & Rucci, 2016; Steinman et al., 1973). It is in itself interesting that this should be so, as the amplitudes of these microsaccadic gaze shifts are so small in relation to size of the high-resolution fovea. The fovea is generally considered to be a few degrees in diameter, and the rod-free region with highest cone density (foveola) about 1° . So one would not think microsaccades would be necessary from the perspective of “foveation” of targets – often taken to be equivalent to high-acuity sampling, and synonymous with “fixation”. However, the preferred retinal locus of fixation is much smaller (and actually does not correspond anatomically to the locus of highest cone density, Putnam et al., 2005). Careful measurements have shown that visual acuity for minute detail drops already a few arc min from the preferred fixational locus, and that when the very finest highest detail is relevant to the task the detail is scanned with microsaccades. Thus the retinal region *used* in highest-accuracy vision is considerably smaller than the anatomical fovea (Ko, poletti & Rucci, 2010; Poletti, Listorti & Rucci, 2013; but see also Valsecchi & Gegenfurtner, 2015).

A number of studies assessed the potential role of microsaccades in preventing fading or restoring vision of faded targets (especially of low-contrast peripheral targets). As fading is assumed to be caused by retinal adaptation, the idea is that “refreshing” the retinal image with a sudden visual transient (not to be conflated with temporally structuring the signal with slow control, above) could be useful for maintaining clear vision (Donner & Hemilä, 2007; Engbert & Mergenthaler, 2006; Martinez-Conde et al., 2006; McCamy et al., 2012 Rolfs, 2009; Gaarder, 1966, cited in Rolfs, 2009,). This function has been contested, however (e.g. Ahissar et al., 2016; Kagan, 2012; Mostofi, Boi & Rucci, 2015; Collewyn & Kowler, 2008, and Poletti & Rucci, 2016, review several arguments). And, anyway, in naturalistic tasks frequent saccades, head movement and optic flow from locomotion make retinal adaptation an unlikely problem.

2.3. Slow eye movements (SEM): multiple systems for image stabilization

Smooth or “slow” eye movements are a class of eye movement where the eye moves relative to the head in a smooth rather than a jerky manner. Despite their name, eye-velocity is in the range of smaller saccades. So the terms “rapid” and “slow” are somewhat misleading. SEM are typically studied in the laboratory by presenting moving stimuli to stationary subjects, or by inducing active or passive head-movement while presenting stationary stimuli. Unlike saccades, slow eye movements cannot be performed “voluntarily”, i.e. initiated without an appropriate moving target stimulus. Smooth pursuit eye movements (SPEMs) are studied with moving dots on a textured or featureless background (focal stimuli), and optokinetic responses with coherent visual flow (ambient stimuli) such as an optokinetic barrel – both of them typically with head stabilized subjects. The vestibulo-ocular response (VOR) is studied with head movement and stationary stimuli. Note that the implicit idea here is to experimentally “isolate” separate oculomotor subsystems with specific functions by presenting stimulus conditions which, unlike real-world situations, only vary on one or few stimulus parameters– parameters to which *only the system of interest* is thought to be responsive.

The optokinetic and vestibulo-ocular response are sometimes called compensatory eye movements (CEM), as their function in natural conditions is thought to be compensate for head instability (based on the bottom-up error signals of retinal image slip and head acceleration). SPEM, on the other hand, is a “tracking” behavior, rotating gaze to maintain alignment on a target whose visual direction changes. SPEM thus resembles fixation functionally (the goal is to maintain gaze on a focal target) *and* saccades mechanically (quick, endogenously programmed rotation of the eye in its socket).

When the head or the body move, the vestibulo-ocular response is essential for the stability of gaze and posture, accurate perceptual orientation, and the sensation of self-motion. The vestibular system of the inner ear detects linear (otolith organs) and rotational (semicircular canals) accelerations of the head, and generates sensory signals that are used in the VOR (translational and rotational). The VOR stabilizes gaze and serves postural control by compensating for passive head rotation.

During VOR the eyes counter-rotate in opposite direction to the rotation and/or translation of the head. (Sustained acceleration produces nystagmus, i.e. alternating slow phases opposite to the direction of rotation and saccadic quick phases in the direction of rotation). For targets at infinity, the absolute rotational velocity of the eye and the head should be the same, and translation need not be compensated for. For nearby targets, eye rotation is scaled up with decreasing target distance to account for motion parallax and the displacement of the eye from the axis of head rotation (Hine & Thorn, 1987; Busettoni et al., 1994; Viirre & Demer, 1996; Crane, Viirre & Demer, 1997; for review see Raphan & Cohen, 2002).

The VOR can be elicited as a very fast response to head-acceleration (well under the latency for compensatory visual responses). A three-neuron reflex arc has been identified as the most basic level of organization of the response. But when integrated in natural behavior quite complex neural organization is involved. Consider for example the initiation of a “saccadic” head/eye movement (combined eye-head gaze shift, see below), that re-orients gaze to a new visual target. The behavioral goal (rotating towards the target) is opposite to the reflex response (counter-rotating the eyes). Accordingly, during *active* movements the vestibular response needs to be suppressed, which means the VOR control system needs to be able to distinguish between vestibular stimulation due to active and passive head-movement. That the suppression is not present during physiologically similar passive head rotation indicates that the motor plan (or sensory prediction based on efference copy) is involved, at least for slower rotations (Herdman Schubert & Tusa., 2001; Roy & Cullen, 2001). Rather than a simple mechanical response to vestibular stimulation, the VOR thus displays flexible adaptation to task, due to cortical and cerebellar modulation. This is essential to synergistic coordination of the reflex response and higher-level motor goals.

Optokinetic nystagmus (OKN) is a periodic eye movement pattern that consists of a slow phase (SP) following coherent image motion (flow) on the retina/in the visual scene and fast re-setting saccadic quick phases (QP). Two subtypes of OKN have been identified: “bottom-up” stare OKN and “top-down” look OKN.

Stare OKN (*Stier-Nystagmus*) can be observed, when no gaze target as such is selected. For example, when a person is looking out of a train window at an angle relative to current heading. Here the SP may be considered an *optokinetic response* (OKR) elicited by coherent image motion (image slip) spanning the entire visual field, or at least a large portion of it. OKR stabilizes the retinal image, removing or reducing retinal image slip caused by difference in angular velocity between the eye and the textured stimulus, and is a form of response found in lower vertebrates lacking a cortical visual system. Like the VOR, OKR can be considered a reflex-response for stabilizing gaze with respect to allocentric targets by compensating for head and body motions, but on the basis of an optical (retinal) motion signal rather than inertial (vestibular) information. Also, like VOR, OKR responses are produced to compensate for rotational and translational motion. (Compensation for translation involves vergence response as well as downward rotation, is scaled for viewing distance, and translational OKR uses depth information from binocular disparity. For review see Angelaki & Hess, 2005).

Look OKN (*Schau-nystagmus*) occurs when the subject attentively observes moving targets. For example, when a person is standing by the railroad and looking at a railway car moving past. It is partially cortically driven (not only retinally), is not as stereotypically periodic, and look OKN SP phase may actually be indistinguishable from smooth pursuit (Wyatt & Pola, 1988).

It is thus important not to confuse optokinetic *nystagmus* (the observed eye movement pattern) with optokinetic *reflex* (the retinal slip correcting closed-loop mechanism that may be just one of the mechanisms underlying OKN SP, perhaps prevalent in stare OKN but not less so in look OKN). The QP of

look OKN - but not stare OKN - also falls on the same main sequence as between-fixation saccades (Kaminiarz, Königs & Bremmer, 2009); its oculomotor characteristics also suggest look OKN QP may be organized by the same saccade planning networks as “non-OKN” saccades (Harrison, Freeman & Sumner, 2014).

Smooth pursuit eye movements (SPEM) track a visual object moving in relation to the head (usually against a textured background). Unlike saccades that can be made “voluntarily” towards locations “in empty space”, pursuit movements require a target to track (but not necessarily a visual one, see below). An attempt to move the eyes smoothly without a moving target usually results in a series of saccades.

Early models of SPEM were reflex-like closed loop models, describing a stimulus-dependent feedback system where pursuit is elicited and maintained by correcting retinal position error of the stimulus – typically a small moving spot (for a review of review early work see Lisberger et al., 1987; for a critique of shortcomings of closed loop response models see Steinman, 1986; for review of more recent work see Krauzlis, 2004).

However, in more naturalistic tasks the observer will be familiar with the underlying dynamics and hence the tracked stimuli often behave in a *predictable* way. It has long been known that in naturalistic conditions oculomotor control of SPEM – like saccades – can utilize predictive information about anticipated rather than observed motion (Kowler, 1989; see also Hayhoe et al., 2012 and section 4.2., below). Smooth eye movement can be initiated in anticipation of predictable target motion (Barnes & Asselman, 1991; Kowler & Steinman, 1979) and humans are able to maintain pursuit (with catch-up saccades) of a target that is momentarily occluded (Bennett & Barnes, 2003, 2006; Orban de Xivry, Missan & Lefèvre, 2008). This latter result is important, because it shows that pursuit can be driven by a spatial memory representation even in the absence of a visual stimulus (see Orban de Xivry et al., 2013 for a computational model). Moreover, this spatial representation can guide positionally accurate tracking even though the desaccaded pursuit velocity “decays” during occlusion. So overall, dependence on the presence of a *retinal* stimulus target is not as strict as originally envisioned: a *proprioceptive* target is sufficient and in suitable conditions (brief occlusion), pursuit can continue even without sensory target.

Humans can successfully perform smooth pursuit to visual targets moving over 180 deg/s but especially at higher eye-velocities (>30 deg/s) gaze tends to lag behind target. This is compensated by catch-up saccades. Unlike OKN QP, of course, the pursuit saccades are in the same direction as the smooth movement. Except for very low angular velocities, the actual observed movement during pursuit is thus a complex movement pattern consisting of smooth and jerky components. Consequently, the maximum speed of “pure” pursuit is sometimes said to be around 30°/s, and about 150°/s has been reported as saturation limit for the *desaccaded* velocity component (Lisberger et al., 1981), which is not the same thing as the maximum reliable tracking velocity. Because pursuit gain (< 1.0 except for slow speed) is computed by comparing *desaccaded* eye-velocity to angular velocity of the stimulus, however, it is not an accurate measure of the success of foveation. The rationale is instead that it is an index of pure pursuit generator output, as it were, and in tracking a “slow (pursuit generator) system” and a “rapid (saccade generator) system” are working in cooperation.

2.4. Oculomotor event types as indices of “modular” control?

In a typical laboratory setting (e.g. reading a text, looking at pictures, or performing visual search or pursuit on a computer display) the body and the head do not move. Oculomotor control and the gaze shifts needed to pick up relevant information therefore only involve movement of the eyes in their sockets. But in the natural ecological context of human cognition, completely sedentary activities are the exception, not the rule, making *gaze* control the fundamental control task. So, how are the classical eye movement “types” related to gaze control in complex streams of motor activity typical of real-world behavior?

A powerful assumption is that the classification reflects distinct subsystems for oculomotor control subserved by distinct neural control circuits, “generators”. Each generator is thought to be

capable of generating one of a small suite of possible movements and recruited at different phases of the activity depending the overall goals and constraints a particular task. The idea is, in other words, that even if the real world does not present itself as saccade tasks, pursuit tasks or fixation tasks, nevertheless presenting the subject with a *saccade task* one can isolate *the saccadic system*, by having them visually *track* a moving target one can engage *the pursuit system*, by introducing endogenously or exogenously generated head rotations one can study the compensatory *vestibulo-ocular system* etc. Experimental control would thus allow us to observe each eye movement *type* in their “pure” form. This, then, allows physiologists to identify and trace anatomically bounded “saccade generators”, “pursuit generators”, “vestibulo–ocular reflex circuits”, and their circuit behavior (any putative “generator” is a widely distributed network, rather than a localized nerve center, Munoz, 2006). This typological² approach has served to organize research into neurobiological underpinnings of oculomotor control.

If complex naturalistic behavior is in this way composed from the classical eye movement types, it would mean that the oculomotor system carries the hallmarks of *modularity* (most researchers’ actual position probably falls somewhere between strict modularity and completely equipotential holism). Modularity in motor control has a long history. It is found in many guises, such as stimulus-driven *reflexes* (Easton, 1972; Pavlov, 1927; Sherrington, 1906;), *central pattern generators* for endogenous rhythmic behaviors (Brown, 1911; Grillner & Wallén, 1985; Grillner, 2006; Marder & Bucher, 2001; Pearson, 2000), learned *motor programs* or *schemata* underlying skilled behavior (Keele, 1968; Lashley, 1951; Newell, 2003; Schmidt, 1975, 2003; Shea & Wulf, 2005; Summers & Anson, 2009), and neuromotor *synergies* (Bernstein, 1967; Lee, 1984; Turvey, 2007; Ting & McKay, 2007) in the biomechanics literature, where it has recently been given a computational expression (Bizzi, Mussa-Ivaldi & Giszter, 1991; D’Avella, Saltiel & Bizzi, 2003; Mussa-Ivaldi & Bizzi, 2000; Mussa-Ivaldi, Giszter & Bizzi, 1994)³.

It is thus relevant to ask on the one hand whether and to what extent it is possible to individuate different *neural circuits* responsible for different eye movements and how strict the traditional *classification boundaries* themselves are. Actually, in general the historical tendency in eye movement research has actually been to identify subtle commonalities in the oculomotor “classes”, thus eroding the categorical textbook dichotomies.

We have already discussed how saccades, microsaccades and look OKN QP might emerge from a unified REM control system in different experimental task settings, rather than being completely different “types”. Also generation of pursuit and saccadic movement may be handled by largely similar circuitry – perhaps in the case of pursuit merely driven by an additional *visual motion estimate signal* from parietal cortex (Harrison, Freeman & Sumner, 2014; Krauzlis, *ibid*). This blurs the REM/SEM distinction. On the other hand, as already discussed, it is not always straightforward to differentiate between pursuit from some cases of OKN SP, either⁴.

It has also been debated for some time to what extent fixation and smooth pursuit really are distinct oculomotor processes, or whether they actually reflect the same cognitive neural resources, to

² Typological here means: treating the classification of eye-movements not as an abstract means for data reduction, but reified as phenomena in their own right.

³ As the others are verbally rather than computationally formulated, they are not strictly mutually exclusive – and, indeed, are sometimes used almost interchangeably so that the different explanations of how the very same complex movement patterns are generated may differ in name only (compare e.g. Easton, 1972; Grillner, 2006).

⁴ The main difference between optokinetic response and smooth pursuit is that the latter refers to stabilization of gaze on a focal target object based on a *cortical* estimate of target velocity, not just *retinal* image slip. Stabilization of coherent retinal image pattern is considered optokinetic response. But, except for completely unpredictably moving large-scale texture, this distinction is not a very clear one. In neither case can a focal, moving visual stimulus or retinal slip stimulus account for the eye-movement. The distinguishing feature of pursuit is “higher-order” cortical representations of target motion that integrates visual, somatosensory and predictive/memory information.

the extent that pursuit could even be considered fixation of a target moving (relative to the head) and fixation considered “pursuit” of a stationary target: two modes (emergent behaviors) of a complex system coordinating the stable orientation of the visual axis relative to focal targets (Krauzlis, 2004; Orban de Xivry & Lefèvre, 2007). Murphy, Kowler & Steinman (1975) already suggested that fixational drift is pursuit of a stationary target. These considerations, and the hypotheses of drift as a form of active slow control (Ahissar & Arieli, 2012; Ahissar et al., 2016; Aytekin, Victor & Rucci, 2014; Kuang et al., 2012; Rucci & Casile, 2005; Rucci & Victor, 2015; Steinman et al., 1973) blur the FEM/SEM distinction.

Of course, pursuing a focal target against a textured background puts the pursuit and OKR movement in an antagonistic relationship, in that successful pursuit will create precisely the kind of global retinal image motion that OKR attempts to cancel (a similar situation that was discussed above in relation to active head-movement and VOR). It is not, however, obvious that two completely distinct and opposite systems need be involved, as one can equally well imagine that the pursuit and OKR can be brought into *synergy* by limiting the retinal region used to drive the OKR to the region spanned by the pursuit target (cf. Barnes, 2008).

Even the most fundamental distinction of fixations and saccades is not entirely dichotomous because of microsaccades during “sustained” fixation (i.e. instructed fixation). There seems to be no consensus on a clear physiological means to define when a saccadic movement is “really” a REM (separating successive fixations) and when it is “only” a FEM (occurring during fixation), beyond conventional amplitude criteria⁵. Otero-Millan et al. (2013) further show that in a task of free-viewing a naturalistic image, reducing the size of the image leads to smaller and smaller saccades, proposing that rather than opposite sides of a fixation vs. saccade dichotomy, microsaccades and saccades are just different parts of an “oculomotor continuum” of exploration of areas of different spatial scales. Ko, Poletti & Rucci (2010) and Poletti, Listorti & Rucci (2013) actually show that the range of “exploratory” saccade amplitude reaches all the way down to the “classical” range of 5 min arc, well below the “modern” standard heuristic thresholds of 0.5°–1°. This, together with evidence on common neurophysiological basis for saccades of all sizes (Hafed Goffard & Krauzlis, 2009; Hafed & Krauzlis, 2012) has led to the suggestion that the difference between fixational REM and other REM is not fundamental; so that generation of microsaccades during oculomotor fixation can even be “functionally equivalent to visual exploration on a spatially focused scale” (Otero-Millan et al., 2013, p.6175; cf. Martinez-Conde et al., 2009; Otero-Millan et al., 2008; Polett & Rucci, 2016). Note, though, that this is not the case for slow control FEM: drift velocity does *not* scale with the size of the explored visual workspace.

In summary, considering pursuit, OKN, saccades and fixation as entirely distinct behaviors is not as fundamental as one might think. OE classes are a powerful *descriptive* framework for communicating details of eye movement in traditional tasks, but the case for their being an *explanatory* framework is not as solid. Historically, there has been a shift towards analysing the complex movement patterns of the eyes from a more goal-driven, functional, or task-based perspective – not just as instances of eye movement “types” akin to separate reflex arcs responding to specific stimulus variables (Steinman, 1986; Steinman, Kowler & Collewyn, 1990; Krauzlis, 2004; Otero-Millan et al., 2013; Orban deXivry & Lefèvre, 2007).

The circuitry responsible for selecting and maintaining vision on targets, even in restricted laboratory tasks, seems too intricate and flexible to be captured by simple typologies. And, as we shall see, this issue is compounded in naturalistic settings where the gaze must be coordinated and controlled in the context of observer motion where understanding gaze behavior calls for analysis of integrated

⁵ It is often said fixation is not “released” during a FEM, unlike saccades and smooth pursuit. This “release”, however, is not always well defined physiologically, but may refer only to judged compliance with instruction. I.e. the subject *attempts* to fixate, but the voluntary fixation is interrupted by *involuntary* movements. “Volition” in this case remains an ephemeral concept without a rigorously defined neurological basis.

representations of space and motor actions across multiple sensory systems, multiple motor effector systems, and multiple frames of reference.

3. Eye movement behavior in the wild

Measuring and modeling visual behavior in natural ecological contexts, rather than specially designed laboratory settings, is a relatively new sub-field in eye movement research. With the advent and gradual spread of wearable and vehicle-mounted eye tracking devices and immersive virtual reality environments, the study of eye movement behavior has ventured outside the confines of typical laboratory tasks. Which objects or locations are fixated (to glean high resolution foveal information and/or stabilize gaze), and in what order (how the “scanpath” relates to global scene structure and task organization) have been explored in a variety of natural tasks. These include making tea (Land, Mennie, & Rusted, 1999), making a sandwich (Hayhoe, Shrivastava, Mruczek & Pelz, 2003), steering a car (Land, 1992; Land & Lee, 1994; Lappi, Pekkanen & Itkonen, 2013) and sports such as cricket (Land & McLeod, 2000; Mann, Spratford & Abernethy, 2013) and squash (Hayhoe et al., 2012). This research has begun to reveal recurring patterns of gaze behavior that are surprisingly regular and repeatable (reviewed in Hayhoe and Ballard (2005), Kowler, (2011), Land (2006), Regan and Gray (1999), Tatler & Land (2011) and Tatler et al. (2011)). Many of the common findings can be summarized in terms of seven “qualitative laws” of gaze behavior in the wild, i.e. recurring patterns that tend to go together (the more so the more naturalistic the setting, *all* of them would be expected in most extended sequences of fully naturalistic behavior).

1. *Gaze behavior is often highly repeatable.* Given task and context, gaze tends to scan the same set of visual targets, and focus predominantly on a few. This *stereotypy* applies both between trials of the same individual and even across individuals (Hayhoe, Shrivastava, Mruczek & Pelz, 2003; Land, 2006).

2. *Gaze is focused on task-relevant targets and locations.* Explaining the distribution of fixations in a scene requires “top down” analysis of task requirements, as opposed to predicting scanning patterns from “bottom-up” visual saliency of stimuli (Tatler et al., 2011, cf. Torralba et al, 2006; Yarbus, 1967).

3. *Individual fixations have interpretable functional roles.* The roles, though, are not always transparent or intuitive – the pattern of gaze can often be quite surprising to the subject him/herself (e.g. Land & Lee, 1994). They can sometimes be understood in terms of the reward or the value of the information gleaned by fixating the target (Hayhoe & Ballard, 2005; Tatler et al., 2011) and/or actively shaping the visual input (e.g. Kim & Turvey, 1999; Wann & Swapp, 2000), thereby possibly simplifying computational problems in perception and motor control (Ballard, 1991; Fermüller & Aloimonos, 1995). Thus, the interpretation may not always be transparent at the level of our commonsense understanding of the task/scene semantics.

4. *If possible, targets are fixated “just in time”,* as opposed to looking at targets well ahead of time - which would require maintaining information gleaned during the fixation in short-term memory (Ballard et al., 1995). In complex tasks, the deployment of gaze is thus tightly coupled to the information requirement of the imminent *subtask*. Unless the (sub)task requires continuous monitoring/tracking, gaze disengages – i.e. switches to a new target – before (sub)task completion (Land, Mennie & Rusted, 1999). Typically gaze leads action by about 1s (Furieux & Land, 1999; Land & Lee, 1994), although see Land (1992) and Herst, Epelboim & Steinman (2001), showing that this is not a rigid pattern but partly depends on task, stimulus layout, and predictability/prior knowledge.

5. *In skilled/routine behavior these “just-in-time” or guiding fixations are interleaved with look-ahead fixations* – that is, fixations to objects and locations that will be relevant in a later subtask, implying some short-term memory (Pelz & Canosa, 2001; Mennie & Hayhoe, 2007; Lehtonen et al., 2013).

6. *Integration of visual and spatial information across saccades and changing vantage points does occur.* Memory can be used to re-orient in 3D, even to targets currently outside the field of view, implying longer-term spatial memory (Tatler & Land, 2011). Also the fact that when moving to a new sub-task

there are no fixations to irrelevant objects (visual search) implies stable contextual representation of where the now task-relevant objects and locations are in 3D space (cf. e.g. Land, Mennie & Rusted, 1999).

7. *Gaze control is part of coordinated eye/head/body/locomotor control.* On the one hand *gaze shifts* are achieved by rotating not only the eyes but also rotation and translation of the head and the body. On the other hand, head and body movements in space are compensated for by *gaze-stabilizing* eye, head and body rotations (e.g. Steinman, Kowler & Collewyn, 1990; Collewyn et al., 1992; Grasso et al., 1996; Land, 2004).

For the present purposes, 6 and 7 are the most pertinent. This is because 6 & 7 exhibit how control of gaze is in essential ways embedded in complex motor coordination and spatial representation in multiple coordinate systems, whereas 1-5 can also be applied to sedentary tasks. However, before considering coordinate transformations and spatial representation in gaze control in more detail, we first need to discuss how the concept of “fixation” in 1-5 differs from the OE concept discussed in the previous section.

3.1. Oculomotor events vs. gaze behavior: how to define “a fixation” in natural behavior?

The eye movement pattern in most sedentary tasks – such as scene perception, visual search and reading is characterized by fixate-saccade-fixate sequences. Naturalistic studies of many real-world behaviors also often report “fixation behavior”. But what this actually *means* in naturalistic eye movement studies is most often “fixating” a point of regard (typically a putative visual target manually identified in a head-mounted scene camera image), and/or aggregate parameters such as total dwell time within an area of interest, AOI). “Fixation”, thus, can refer to very different underlying physiological states in real-world contexts and “fixation tasks” in the laboratory. This potentially conflates quite different neurophysiological states under a catch-all label of “fixation”. In this light it is perhaps a sign of immaturity of the field that gaze behavior is described in terms of “fixation behavior”, and the lab-derived *terminology* - “fixation”, “pursuit”, “(micro)saccade” - is applied beyond their original domain of application (rather than carefully describing the actual *movement pattern* of the eyes during a natural task).

Consider the case of a car driver “fixating” a point on their future path. Perhaps s/he is looking at an obstacle such as a pothole or a puddle on the road appearing from behind an occlusion in a bend, or perhaps s/he is just taking the visual fix on a point in the road s/he is intending to travel over. As s/he approaches the visual target, the horizontal eccentricity and the vertical declination of the target point move continuously. Thus, a “fixation” that maintains the target on the fovea is actually a *pursuit movement* in a head-centered (egocentric) frame of reference. As this pursuit movement corresponds in magnitude and direction to the large-scale optical flow of road texture at and around the location of interest, *optokinetic* control (retinal slip compensation) may be involved as well. What this example illustrates is that when taken outside the laboratory, “fixation” becomes an underspecified concept, and the required oculomotor control is highly task-dependent⁶.

“Fixation” has a physiological meaning – stabilization of the eye in relation to the head, a specific class of oculomotor event - and another meaning as a type of gaze behavior; and in the general case these are not equivalent (Figure 3). To be clear, what is meant here by an *oculomotor event* is the *systematic pattern in the movement of the eye in its socket*. *Gaze behavior* in turn refers to *the movement of the gaze vector in 3D space*; i.e. the rotation and translation of the line of sight from the point of vantage (at the

⁶ Sometimes the expression “fixation of a moving target” is used, clearly implicating smooth pursuit eye movement, and Kim & Turvey (1999, p.245) even speak of “pursuit fixations” and Lappi, Pekkanen & Itkonen (2013) about “optokinetic pursuit”. To avoid the use of such oxymoronic terms, some prefer to use the term *foveation*. This is not satisfactory, however, as it suggests gaze control is almost *by definition* driven by the need to target *foveal* vision, rather than stabilizing *peripheral* vision.

eye) to the point of fixation (in the scene), either of which may be stationary or moving in the 3D scene frame of reference (Figure 4). In the naturalistic eye movement literature “fixation” usually means gaze fixation (fixing the point of regard on an object or a location), while the laboratory/physiological literature uses the term to refer to oculomotor fixation. Confusion may occur if these two meanings are conflated. In loose discourse it is of course acceptable to talk about “fixation strategies” (Land, 2006) or the cognitive mechanisms underlying “guiding fixations” (Mennie, Hayhoe & Sullivan, 2007) or “just in time fixations” (Ballard, Hayhoe & Pelz, 1995), as long as one is careful to make clear in the context just what is “fixed” - and in which frame of reference.

--- INSERT FIGURES 3 & 4 HERE ---

3.2. Frames of reference

Changes in eye-in-head position only make sense in the context of known frame of reference transformations. “Where people look” or “how subjects move their eyes” in natural behavior are not well-defined questions unless the relevant coordinate system is specified⁷, as locations and movements are always *relative* to some frame of reference.

Representing the multiple coordinate systems with appropriate transformations is a *methodological* challenge: if the head and body are not forced to remain at rest, gaze direction and gaze landing points can only be computed if the 3D head and body movements are explicitly modeled. A rather obvious methodological upshot is that OE detection based on the eye-in-head position signal may not work. For example, the fixation duration and counts or saccade amplitudes or velocities may be spurious if head movement (which elicits optokinetic and vestibulo-ocular eye movements) are not handled properly (Kinsman et al., 2012). Handling the necessary coordinate transformations presents challenges which have gradually begun to be addressed in the eye tracking technical literature (e.g. Cesqui et al., 2013; Diaz et al., 2013; Duchowski et al., 2002; Epelboim et al., 1995; Kinsman et al., 2012; Lanata et al., 2015; Munn, Stefano & Pelz, 2008; Munn & Pelz, 2009; Reimer & Sodhi, 2006; Tomasi et al., 2016; Vidal, Bulling & Gellersen, 2011).

But the choice of coordinate system is essential also for posing meaningful hypotheses defining the phenomena of interest: a *theoretical* challenge. The computational problem for the brain that we are interested in here is that sensory information is encoded in multiple reference frames. Thus, the sensorimotor transformations responsible for *oculomotor* control must at the same time ensure the movement of the eye in its socket is appropriately coordinated with motor commands and stimulus events in the multiple coordinate systems relevant to *manipulation* and *locomotion*. Behaviorally significant frames of reference are shown in Figure 5.

--- INSERT FIGURE 5 HERE ---

Frames of reference are usually divided into *allocentric* (or *spatiotopic*) frames of reference that are fixed to objects, locations or reference directions in the 3D scene, and *egocentric* frames of reference. Egocentric frames include eye-centered (retinotopic), head-centered (craniotopic), body-centered and locomotor frames of reference.

⁷ A *frame of reference*, as the concept is used here, is a set of objects, locations or reference directions that maintain their spatial arrangement over time, and can be used to represent space, as a basis for a coordinate system. A *coordinate system*, on the other hand, is a system for assigning sets of numbers, coordinates, to objects or locations in physical space in some frame of reference. It requires a point of origin and a distance metric. For present purposes, however, the terms can be used interchangeably.

The term retinotopic is here used for any frame of reference fixed to the eye, although the retina only covers part of the inner surface of the eye and targets behind the eye have no *retinal* projection as such. Also, body-centered here means locations and motions to some convenient reference frame, such as the midline and dorsoventral axes, but of course the body is not a rigid object and body parts - limbs in particular - move relative to one another, which means “the body” actually comprises of multiple coordinate systems. The locomotor frame of reference for terrestrial locomotion can be defined by instantaneous heading and its normal in the horizontal ground plane as the reference axes.

The physical targets for eye and body movements (and the eye and the body themselves) reside in an allocentric 3D scene, but in general allocentric spatial locations and relations are not directly specified in any of the sensory frames of reference directly available to the brain (Figure 6). Visual information is encoded in retinotopic coordinates, i.e. the signal in the optic nerve changes with change in retinal image irrespective of whether this change is a result of target motion, eye movement or locomotion. The same goes for auditory information (craniotopic). Also vestibular information is encoded in craniotopic coordinates, and ambiguous with respect to gravitational vs. locomotor acceleration. Somatosensory information (kinaesthetic, touch, pain, heat) comes in body coordinates, and is ambiguous with respect to active vs. passive contact forces - whether for example a change in pressure at the foot is caused by body or ground movement.

Head coordinates of a target can be recovered from its retinal position when this information is combined with known eye position. Body-centered coordinates can in turn be recovered from retinal coordinates, eye-position, and head pose combined etc. In general, then, target location and motion in the 3D scene, as well as changes in ego location and heading must be inferred by integrating multiple sensory signals (retinal, proprioceptive, vestibular...), and motor efference copy, and very likely spatial memory as well.

Worth pointing out is that the egocentric vs. allocentric distinction and the directly accessible vs. indirectly inferred distinction do not overlap perfectly. Specifically, the locomotor axis is an egocentric frame of reference – objects’ and locations’ positions in this frame of reference depend on the self-motion of the observer – this is not an egocentric reference frame that the brain has “direct access” to. (This is, incidentally, also the case for *egocentric vertical* – important in the control of an “upright” posture). It seems intuitively clear enough that the brain has no direct access to *allocentric* spatial relations in the external environment - such as the distance to targets and obstacles (depth) or distances between objects (e.g. gaps between obstacles) which nevertheless need to be judged visually to make locomotor decisions, but it may not be as intuitively obvious that the brain does not have direct access to all egocentric frames. But because no sensory signal alone directly gives directions or distances relative to egocentric locomotor heading, if heading or visual direction relative to heading is to be used for locomotor control, then it must be recovered from sensory information, just like allocentric relations, through multisensory integration.

--- INSERT FIGURE 6 HERE ---

3.3. Internal representation of space: multisensory integration & sensorimotor coordinate transformations

Multisensory integration refers here to the way multiple sensory cues are combined with each other (and efference copy signals and spatial memory) to derive robust estimates of self-motion and target motions. This kind of disambiguation of motion signals is a fundamental processing principle of sensorimotor and spatial information in the brain (Sperry, 1950; vonHolst & Mittelstaedt, 1950), important in controlled visual orientation in 3D space, gaze fixation and visual stability (Bremmer, 2011; Green & Angelaki, 2010; Roy & Cullen, 1998, 2001, 2002).

Every time we move our eyes, spatial updating of (retinotopic) target locations is needed. *Remapping* cortical representations across saccades helps us to maintain perceptual and gaze stability (for review see Wurtz, 2008). Recent experiments suggest that even in sedentary updating both

oculomotor efference copy and proprioceptive feedback are used to update gaze/target representation (Poletti, Burr & Rucci, 2013). When we locomote, similar updating of changing spatial relations relative to the point of vantage may be involved.

For example, when subjects are allowed to preview a visual target in a natural scene (at distances of 2–22m) and then walk to them without vision, they can do this very accurately on average (constant error only some decimeters, even for the longer distances), with increasing variability with increased walking distance (Thomson, 1983; Rieser et al., 1990; Loomis et al., 1992; see also Fukusima, Loomis & DaSilva, 1997 and for review Loomis et al., 1996). This indicates that the subjects are able to derive accurate metric information about the direction and distance to the gaze target in egocentric space, translate the visual target into a *locomotor* target (select appropriate locomotor action to follow a path to the point of fixation), *spatially update* the internal representation based on vestibular, somatosensory and proprioceptive cues (ambient auditory cues were prevented by playing the subjects noise through headphones), and, finally, to use the updated representation to control the unfolding motor responses. That this pattern is not materially affected by delaying the response or by walking speed indicates that the memory error depends on the spatial distance, so that the error may be generated from the number of “update operations”, and their associated sensorimotor coordinate transformations. This amodal egocentric spatial representation has been called a *spatial image* (Loomis, Klatzky & Giudice, 2013; see also Tatler & Land, 2011).

Gaze also appears to be controlled relative to the spatial image: when subjects walk on curved trajectories in darkness, gaze rotation leads head and body rotation as it does in conditions of normal visibility (Grasso et al., 1996, Prévost et al., 2003, Authié et al., 2015; cf. discussion of SPEM of actively moved own body parts in darkness, above). Crucially, this is true both of walking forward and walking *backwards* (Grasso et al., 1998). This means that (1) the orientation of gaze in the direction of path curvature is a true spatial orientation response, not a hardwired pattern of motor coordination (in walking backwards the same orienting rotation relative to the *path* is in opposite direction relative to the *body*), and (2) both locomotion and gaze are controlled on the basis of on an internal spatial representation (in darkness neither can be driven by visual cues, so if the visual targets are the same as locomotor targets the future path must be “projected” into the spatial image).

To redirect *gaze* to an eccentric but visible visual stimulus, the retinal displacement of the target from fovea specifies the required eye rotation. However, to move the *hand* to the target, or to *walk* to the target whose position is known from the retinal image, limb and locomotor trajectories must be planned and appropriate motor programs selected - and these are encoded in proprioceptive and muscle coordinates. At some stage of processing the eye and body reference frame information must therefore connect. This is involved, for example, when you reach out across the sofa to pet a purring cat. Anatomically, posterior parietal areas are generally thought to be involved in representing egocentric space and sensorimotor transformations across modalities (Andersen et al., 1993; Pouget et al., 2002; Crawford et al., 2011; McGuire & Sabes, 2009). The hippocampal system in turn is thought to be involved in construction of allocentric “cognitive maps” (O’Keefe & Nadel, 1978; reviewed by Barry & Burgess, 2014). This is involved in knowing that the other cat is in the kitchen.

These egocentric and allocentric systems are anatomically and functionally connected via posterior cingulate/retrosplenial cortex, which is implicated in view-independent processing of scene information that integrates spatial information and “pieces together” information from locally observed snapshots (Park & Chun, 2009; Vann, Aggleton & Maguire, 2009). This is involved when you plan a route from the sofa, through the hall and into the kitchen.

How does the brain handle the coordination of sensorimotor activity across these multiple coordinate systems? One possibility is that there is one master representation for all of spatial perception and action used to plan action. This would imply that *all* sensory information would re-encoded for action planning into a single frame of reference, irrespective of target sensory modality or the motor system used for the response.

If there is such a master representation, introspection of a stable world might suggest this is allocentric. But much of the sensory analysis and motor preparation takes place in cortical areas that are retinotopically arranged - not just planning for eye movements but reaching as well (Batista et al., 1999; Cohen & Andersen, 2002) - suggesting the common code could be retinotopic. Or, there might be one transient egocentric “spatial image” in working memory, and one allocentric “cognitive map” in long-term memory. A further possibility is that *all* different elementary sensorimotor responses use specific coordinate transformations that are only resolved locally (perhaps in modular circuits responsive to very specific stimulus information available at the periphery).

These alternatives are not exhaustive, however. They are predicated on the intuitively appealing assumption that the 3D coordinate systems in Figure 5 are represented, explicitly and individually, in separate systems in the brain. That is: specific circuits responsible for maintaining representations in specific frames of reference, responsible only for coordinate transformations to and from those coordinates. However a somewhat counterintuitive alternative is suggested by research in computational neuroscience (see Andersen et al., 1993; Pouget & Sejnowski, 1997; Colby & Goldberg, 1999; Crawford et al., 2011; McGuire & Sabes, 2009). Multisensory neurons with large *gain fields* have no single proprietary frame of reference (e.g. instead of a classical *receptive field* in a retinotopic frame of reference the gain of a visual neuron may change with head rotation). In this case, the representation of space may be based on basis functions, effectively representing sensory and motor events in *multiple* frames of reference simultaneously (Pouget et al., 2002).

The way space is represented in the brain has potentially quite important implications for understanding the mechanistic basis of gaze control in eye movement and locomotor processes. *If* the sensorimotor transformations for converting visual (retinotopic) target information into locomotor and gaze (allocentric) behavior did correspond 1:1 to the 3D frame of reference transformations (see Figure 6), *then* there would be a niche for pure oculomotor control programs in the eye-in-head transformations. (Each OE might be indicative of a separate OE subsystem). However, as discussed in Section 2.4, the OE classification might be a powerful descriptive framework without providing an explanatory framework of the actual underlying mechanistic realization. Analogously, while the 3D frame of reference transformations are essential to accurately *describe* the *overt* gaze behavior and analyze the computational task the brain has to solve, the story of how the brain *internally* performs the task might turn out to be different. I.e. representation of 3D space need not decompose into separate neural representations individually corresponding to each of the geometric operations. And if spatial representation and central motor coordination is indeed integrated, across multiple sensory modalities and effector systems, it may not be possible to assign responsibility for each OE to specific neural circuits as eye movement control would be inseparably embedded in head-body-locomotor control. (This type of organization might go some way to explaining the historical trend to identify new oculomotor “types” only for them to be “unified” with other types).

4. Active gaze & feedback control as unifying themes

Trying to put together ideas of oculomotor event types as indices of separate modular systems, on the one hand, and the control of gaze embedded in complex sensorimotor tasks, on the other, does not always give a smooth fit. In the previous two sections we have focused on gaps. In this final section we take a more constructive look at a few unifying themes that could be invoked to smooth over some of the cracks.

4.1. Active gaze (and optic flow)

A major feature of the *active gaze* approach (Ballard, 1991; Ferrière & Aloimonos, 1995) is that rather than relying on invariant representations of 3D space, online visuomotor control is achieved on a fixation-by-fixation basis, with at most only sparse and task-specific information carried in trans-saccadic memory (Ballard, Hayhoe & Pelz, 1995).

Use of representations that are viewpoint-centered (point of fixation frame of reference) and simple to compute with active control of gaze (point of fixation) can be used to build efficient visual strategies that reduce the computational load of visuomotor transformations. Structuring feedback with active movement of the sensor is the essential feature (cf. the discussion on fixational eye movements in section 2.2). It simplifies downstream signal processing and constrains the problem space and; then computational problems that are ill-posed or intractable in general form can become well-posed and tractable in practice when the movement of the sensor is controlled and known. The problems of recovering shape, scene layout and 3D structure from flow being particularly pertinent for visual behavior in tasks involving locomotion).

Of course, the incoming signal is not structured by the movement of *only* the eye, but by the total movement of the eye, the head and the body in locomotion. Therefore, it is relevant here to consider optic flow as a visual stimulus. For an observer located at a geometrical point of vantage moving along a continuous trajectory through a textured environment, the pattern of visible light impinging on the current location is called the optic array of that point. Changes in the optic array to when the observer locomotes constitutes the optic flow experienced by the observer. In some situations optic flow directly specifies self-motion, and could therefore be useful for the control of locomotion (Gibson, 1958). For example, during linear motion, the focus of expansion in the flow field specifies current heading, which can be used as the basis of a *control law*: signed displacement of FOE from the desired locomotor target signals heading error, that needs to be compensated by steering in the opposite direction. It has been suggested that this is an attractive way of looking at the visual control of locomotion, as the information in the optic array can be directly connected to locomotor control by a simple control law without the need to consider complex sensorimotor transformations.

However, an organism clearly can only sample the optic flow field subject to constraints imposed by the movement and the orientation of its eyes. This creates a problem in that the focus of expansion may be “masked” by head and eye-rotation in the *retinal flow* (Raudies Mignolla & Neumann, 2012; Regan & Beverley, 1982). Retinal flow is the flow pattern in the retinal image. This proximal visual stimulus available to the brain is a combination of optic flow from locomotion and head and body rotations. Every time we move ourselves *or* our head *or* our eyes retinal flow occurs.

So the proximal stimulus for the visual system depends not only on changes in vantage point, but also rotation of the visual axes of the eyes relative to the locomotor frame because we do not always align the visual axis with the locomotor axis (e.g. when looking at an eccentric target one is not heading toward, or at the locomotor target on a curved path). Especially when looking at a waypoint on one’s future path during curvilinear motion the optic flow field itself is complex, and gaze rotation due to visual anticipation and optokinetic nystagmus creates a quite a complex relation between the optic flow field and retinal flow (Kim & Turvey, 1999; Li & Warren, 2000, 2002; Wann & Swapp, 2000; Authié & Mestre, 2012; Lappi, 2014).

Extraretinal efference copy of the oculomotor commands, and vestibular signals signifying accelerations relative to the external world, are available for the brain and potentially useful for disambiguate the components of the observed retinal flow pattern due to movement and due to eye-rotation. Yet, while accurately determining heading from optic flow stimuli is something humans can do (Lappe, Bremmer & vandenBerg, 1999; Warren & Hannon, 1988) and it can be in principle recovered (in some cases even without extraretinal information, Longuet-Higgins & Pradny, 1980), it seems like a costly solution to have to recover it from the pattern of retinal flow. Also, note that the original rationale for the recovery of optic flow is that it can be used to determine current heading “directly” – having to estimate optic flow through perhaps complex computations may seem backward (cf. discussion in Fajen & Warren, 2000; Wann & Land, 2000; Wann & Land, 2001).

Consideration of eye movements in the context of multiple frame of reference transformations suggests an alternative analysis, however. In active gaze visual strategies eye-position is controlled so that it structures the *retinal flow* in such a way that it becomes a more useful control signal for locomotion (Kim & Turvey, 1999; Wann & Swapp, 2000). Eye movements and locomotion are thus coupled into a

unified visual strategy for controlling heading “directly”, i.e. without the need to compute optic flow, current heading, or physical parameters such as path curvature or depth distance.

This possibility is based on a rather remarkable geometrical property of the retinal flow field - *if* it is under active gaze control. (The POF must remain fixed in the 3D scene, at the desired locomotor target; i.e. some model-external gaze control process is assumed to enable you to “look where you want to go”, see below). If the observer is on a linear or a circular trajectory which will take him to the fixated point, then all the visual flow lines will be straight, rather than curved. What is more, all those flow lines that fall on the observer's future path will be vertical (see Kim & Turvey, 1999 and supplement to Wann & Swapp, 2000, for mathematical details). On the other hand if the current trajectory is either too tightly curved or too straight, the flow lines will be curved in the opposite direction, specifying the required steering correction. Stated as a rule for control, the observer should: fixate a waypoint on one's future path, and, while maintaining fixation with an oculomotor pursuit, keep the visual flow lines approximately straight by steering into the direction of retinal flow line curvature. Although speculative (it is not known, for example, whether retinal-flow curvature in realistic settings would be detectable), there are some attractive aspects in this idea; not least of which is that eye movements are here part of the overall visuomotor *strategy*, not just a hindrance to veridical heading perception. They are a means for the visual system to find or actively *create* relevant invariants in the proximal visual input.

4.2. Feedback control (and internal representation revisited)

Control theoretical closed loop modeling has been extensively applied in the research on oculomotor circuits (reviews: Girard & Berthoz, 2005; Lisberger et al., 1987). The classical approach is that this reveals the workings of internal modules, i.e. OE specific circuits that are parallel and largely independent, and respond to eliciting stimuli isolated by the conventional oculomotor tasks in bottom-up, retinal-image-driven manner (see section 2.4.; for a history and critique see Steinman, 1986). But this “telephone wire” model of oculomotor circuitry is not by any means the only way to approach the role of *feedback* in gaze control.

A model of eye head and body control in saccadic gaze shifts recently put forward by Daye et al. (2014) illustrates one important advance over the classical picture. This model is based on control theory, with multiple feedback control loops for gaze, head and body orientation. The notable and counterintuitive feature of the model is that oculomotor control is *not* based on an eye-in-head signal that would be returned to the system and compared to desired eye-rotation to determine OE error. Instead, the feedback signals are combined with desired *gaze* (eye + head) and head displacement. This is, of course, in contrast with most of the traditional oculomotor control theory literature that aims to identify “controllers” for OE classes, or even their subtypes.

What is important here is that the analysis of eye movement patterns is not restricted to the traditional eye-in-head analysis: the control problem is framed in terms of control of *gaze*, not control of the eye as such. This takes the analysis towards the question of coordinating gaze (and the body) in 3D space, and the multiple coordinate system transformations this entails in conditions of free movement in complex scenes, typical of natural behavior. Full-fat locomotor models for the internal mechanisms of top-down gaze control in the wild remain to be developed, though.

The Wann & Swapp (2000) retinal flow model is an example of a another type of closed-loop control model that is not based on the OE circuit approach. It specifies control laws which *directly* relate action to stimulus features (Figure 7, dashed lines). These types of *vision-action models* (see Wann & Land, 2000) or online models (Zhao & Warren, 2015) are well suited to real-world studies of visual behavior because model testing and development can proceed without detailed knowledge of the underlying mechanisms producing the sensorimotor transformations. The downside, which the figure tries to convey, is that it does not describe how the brain encodes sensory information and integrates it with existing representations (Figure 7, solid lines; see discussion in section 3.3., above).

FIGURE 7 HERE

Not all approaches to motor control are based on internal models, though. The ecological approach to vision (Gibson, 1958, 1986), some radical interpretations of embodied cognition (Wilson & Golonka, 2013), and the online vs. model-based approach (Zhao & Warren, 2015) take different view: in these frameworks a good theory of visuomotor control describes how motor action is “attuned “ to sensory patterns which, under normal ecological constraints, specify environmental conditions that make that action adaptive. These theories do not construe the brain’s task as re-encoding the stimulus information into stable internal representations (and thus would not consider the lack of this as a shortcoming). So, from this ecological perspective, the retinal flow active gaze model is *not* in need of further elaboration and explanation in terms of internal processes. Generally, though, powerful yet efficient representation is considered an essential feature of brain function (Barlow, 2001; Gershman & Niv, 2010), and in Artificial Intelligence extraction of sufficiently abstract representations is seen as the key to understanding human-level performance in complex tasks (Dennett, 1984; Lecun, Bengio & Hinton, 2015; Mnih et al., 2015; Botvinick et al., 2016).

Feedback refers to passing the output of a system back as input. This is also called *closed loop control* as the feedback can be continuously compared to target value(s) of the controlled variable(s) to determine the next motor command, as opposed to *open loop control*, where the control commands are planned before execution. The rationale for internal models in motor control is that physiological sensory delays make closed loop control based on using raw sensory feedback not viable in rapid skilled action, and can lead to instability and oscillatory behavior (if simple feedback control is applied to a control task where the controlled state is changing rapidly relative to feedback delays).

If the current state (e.g. gaze target direction) were estimated solely based on sensory feedback, motor plans would be generated on partially outdated information. Because of this, until the 90’s prevailing view in the literature on motor control was that well-practiced and skilled action is generated as “preprogrammed” patterns of motor activation stored in the course of skill learning. That is, motor sequences were assumed to be generated by open loop motor programs that are set up based on preview information acquired before movement onset, but, once initiated by a sensory cue, the programs would their course independent of feedback (Keele, 1968). The modern approach to solve the feedback delay problem is to use internal models in feedback control: predictive forward models in the nervous system can produce an internal “feedback” signal before sensory feedback arrives, enabling zero-lag and even anticipatory (predictive) control (Miall & Wolpert, 1996; Wolpert & Kawato, 1998; for a historical overview see Desmurget & Grafton, 2000).

In modern computational movement neuroscience (Todorov, 2004; Wolpert & Kawato, 1998; Wolpert & Ghahramani, 2000; Wolpert, Diedrichsen & Flanagan, 2011) “internal feedback” created by predictive internal models is an essential source of control information. Such *forward models* are ubiquitous in the nervous system and have been extensively studied in at multiple levels, from low-level motor system models for feedback prediction in sensorimotor control (Miall & Wolpert, 1996) to cognitive maps of allocentric space (Barry & Burgess, 2014).

Forward models explain why you cannot tickle yourself: the stimulation is predictable from the motor plan and this signal affects neural processing, already at the level of the spinal cord (Blakemore, Wolpert & Frith, 1998). As discussed in section 2.2., when we turn our head and eyes (gaze saccade), efference copy inhibits the vestibulo-ocular reflex that would counter-rotate the eyes to compensate for head motion when angular acceleration is registered in the semicircular canals (Roy & Cullen, 1998, 2002). Indeed, the concept of forward models can be considered a generalization of the corollary discharge/efference copy theory (Sperry, 1950; vonHolst & Mittelstaedt, 1950) where the motor signal causing an active movement of the eye “cancels” the change in retinal image so that the image shift is not interpreted as a change in the scene whenever you move your eyes. In these examples, the simplest efference copy (corollary discharge of the motor command itself) is used to predict the simplest type of feedback (reafferent sensory stimulation) in a *sensory forward model*. At a higher level, *dynamical forward models* can maintain real-time, lag-free, estimates outcomes of more complex actions. They can use the

history of previously observed states (or re-estimates taking advantage of subsequent sensory information), current state estimate, efference copy of the current motor plan, and an internal representation that captures the dynamics of the environment. In this framework, maintaining a coherent spatial image during locomotion (Loomis et al., 1996; Loomis, Klatzky & Giudice, 2013; see section 3.3.) could be interpreted as high-level, multisensory dynamical forward models, updating the egocentric directions and distances of targets (and possibly allocentric self-location and orientation).

Recently, also *inverse models* have begun to receive increased attention (Spiers & Barry, 2015). Inverse models can be used to select motor commands from among a suite of alternatives in order to achieve the present goal, given the present state estimate of forward models. Identifying your current point of fixation as your locomotor goal in the blind walking task ("going where you are/were looking"), or conversely your locomotor target - such as a waypoint on your predicted future path - as a gaze target ("looking where you are going") would involve inverse models. At some level, the goals for oculomotor and locomotor inverse models might come from some higher decision process planning *both* gaze and body movements in a unified spatial representation (when we "look where we want to go", i.e. reorient gaze *and* locomotion in relation to predicted path information; cf. Hayhoe et al., 2012).

With a combination of forward and inverse models, a motor plan (as opposed to a rigid motor program) does not unfold completely ballistically, nor is it driven by sensory events (as opposed to an online closed-loop control law). The timing, amplitude and direction of actions involved are constantly fine tuned by internally generated feedback, and actions can be selected with reference to adaptively changing goals, rather than fixed set points. Understanding the role and prevalence in the feedback control of gaze of abstract internal models (forward and inverse), and their relation to online (active gaze) strategies, are essential open areas of research, and may help to integrate the ecological, neurophysiological and cognitive perspectives on naturalistic behavior.

5. Conclusions

We have looked at eye movement behavior in the wild from two complementary perspectives. One is to take what is known about basic oculomotor events and the neural processes controlling them, and to work out how they could be combined in more complex behaviors. The other is to start with qualitative principles of gaze control in naturalistic conditions, and work out how it could be composed out of more elementary patterns. Potential points of conflict, points of integration, and areas in need of theoretical and methodological development were diagnosed.

We emphasized how eye movement control "in the wild" needs to be considered in the context of overall movement control, because naturalistic oculomotor control is always embedded within eye-head and head - body control for controlling gaze, rather than a separate control problem. We evaluated critically the idea of using the eye movement types as indices of a "modular" neuromotor architecture, i.e. separate response systems for OE types as building blocks for complex behavior. Even in the laboratory context the OE classes are not so clear-cut and in complex real-world gaze behavior depends on multisensory spatial representations, and the multiple frames of reference and mean that even elementary classifications - such as what a "fixation" is - become underspecified. We also raised the conceptual distinction between 3D frames of reference - needed to accurately represent coordination of real-world gaze behavior, i.e. *what* the brain's control task is - and *how* internal coordinate transformations are actually represented in the brain.

In addition to points of divergence, two major themes of potential converge were discussed. *Active gaze* structuring the retinal input to facilitate efficient perceptual and/or motor processing is a principle that has been applied in both sedentary lab settings and active locomotion. *Feedback control* is another fundamental feature of gaze control. Here we emphasized that instead of "OE modules" responding in a reflex-like way to retinal events when interpreted, feedback control can be analyzed from the point of view of gaze control embedded into overall motor control, and driven in part by multisensory representation of space.

Tightly controlled laboratory experiments are needed to reveal the detailed workings of oculomotor mechanisms, but the level detail comes at the cost of restricting the behavioral context to much simplified tasks. And sometimes the connection to real-world behavior can be a tenuous one, even relying on suggestive analogies between the task settings and some parameters of a natural task only outlined in an introductory vignette. Experiments in naturalistic settings are necessary to uncover the cues and gaze strategies people actually use in coping with the complexity and ambiguity of real-world tasks - but the complexity of the tasks and stimuli and lack of experimental control mean underlying causal mechanisms are often not conclusively established. Ideally, the bottom-up approach and top-down approach should meet in the middle, as this will increase our understanding how gaze control fits in the context of general motor control in locomotor tasks such as running, driving, dancing, or sports, and, indeed, most everyday tasks such as cooking, infant care, tool use and many forms of social interaction and cooperation. After all, in the natural ecological context of human and animal vision, purely sedentary tasks are the exception, not the rule.

--- INSERT TABLE 3 at the very end of the manuscript, for easy reference ---

References

- Ahissar, E. & Arieli, A. (2012). Seeing via miniature eye movements: a dynamic hypothesis for vision. *Frontiers in Computational Neuroscience*, 6, 1–27.
- Ahissar, E., Arieli, A., Fried, M., Bonneh, Y. (2016) On the possible roles of microsaccades and drifts in visual perception, *Vision Research*, 118, 25-30.
- Andersen, R. A., Snyder, L. H., Li, C. S., & Stricanne, B. (1993). Coordinate transformations in the representation of spatial information. *Current opinion in neurobiology*, 3(2), 171-176.
- Angelaki, D. E., & Hess, B. J. (2005). Self-motion-induced eye movements: effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6(12), 966-976.
- Authié, C. N., Hilt, P. M., N_guyen, S., Berthoz, A., & Bennequin, D. (2015). Differences in gaze anticipation for locomotion with and without vision. *Frontiers in Human Neuroscience*, 9, 312.
- Authié, C. N., & Mestre, D. R. (2012). Path curvature discrimination: dependence on gaze direction and optical flow speed. *PloS One*, 7(2), e31479.
- Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, 24(3), 191-204.
- Ballard, D.H. (1991). *Animate Vision*. *Artificial Intelligence*, 48, 57–86.
- Ballard, D., Hayhoe, M., & Pelz, J. (1995). Memory representations in natural tasks, *Journal of Cognitive Neuroscience*, 7(1), 66-80.
- Barlow, H. (2001). Redundancy reduction revisited. *Network: Computation in neural systems*, 12(3), 241-53.
- Barnes, G. R. (2008). Cognitive processes involved in smooth pursuit eye movements. *Brain and cognition*, 68(3), 309-326.
- Barnes, G. R., & Asselman, P. T. (1991). The mechanism of prediction in human smooth pursuit eye movements. *The Journal of Physiology*, 439(1), 439-461.
- Barry, C., & Burgess, N. (2014). Neural mechanisms of self-location. *Current Biology*, 24(8), R330-R339.
- Batista, A. P., Buneo, C. A., Snyder, L. H., & Andersen, R. A. (1999). Reach plans in eye-centered coordinates. *Science*, 285(5425), 257-260.
- Becker, W., & Fuchs, A. F. (1969). Further properties of the human saccadic system: eye movements and correction saccades with and without visual fixation points. *Vision research*, 9(10), 1247-1258.
- Bennett, S. J., & Barnes, G. R. (2003). Human ocular pursuit during the transient disappearance of a visual target. *Journal of Neurophysiology*, 90(4), 2504-2520.
- Bennett, S. J., & Barnes, G. R. (2006). Combined smooth and saccadic ocular pursuit during the transient occlusion of a moving visual object. *Experimental brain research*, 168(3), 313-321.
- Bernstein, N. A. (1967). *The co-ordination and regulation of movements*. New York: Pergamon Press.
- Bizzi, E., Mussa-Ivaldi, F. A., & Giszter, S. (1991). Computations underlying the execution of movement: a biological perspective. *Science*, 253(5017), 287-291.
- Blakemore, S., Wolpert, D.M. & Frith, C.D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, 1, 635–640.
- Bremmer, F. (2011). Multisensory space: from eye-movements to self-motion. *The Journal of physiology*, 589(4), 815-823.
- Brown, T. G. (1911). The intrinsic factors in the act of progression in the mammal. *Proceedings of the Royal Society of London. Series B, containing papers of a biological character*, 308-319.
- Botvinick, M., Weinstein, A., Solway, A., & Barto, A. (2015). Reinforcement learning, efficient coding, and the statistics of natural tasks. *Current Opinion in Behavioral Sciences*, 5, 71-77.
- Busettini, C., Miles, F. A., Schwarz, U., & Carl, J. R. (1994). Human ocular responses to translation of the observer and of the scene: dependence on viewing distance. *Experimental Brain Research*, 100(3), 484-494.
- Cesqui, B., van de Langenberg, R., Lacquaniti, F., & d'Avella, A. (2013). A novel method for measuring gaze orientation in space in unrestrained head conditions. *Journal of vision*, 13(8), 28-28.

- Chattington, M., Wilson, M., Ashford, D., & Marple-Horvat, D. E. (2007). Eye-steering coordination in natural driving. *Experimental brain research*, 180(1), 1-14.
- Cohen, Y. E., & Andersen, R. A. (2002). A common reference frame for movement plans in the posterior parietal cortex. *Nature Reviews Neuroscience*, 3(7), 553-562.
- Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual review of neuroscience*, 22(1), 319-349.
- Collewyn, H., & Kowler, E. (2008). The significance of microsaccades for vision and oculomotor control. *Journal of Vision*, 8(14), 20.
- Collewyn, H., Steinman, R. M., Erkelens, C. J., Pizlo, Z., & van der Steen, J. (1992). Effect of freeing the head on eye movement characteristics during three-dimensional shifts of gaze and tracking. The head-neck sensory motor system. Oxford University Press, Oxford, 412-418.
- Collewyn, H., Steinman, R. M., Erkelens, C. J., Pizlo, Z., Kowler, E., & Van der Steen, J. (1992). Binocular gaze control under free-head conditions.
- Crane, B. T., Viirre, E. S., & Demer, J. L. (1997). The human horizontal vestibulo-ocular reflex during combined linear and angular acceleration. *Experimental Brain Research*, 114(2), 304-320.
- Crawford, J. D., Henriques, D. Y., & Medendorp, W. P. (2011). Three-dimensional transformations for goal-directed action. *Annual review of neuroscience*, 34, 309-331.
- d'Avella, A., Saltiel, P., & Bizzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior. *Nature neuroscience*, 6(3), 300-308.
- Daye, P. M., Optican, L. M., Blohm, G., & Lefevre, P. (2014). Hierarchical control of two-dimensional gaze saccades. *Journal of computational neuroscience*, 36(3), 355-382.
- Dennett, D.C. (1984). Cognitive wheels: the frame problem of AI. In: C. Hookway (ed.): *Minds, Machines and Evolution*, Cambridge University Press, pp. 129-150.
- Desmurget, M. & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, 4(11), 423-431.
- Diaz, G., Cooper, J., Kit, D., & Hayhoe, M. (2013). Real-time recording and classification of eye movements in an immersive virtual environment. *Journal of vision*, 13(12), 1-14.
- Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature*, 170, 36-37.
- Ditchburn, R. W. (1980). The function of small saccades. *Vision Research*, 20(3), 271-272.
- Donner, K. & Hemilä, S. (2007). Modelling the effect of microsaccades on retinal responses to stationary contrast patterns, *Vision Research*, 47(9), 1166-1177.
- Duchowski, A., Medlin, E., Cournia, N., Murphy, H., Gramopadhye, A., Nair, S., ... & Melloy, B. (2002). 3-D eye movement analysis. *Behavior Research Methods, Instruments, & Computers*, 34(4), 573-591
- Easton, T. A. (1972). On the normal use of reflexes: The hypothesis that reflexes form the basic language of the motor program permits simple, flexible specifications of voluntary movements and allows fruitful speculation. *American Scientist*, 591-599.
- Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. *Progress in brain research*, 154, 177-192.
- Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences*, 103(18), 7192-7197.
- Epelboim, J., Steinman, R. M., Kowler, E., Edwards, M., Pizlo, Z., Erkelens, C. J., & Collewyn, H. (1995). The function of visual search and memory in sequential looking tasks. *Vision research*, 35(23), 3401-3422.
- Epelboim, J., Steinman, R. M., Kowler, E., Pizlo, Z., Erkelens, C. J., & Collewyn, H. (1997). Gaze-shift dynamics in two kinds of sequential looking tasks. *Vision research*, 37(18), 2597-2607.
- Epelboim, J. (1998). Gaze and retinal-image-stability in two kinds of sequential looking tasks. *Vision research*, 38(23), 3773-3784.
- Fajen, B. R., & Warren Jr, W. H. (2000). Go with the flow. *Trends in Cognitive Sciences*, 4(10), 369-370.
- Fermüller, C. & Aloimonos, Y. (1995). Vision and action. *Image and Vision Computing*, 13(10), 725-744.
- Fukushima, S. S., Loomis, J. M., & Da Silva, J. A. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1), 86-100.

- Furneaux, S. & Land, M.F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proceedings of the Royal Society of London B*, 266, 2435-2440.
- Gaarder, K. (1966). Transmission of edge information in the human visual system. *Nature*, 212, 321-323.
- Gellman, R. S., Carl, J. R., & Miles, F. A. (1990). Short latency ocular-following responses in man. *Visual neuroscience*, 5(02), 107-122.
- Gershman, S. J., & Niv, Y. (2010). Learning latent structure: carving nature at its joints. *Current opinion in neurobiology*, 20(2), 251-256.
- Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British journal of psychology*, 49(3), 182-194.
- Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Girard, B. & Berthoz, A. (2005). From brainstem to cortex: Computational models of saccade generation circuitry. *Progress in neurobiology*, 77, 215-251.
- Goodale, M. A. (2011). Transforming vision into action. *Vision research*, 51(13), 1567-1587.
- Grasso, R., Glasauer, S., Takei, Y., & Berthoz, A. (1996). The predictive brain: anticipatory control of head direction for the steering of locomotion. *Neuroreport*, 7(6), 1170-1174.
- Grasso, R., Prévost, P., Ivanenko, Y. P., & Berthoz, A. (1998). Eye-head coordination for the steering of locomotion in humans: an anticipatory synergy. *Neuroscience letters*, 253(2), 115-118.
- Green, A. M., & Angelaki, D. E. (2010). Multisensory integration: resolving sensory ambiguities to build novel representations. *Current opinion in neurobiology*, 20(3), 353-360.
- Grillner, S. (2006). Biological pattern generation: the cellular and computational logic of networks in motion. *Neuron*, 52(5), 751-766.
- Grillner, S., & Wallén, P. (1985). Central pattern generators for locomotion, with special reference to vertebrates. *Annual review of neuroscience*, 8(1), 233-261.
- Hafed, Z. M., Goffart, L., & Krauzlis, R. J. (2009). A neural mechanism for microsaccade generation in the primate superior colliculus. *science*, 323(5916), 940-943.
- Hafed, Z. M., & Krauzlis, R. J. (2012). Similarity of superior colliculus involvement in microsaccade and saccade generation. *Journal of neurophysiology*, 107(7), 1904-1916.
- Harrison, J.J., Freeman, T.C. A. & Sumner, P. (2014). Saccade-like behavior in the fast-phases of optokinetic nystagmus: An illustration of the emergence of volitional actions from automatic reflexes. *Journal of Experimental Psychology: General*, 143(5), 1923-1938.
- Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in cognitive sciences*, 9(4), 188-194.
- Hayhoe, M. M., McKinney, T., Chajka, K., & Pelz, J. B. (2012). Predictive eye movements in natural vision. *Experimental Brain Research*, 217(1), 125-136.
- Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of vision*, 3(1), 49-63.
- Herdman, S. J., Schubert, M. C., & Tusa, R. J. (2001). Role of central preprogramming in dynamic visual acuity with vestibular loss. *Archives of Otolaryngology-Head & Neck Surgery*, 127(10), 1205-1210.
- Herst, A. N., Epelboim, J., & Steinman, R. M. (2001). Temporal coordination of the human head and eye during a natural sequential tapping task. *Vision research*, 41(25), 3307-3319.
- Hine, T., & Thorn, F. (1987). Compensatory eye movements during active head rotation for near targets: effects of imagination, rapid head oscillation and vergence. *Vision research*, 27(9), 1639-1657.
- Ibbotson, M., & Krekelberg, B. (2011). Visual perception and saccadic eye movements. *Current opinion in neurobiology*, 21(4), 553-558.
- Ilg, U. J. (1997). Slow eye movements. *Progress in neurobiology*, 53(3), 293-329.
- Imai, T., Moore, S. T., Raphan, T., & Cohen, B. (2001). Interaction of the body, head, and eyes during walking and turning. *Experimental brain research*, 136(1), 1-18.
- Kagan, I. (2012). Active vision: fixational eye movements help seeing space in time. *Current Biology*, 22(6), R186-R188.

- Kaminiaz, A., Königs, K., & Bremmer, F. (2009). Task influences on the dynamic properties of fast eye movements. *Journal of vision*, 9(13), 1-11.
- Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin*, 70 (6p1), 387-403.
- Kim, N. G., & Turvey, M. T. (1999). Eye movements and a rule for perceiving direction of heading. *Ecological Psychology*, 11(3), 233-248.
- Kinsman, T., Evans, K., Sweeney, G., Keane, T., & Pelz, J. (2012). Ego-motion compensation improves fixation detection in wearable eye tracking. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 221-224). ACM.
- Ko, H. K., Poletti, M., & Rucci, M. (2010). Microsaccades precisely relocate gaze in a high visual acuity task. *Nature neuroscience*, 13(12), 1549-1553.
- Kowler, E., & Steinman, R. M. (1979). The effect of expectations on slow oculomotor control—I. Periodic target steps. *Vision research*, 19(6), 619-632.
- Kowler, E., & Steinman, R. M. (1980). Small saccades serve no useful purpose: reply to a letter by RW Ditchburn. *Vision research*, 20(3), 273-276.
- Kowler, E. (1989). Cognitive expectations, not habits, control anticipatory smooth oculomotor pursuit. *Vision research*, 29(9), 1049-1057.
- Kowler, E. (2011). Eye movements: The past 25 years. *Vision research*, 51(13), 1457-1483.
- Krauzlis, R. J. (2004). Recasting the smooth pursuit eye movement system. *Journal of neurophysiology*, 91(2), 591-603.
- Kuang X, Poletti M, Victor JD, Rucci M (2012) Temporal encoding of spatial information during active visual fixation. *Cur Biol* 22(6): 510- 514.
- Lanata, A., Greco, A., Valenza, G., & Scilingo, E. P. (2015) Robust Head Mounted Wearable Eye Tracking System for Dynamical Calibration. *Journal of eye movement research*, 8(5):2, 1-15.
- Land, M. F. (1992). Predictable eye-head coordination during driving. *Nature*, 359, 318-320.
- Land, M.F. (2004). The coordination of rotations of the eyes, head and trunk in saccadic turns produced in natural situations. *Experimental Brain Research*, 159, 151-160.
- Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in retinal and eye research*, 25(3), 296-324.
- Land, M.F. & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society B*, 352: 1231-1239.
- Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature neuroscience*, 3(12), 1340-1345.
- Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception-London*, 28(11), 1311-1328.
- Lappe, M., Bremmer, F., & Van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in cognitive sciences*, 3(9), 329-336.
- Lappi, O., & Lehtonen, E. (2013). Eye-movements in real curve driving: pursuit-like optokinesis in vehicle frame of reference, stability in an allocentric-reference coordinate system. *Journal of Eye Movement Studies*, 6(1):4, 1-13.
- Lappi, O., Pekkanen, J., & Itkonen, T. H. (2013). Pursuit eye-movements in curve driving differentiate between future path and tangent point models. *PLOS ONE*, 8(7), e68326.
- Lappi, O. (2014). Future path and tangent point models in the visual control of locomotion in curve driving. *Journal of vision*, 14(12), 1-22.
- Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112-131). New York: Wiley.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
- Lehtonen, E., Lappi, O., Kotkanen, H. & Summala, H. (2013). Look-ahead fixations in curve driving. *Ergonomics* 56(1), 34-44.

- Lee, W. A. (1984). Neuromotor synergies as a basis for coordinated intentional action. *Journal of motor behavior*, 16(2), 135-170.
- Li, L., & Warren, W. H. (2000). Perception of heading during rotation: sufficiency of dense motion parallax and reference objects. *Vision research*, 40(28), 3873-3894.
- Li, L., & Warren, W. H. (2002). Retinal flow is sufficient for steering during observer rotation. *Psychological Science*, 13(5), 485-490.
- Lisberger, S.G., Evinger, C., Johanson, G. & Fuchs, A.F. (1981). Relationship between eye acceleration and retinal image velocity during foveal smooth pursuit eye movements in man and monkey. *Journal of Neurophysiology*, 46, 229-249.
- Lisberger, S.G., Morris, E.J. & Tychsen, L. (1987). Visual motion processing and sensory-motor integration from smooth pursuit eye movements. *Annual Review of Neuroscience*, 10, 97-129.
- Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 906-921.
- Loomis, J. M., Silva, J. A. D., Philbeck, J. W., & Fukusima, S. S. (1996). Visual perception of location and distance. *Current Directions in Psychological Science*, 72-77.
- Loomis, J. M., Klatzky, R. L., & Giudice, N. A. (2013). Representing 3D space in working memory: Spatial images from vision, hearing, touch, and language. In *Multisensory imagery* (pp. 131-155). Springer New York.
- Malinov, I. V., Epelboim, J., Herst, A. N., & Steinman, R. M. (2000). Characteristics of saccades and vergence in two kinds of sequential looking tasks. *Vision research*, 40(16), 2083-2090.
- Mann, D. L., Spratford, W., & Abernethy, B. (2013). The head tracks and gaze predicts: how the world's best batters hit a ball. *PloS one*, 8(3), e58289.
- Marder, E., & Bucher, D. (2001). Central pattern generators and the control of rhythmic movements. *Current biology*, 11(23), R986-R996.
- Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, 5(3), 229-240.
- Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: a neurophysiological analysis. *Trends in neurosciences*, 32(9), 463-475.
- Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, 14(2), 83-96.
- McGuire, L. M., & Sabes, P. N. (2009). Sensory transformations and the use of multiple reference frames for reach planning. *Nature neuroscience*, 12(8), 1056-1061.
- Mennie, N., Hayhoe, M., & Sullivan, B. (2007). Look-ahead fixations: anticipatory eye movements in natural tasks. *Experimental Brain Research*, 179(3), 427-442.
- Meyer, C.H., Lasker, A.G. & Robinson, D.A. (1985). The upper limit of human smooth pursuit velocity. *Vision Research*, 25(4), 561-563.
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural networks*, 9(8), 1265-1279.
- Miles, F. A. (1997). Visual stabilization of the eyes in primates. *Current opinion in neurobiology*, 7(6), 867-871.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529-533.
- Mostofi, N., Boi, M., & Rucci, M. (2015). Are the visual transients from microsaccades helpful? Measuring the influences of small saccades on contrast sensitivity. *Vision research*.
- Munn, S. M., & Pelz, J. B. (2009). FixTag: An algorithm for identifying and tagging fixations to simplify the analysis of data collected by portable eye trackers. *ACM Transactions on Applied Perception (TAP)*, 6(3), 16.
- Munn, S. M., Stefano, L., & Pelz, J. B. (2008). Fixation-identification in dynamic scenes: Comparing an automated algorithm to manual coding. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization* (pp. 33-42). ACM.

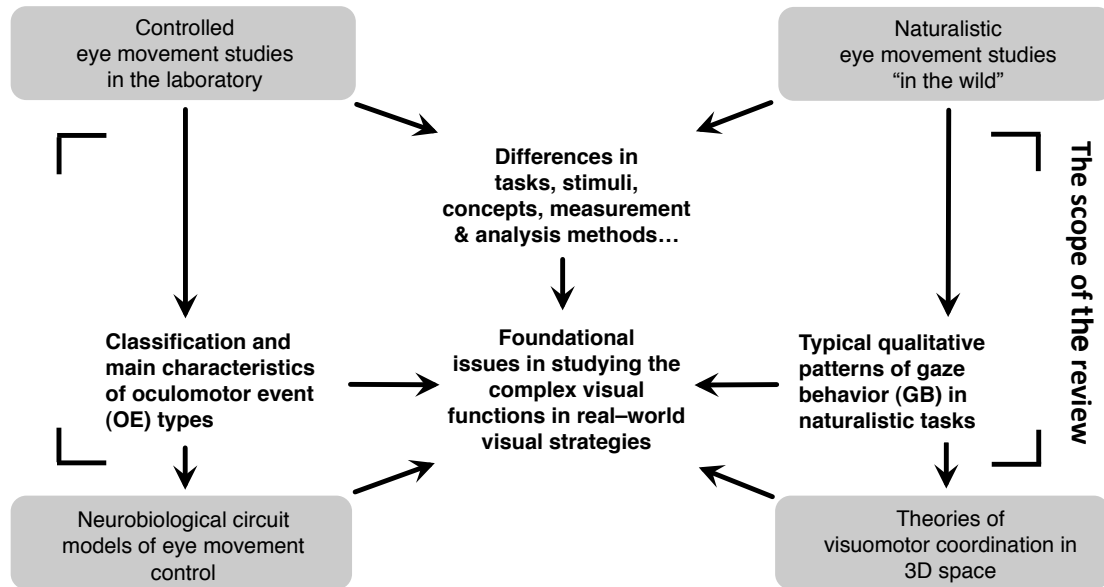
- Munoz, D. P. (2002). Commentary: saccadic eye movements: overview of neural circuitry. *Progress in brain research*, 140, 89-96.
- Murphy, B. J., Kowler, E., & Steinman, R. M. (1975). Slow oculomotor control in the presence of moving backgrounds. *Vision research*, 15(11), 1263-1268.
- Mussa-Ivaldi, F. A., & Bizzi, E. (2000). Motor learning through the combination of primitives. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 355(1404), 1755-1769.
- Mussa-Ivaldi, F. A., Giszter, S. F., & Bizzi, E. (1994). Linear combinations of primitives in vertebrate motor control. *Proceedings of the National Academy of Sciences*, 91(16), 7534-7538.
- Neisser, U. (1976). *Cognition and reality*. San Francisco: W.H. Freeman.
- Newell, K. M. (2003). Schema theory (1975): Retrospectives and prospectives. *Research quarterly for exercise and sport*, 74(4), 383-388.
- Nyström, M., Hansen, D. W., Andersson, R., & Hooge, I. (2016). Why have microsaccades become larger? Investigating eye deformations and detection algorithms. *Vision Research*, 118, 17-24.
- O'Keefe, J. & Nadel, L. (1978). *The Hippocampus as a Spatial Map*. Oxford: Oxford University Press.
- Orban de Xivry, J. J., & Lefevre, P. (2007). Saccades and pursuit: two outcomes of a single sensorimotor process. *The Journal of Physiology*, 584(1), 11-23.
- Orban de Xivry, J. J., Missal, M., & Lefèvre, P. (2008). A dynamic representation of target motion drives predictive smooth pursuit during target blanking. *Journal of Vision*, 8(15), 1-13.
- Orban de Xivry, J. J., Coppe, S., Blohm, G., & Lefevre, P. (2013). Kalman filtering naturally accounts for visually guided and predictive smooth pursuit dynamics. *The Journal of Neuroscience*, 33(44), 17301-17313.
- Otero-Millan, J., Troncoso, X. G., Macknik, S. L., Serrano-Pedraza, I., & Martinez-Conde, S. (2008). Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator. *Journal of Vision*, 8(14), 1-18.
- Otero-Millan, J., Macknik, S. L., Langston, R. E., & Martinez-Conde, S. (2013). An oculomotor continuum from exploration to fixation. *Proceedings of the National Academy of Sciences*, 110(15), 6175-6180.
- Park, S. & Chun, M. (2009). Different roles of the parahippocampal place area (PPA) and retrosplenial cortex (RSC) in panoramic scene perception. *NeuroImage*, 47, 1747-1756.
- Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Oxford: Oxford University Press.
- Pearson, K. G. (2000). Neural adaptation in the generation of rhythmic behavior. *Annual review of physiology*, 62(1), 723-753.
- Pelz, J.B. & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research*, 41, 3587-3596.
- Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 781-795.
- Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Current opinion in neurology*, 17(1), 17-25.
- Poletti, M., Listorti, C., & Rucci, M. (2013). Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Current Biology*, 23(17), 1691-1695.
- Poletti, M., Burr, D. C., & Rucci, M. (2013). Optimal multimodal integration in spatial localization. *The Journal of Neuroscience*, 33(35), 14259-14268.
- Poletti, M., & Rucci, M. (2016). A compact field guide to the study of microsaccades: challenges and functions. *Vision research*, 118, 83-97.
- Poletti, M., Listorti, C., & Rucci, M. (2013). Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Current Biology*, 23(17), 1691-1695.
- Pouget, A., & Sejnowski, T. (1997). Spatial transformations in the parietal cortex using basis functions. *Cognitive Neuroscience, Journal of*, 9(2), 222-237.
- Pouget, A., Deneve, S., & Duhamel, J. R. (2002). A computational perspective on the neural basis of multisensory spatial representations. *Nature Reviews Neuroscience*, 3(9), 741-747.

- Prévost, P., Yuri, I., Renato, G., & Alain, B. (2003). Spatial invariance in anticipatory orienting behaviour during human navigation. *Neuroscience letters*, 339(3), 243-247.
- Putnam, N. M., Hofer, H. J., Doble, N., Chen, L., Carroll, J., & Williams, D. R. (2005). The locus of fixation and the foveal cone mosaic. *Journal of Vision*, 5(7), 632-639.
- Raudies, F., Mingolla, E., & Neumann, H. (2012). Active gaze control improves optic flow-based segmentation and steering. *PloS one*, 7(6), e38446.
- Regan, D., & Beverley, K. I. (1982). How do we avoid confounding the direction we are looking and the direction we are moving?. *Science*, 215(4529), 194-196.
- Regan, D., & Gray, R. (2000). Visually guided collision avoidance and collision achievement. *Trends in Cognitive Sciences*, 4(3), 99-107.
- Reimer, B., & Sodhi, M. (2006). Detecting eye movements in dynamic environments. *Behavior Research Methods*, 38(4), 667-682.
- Rieser, J. J., Ashmead, D. H., Talor, C. R., & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, 19(5), 675-689.
- Rolfs, M. (2009). Microsaccades: small steps on a long way. *Vision research*, 49(20), 2415-2441.
- Robinson, D. A. (1964). The mechanics of human saccadic eye movement. *The Journal of physiology*, 174(2), 245-264.
- Roy, J.E. & Cullen, K.E. (1998). A neural correlate for vestibulo-ocular reflex suppression during voluntary eye-head gaze shifts. *Nature Neuroscience*, 1, 404-410.
- Roy, J. E., & Cullen, K. E. (2001). Selective processing of vestibular reafference during self-generated head motion. *The Journal of Neuroscience*, 21(6), 2131-2142.
- Roy, J.E. & Cullen, K.E. (2002). Vestibuloocular reflex signal modulation during voluntary and passive head movements. *Journal of Neurophysiology*, 87, 2337-2357.
- Rucci, M. & Casile, A. (2005). Fixational instability and natural image statistics: Implications for early visual representations. *Network: Computation in Neural Systems*, 16(2/3), 121-138.
- Rucci, M. & Victor, J.D. (2015). The unsteady eye: an information-processing stage, not a bug, *Trends in Neurosciences*, 38(4), 195-206.
- Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological review*, 82(4), 225-260.
- Schmidt, R. A. (2003). Motor schema theory after 27 years: reflections and implications for a new theory. *Research quarterly for exercise and sport*, 74(4), 366-375.
- Scudder, C. A., Kaneko, C. R., & Fuchs, A. F. (2002). The brainstem burst generator for saccadic eye movements. *Experimental Brain Research*, 142(4), 439-462.
- Shea, C. H., & Wulf, G. (2005). Schema theory: A critical appraisal and reevaluation. *Journal of motor behavior*, 37(2), 85-102.
- Sherrington, C.S. (1906) *The integrative action of the nervous system*. New Haven, CT: Yale University Press.
- Sinn, P. & Engbert, R. (2016). Small saccades versus microsaccades: Experimental distinction and model-based unification, *Vision Research*, 118, 132-143.
- Sparks, D. L. (2002). The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience*, 3(12), 952-964.
- Sperry, R.W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, 43, 482-489.
- Spiers, H. J., & Barry, C. (2015). Neural systems supporting navigation. *Current opinion in behavioral sciences*, 1, 47-55.
- Steinman, R. M. (1986). The need for an eclectic, rather than systems, approach to the study of the primate oculomotor system. *Vision Research*, 26(1), 101-112.
- Steinman, R.M., Cushman, W.B. & Martins, A.J. (1982). The Precision of Gaze. *Human Neurobiology*, 1, 97-109.
- Steinman, R. M., Haddad, G. M., Skavenski, A. A., & Wyman, D. (1973). Miniature eye movement. *Science*, 181(4102), 810-819.

- Summers, J. J., & Anson, J. G. (2009). Current status of the motor program: Revisited. *Human Movement Science*, 28(5), 566-577.
- Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of vision*, 11(5), 1-23.
- Tatler, B. W., & Land, M. F. (2011). Vision and the representation of the surroundings in spatial memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1564), 596-610.
- Thomson, J. A. (1983). Is continuous visual monitoring necessary in visually guided locomotion?. *Journal of Experimental Psychology: Human Perception and Performance*, 9(3), 427.
- Thiele, A., Henning, P., Kubischik, M., & Hoffmann, K. P. (2002). Neural mechanisms of saccadic suppression. *Science*, 295(5564), 2460-2462.
- Thier, P., & Ilg, U. J. (2005). The neural basis of smooth-pursuit eye movements. *Current opinion in neurobiology*, 15(6), 645-652.
- Ting, L. H., & McKay, J. L. (2007). Neuromechanics of muscle synergies for posture and movement. *Current opinion in neurobiology*, 17(6), 622-628.
- Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature neuroscience*, 7(9), 907-915.
- Tomasi, M., Pundlik, S., Bowers, A.R., Peli, E., Luo, G. (2016). Mobile gaze tracking system for outdoor walking behavioral studies. *Journal of Vision*, 16(3), 27, 1-15.
- Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological review*, 113(4), 766-786.
- Turvey, M. T. (2007). Action and perception at the level of synergies. *Human movement science*, 26(4), 657-697.
- Valsecchi, M., & Gegenfurtner, K. R. (2015). Control of binocular gaze in a high-precision manual task. *Vision research*, 110 (Part B), 203-214
- Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, 10(11), 792-802.
- Vansteenkiste, P., Cardon, G., D'Hondt, E., Philippaerts, R., & Lenoir, M. (2013). The visual control of bicycle steering: The effects of speed and path width. *Accident Analysis & Prevention*, 51, 222-227.
- Vidal, M., Bulling, A., & Gellersen, H. (2011). Analysing EOG signal features for the discrimination of eye movements with wearable devices. In *Proceedings of the 1st international workshop on pervasive eye tracking & mobile eye-based interaction* (pp. 15-20). ACM.
- Viirre, E. S., & Demer, J. L. (1996). The human vertical vestibulo-ocular reflex during combined linear and angular acceleration with near-target fixation. *Experimental brain research*, 112(2), 313-324.
- Von Holst, E. & Mittelstaedt, H. (1950). Das reafferenzprinzip. *Naturwissenschaften*, 37, 464-476.
- Wann, J., & Land, M. (2000). Steering with or without the flow: is the retrieval of heading necessary?. *Trends in cognitive sciences*, 4(8), 319-324.
- Wann, J., & Land, M. (2001). Heading in the wrong direction: Reply to Fajen and Warren. *Trends in Cognitive Sciences*, 5(1), 8-9.
- Wann, J. P., & Swapp, D. K. (2000). Why you should look where you are going. *Nature neuroscience*, 3(7), 647-648.
- Warren, W. H., & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature*, 336(6195), 162-163.
- Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, 4(58), 1-13.
- Winterson, B. J., & Collewyn, H. (1976). Microsaccades during finely guided visuomotor tasks. *Vision research*, 16(12), 1387-1390.
- Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 12(12), 739-751.
- Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *nature neuroscience*, 3, 1212-1217.

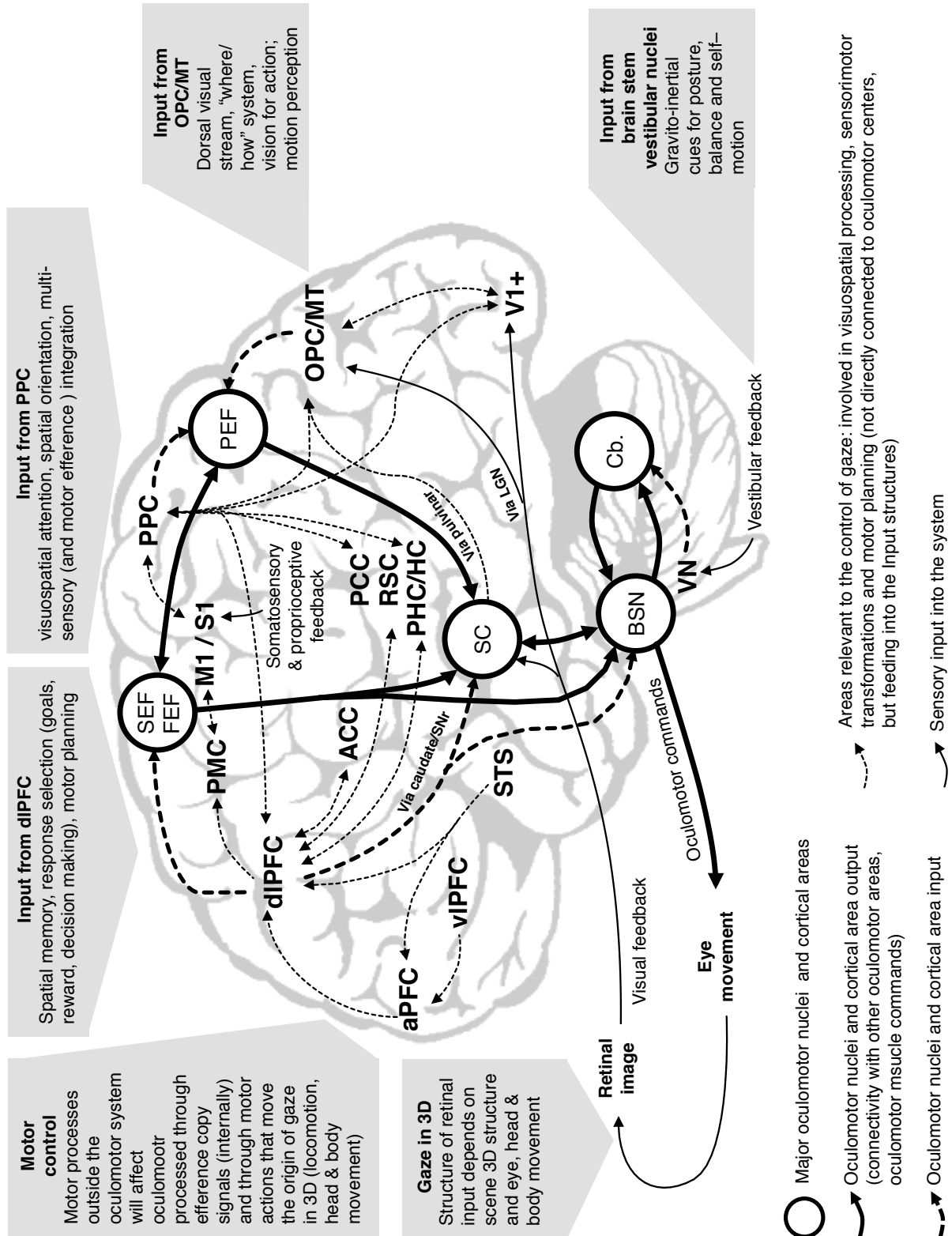
- Wolpert, D. M. & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11, 1317–1329
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision research*, 48(20), 2070-2089.
- Wyatt, H. J., & Pola, J. (1988). Predictive behavior of optokinetic eye movements. *Experimental brain research*, 73(3), 615-626.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum press.
- Zhao, H. & Warren, W. H. (2015). On-line and model-based approaches to the visual control of action. *Vision Research*, 110 (Part B), 190–202.

Figure 1.



Outline of the topics and scope of the review.

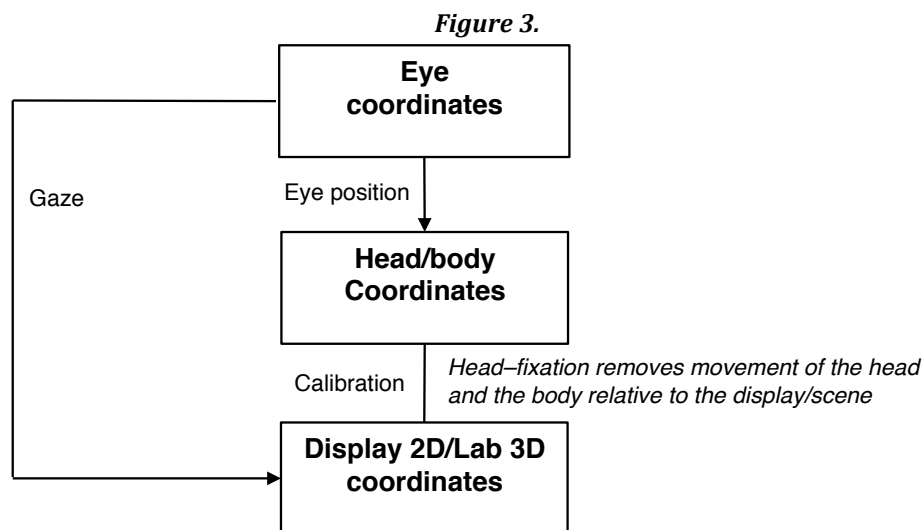
Figure 2.



Some of the main pathways involved in oculomotor control. The major nodes in the network (solid lines) are supplementary and frontal eye fields (SEF/FEF) and parietal eye fields (PEF) in the posterior intraparietal sulcus, superior colliculus (SC) and cerebellum (Cb.). Brain stem nuclei (BSN) in the pons mediate between cerebellar and cerebral system, and reticular motor nuclei represent the final efferent motor pathway. Major inputs (thick dashed lines) come from dorsolateral prefrontal cortex (dlPFC; BA9/46) and posterior parietal cortex (PPC; BA7/40) and the dorsal visual stream (occipitoparietal and

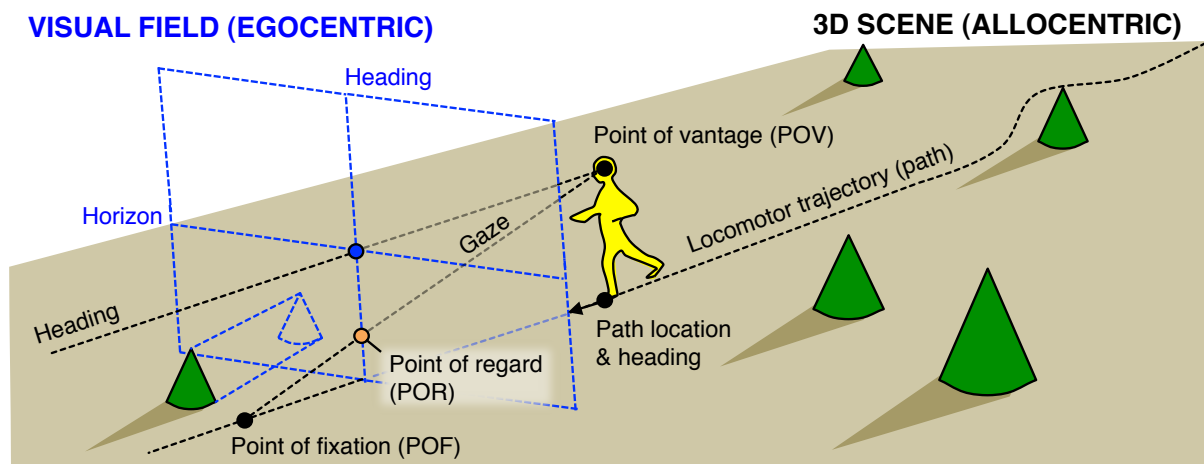
(Figure 2 Caption continued)

middle temporal, OPC/MT; BA19/39). Importantly, these systems are all involved multisensory integration and the representation of 3D space. Other areas important in representation of spatial information – not only for eye-movements but also for posture, balance, navigation and spatial perception, attention and memory – include hippocampus (HC), parahippocampal cortex (PHC; BA36), anterior cingulate cortex (ACC; BA24) and posterior cingulate/retrosplenial cortex (PCC; BA23, RSC; BA29/30). The detailed circuitry is much more complex, with more fine-grained differentiation of connectivity between sub-areas. V1+: primary visual areas (BA17/18). LGN: lateral geniculate nucleus. VN: vestibular nuclei. SNr: substantia nigra pars reticulata. vLPFC: ventrolateral prefrontal cortex (BA44/45). aPFC: anterior prefrontal cortex (BA10), PMC: premotor cortex (BA6/8) M1/S1: precentral and postcentral gyri (BA4 and BA1/2/3). The figure is based on the discussion of connectivity in the reviews by Munoz (2004), Krauzlis (2004), Pierrot-Deseilligny et al. (2004), Goodale (2011), Petrides (2005) and Vann, Aggleton, & Maguire (2009).



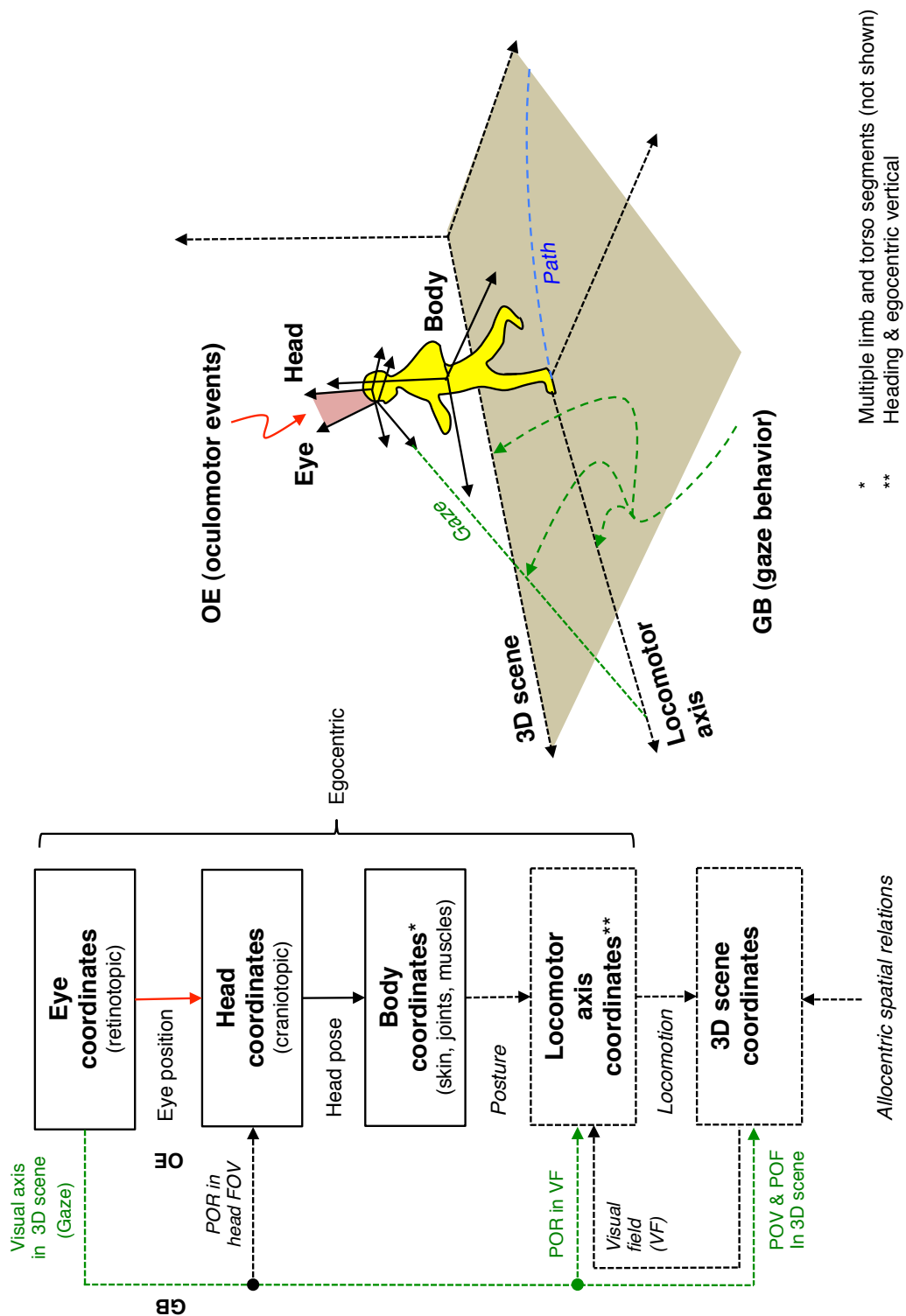
Oculomotor events vs. gaze. Frames of reference (f.o.r.) are represented by boxes and rotations/translations corresponding to the respective coordinate transformations are represented by links. Changes in eye-position define oculomotor events. Gaze refers to the line of sight from the eye to the visual target, determining the projection of the scene to the retinal image. Gaze behavior and oculomotor events coincide, when head and body movement is restrained, e.g. using a bite bar. (The coordinate transformation between head, body and scene reduces to merely a shift of origin). This means that once the eye-tracker is calibrated to a display screen – stationary in the allocentric frame of reference of the lab – a change in eye-position is a gaze shift and vice versa. “In the wild” (free locomotion in a complex 3D scene) this is generally not the case.

Figure 4.



Descriptive terminology used to characterize eye-movements, locomotion and gaze behavior. 3D scene refers to the layout of visible objects surrounding the observer. Trajectory refers to a curve in space described by the successive observer positions in the scene (Note that any egocentric reference point describes a trajectory and the trajectory of the eye and the head are not identical to the locomotor trajectory). Location is the position of the origin of the locomotor axis frame (here: the point in space where control forces are applied to change heading and speed). Point of vantage is where the eye is located. Visual field refers to the projection of the scene to the locomotor frame of reference, with the origin at the intersection of the projections of the heading and the horizon. A point of fixation is the intersection of the visual axes of the eyes with a visible object or location in the scene. The line of sight connecting the point of vantage at the eye to the point of fixation in the scene is also called gaze axis or simply gaze. Note that eye movements may move the point of regard (the projection of gaze into the visual field) even when the point of fixation does not move and vice versa. Also note that locomotion changes the visual directions of targets in the visual field (optic flow), whereas head rotation and eye rotation do not. Also note that POV is under *locomotor* control, and so POF moves in response to both locomotor and oculomotor actions, which need to be very tightly coordinated for accurate gaze control.

Figure 5.



Oculomotor events and gaze behavior in 3D frames of reference. Frames of reference relevant to eye movement behavior during locomotion. Dashed boxes and arrows indicate frames of reference/transformations to which the brain does not have a direct sensory access. Note that these include the locations of the point of vantage (POV) and the point of fixation (POF) in the 3D scene, which together specify gaze, as well as the locomotor heading f.o.r..

SENSORIMOTOR TRANSFORMATIONS

3D FRAMES OF REFERENCE

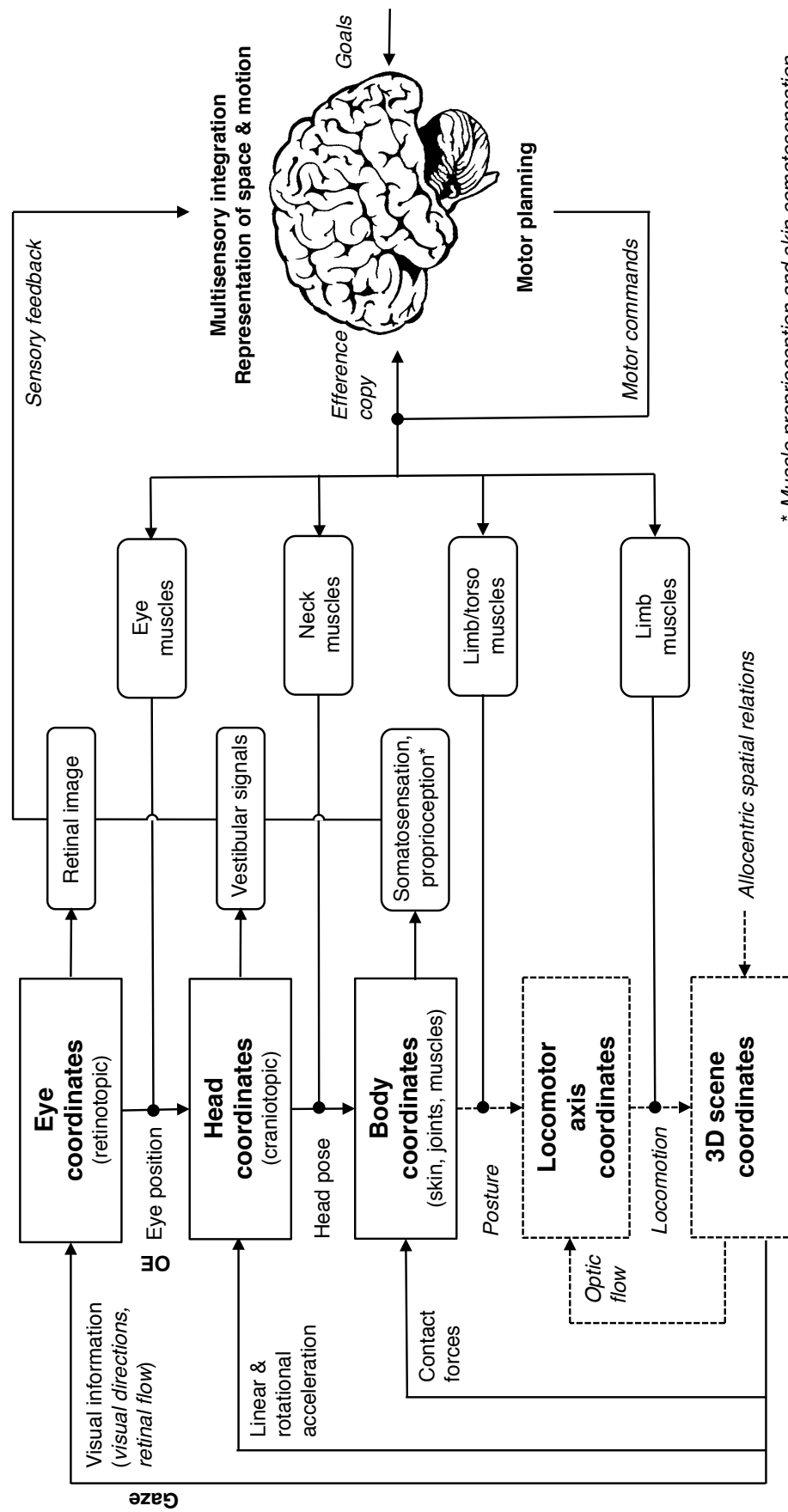
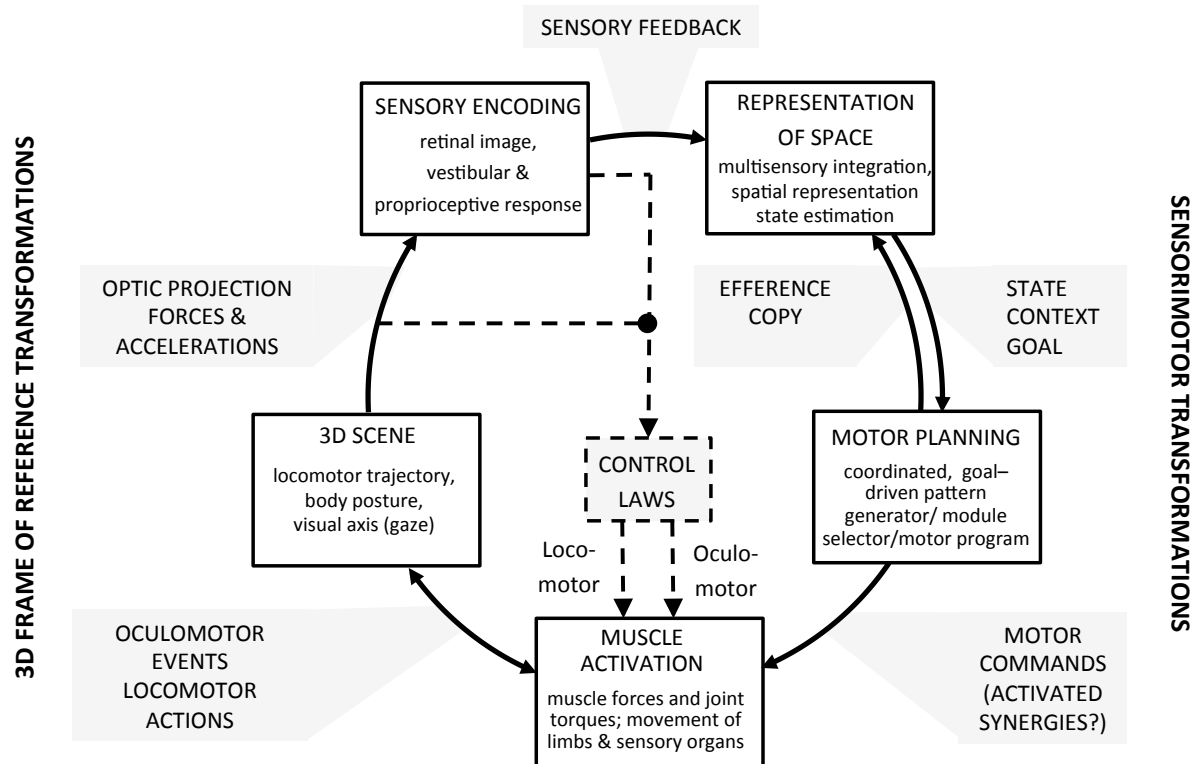


Figure 6.

* Muscle proprioception and skin somatosensation from eye- and head frames of references not shown

3D frames of reference vs. sensorimotor processes/internal representation in the brain. Sensory and motor events are encoded in proprietary frames of reference of motor effectors and sensory organs (including proprioceptive sensations from muscles and joints, please note connections for that eye- and neck muscle proprioception are omitted for clarity). Recovering information about events in frames of reference to which the brain has no direct access (dashed boxes and arrows, italics) requires integrating information across signals in multiple frames of reference. The brain also has to plan and coordinate actions across multiple sensory and motor coordinate systems. OE's in particular are typically part of larger synergistic patterns of motor activity. Thus, analysis of OE's in naturalistic behavior needs to take place in the context of six degrees-of-freedom frame of reference transformations, as well as internal representations and sensorimotor transformations in the brain.

SENSORY FEEDBACK



39

Table 1.

Major characteristics of fixation (including fixational eye movements) and saccades

Major characteristics of fixation (including fixational eye movements) and saccades							
Eye movement type	Oculomotor characteristics				Response characteristics (dependence of movement on optical stimulus)	Function (contributions to gaze behavior)	
	Eye-in-head angular velocity range	Amplitude	Duration	Frequency			
Fixational eye-movements (FEM)	Fixation	n/a	n/a	>100 ms, highly task-dependent	About 2-4 per second, depending on task	n/a	Stabilizing gaze on a stationary target object or location
	Microsaccade	~10 – 120°/s v _{max} . (Martinez-Conde et al., 2009)	< 1°, typically 0.1–0.3° (Martinez-Conde et al., 2009)	About 10–20ms (Martinez-Conde et al., 2009)	About 1 per second	Open loop (pre-programmed, “ballistic”, not corrected by visual feedback during execution)	Gaze shift between objects or locations “on the same target” or very nearby targets in high-accuracy tasks. Possibly refreshing visual image to prevent visual fade by adaptation.
	Drift	4 min arc /s (Collewijn & Kowler, 2008) < 0.5°/s (Rofs, 2009)	A few minutes of arc (Martinez-Conde et al., 2004) 1.5–4 min arc (Collewijn & Kowler, 2008) < 0.13° (Rofs, 2009)	Depends on saccade frequency (fixation duration)	Constant during fixation	Possibly random walk with memory (Engbert, 2005). Degree of conjugacy still unresolved.	Uncertain. Possibly not distinct from OKR/VOR gaze stabilizing responses (or “noise” when the head is immobilized) during fixation.
Rapid eye-movements (REM)	Saccade	125 – 500°/s v _{max} . (Collewijn, Erkelens & Steinman, 1988)	3° – 80° (Collewijn, Erkelens & Steinman, 1988)	25 – 300 ms (Collewijn, Erkelens & Steinman, 1988)	About 2–4 per second, depending on task	Open loop (pre-programmed, no visual or somatosensory feedback).	Gaze shift between objects or locations, re-orienting visual axis in 3D space.
		150 – 600°/s v _{max} . (Baloh et al., 1975)	6° – 90° (Baloh et al., 1975)	40 – 250 ms (Baloh et al., 1975)		Latency 200ms (Sparks, 2002), but down to 100 ms for pre-cued “express saccades” (Fischer & Ramsperger, 1984)	Catch-up saccades in pursuit compensate for <1 smooth eye movement gain.
		10 – 800°/s v _{max} . (Bahill, Clark & Stark, 1975)	0.1 – 60° (Bahill, Clark & Stark, 1975).	10 – 100 ms (Bahill, Clark & Stark, 1975)			Optokinetic nystagmus quick phase.

Table 2.

Major characteristics of “slow” eye movements

Slow eye-movements (SEM)	Eye movement type	Oculomotor characteristics			Response characteristics (dependence of movement on optical stimulus)	Function (contributions to gaze behavior)	
		Eye-in-head angular velocity range	Amplitude	Duration			Frequency
	Vestibulo-ocular response (VOR)	Up to 800°/s (see Sparks, 2002)	n/a	n/a	n/a	Open loop (eye-position or visual signal not fed back to the system), ultra-short latency (15ms, Sparks, 2002)	Stabilizing gaze relative to world. Compensation for head motion during fixation (suppressed during gaze shift and pursuit)
	Optokinetic response (OKR)	Up to about 30°/s (Lisberger, 1997)	n/a	n/a	n/a	Closed loop (visual feedback). Ocular following response latency as fast as 70–75ms (Gellman, Carl & Miles, 1990)	Stabilizing retinal image/gaze relative to world. Compensation for large-scale retinal image slip during fixation.
	Pursuit	Up to about 90°/s eye speed, 100°/s target speed (Meyer, Lasker & Robinson, 1985), 115–150°/s, desaccaded, up to 180°/s target speed (Lisberger, et al., 1981).	n/a	n/a	n/a	Closed loop (visual or somatosensory feedback) with predictive/memory capability. Feedback latency (ocular following response) 65–130ms, (Lisberger et al., 1987), but with predictable target motion zero-lag or anticipatory (negative latency) responses (Barnes & Asselman, 1991; Kowler & Steinman, 1979; Kowler, 1989)	Optokinetic nystagmus slow phase (stare-nystagmus). Stabilizing gaze relative to a moving target object (tracking, sometimes considered equivalent to fixation of a moving target). Alternative interpretation: a slow, motion-driven gaze shift. Optokinetic nystagmus slow phase of look-nystagmus?

Table 3.

List of abbreviations for terms describing oculomotor behavior and gaze. (NB. brain area name abbreviations are given in Figure 2 caption).

OE	Oculomotor event	Stereotypical eye-rotation, characteristics and eliciting conditions, codified in oculomotor experimental paradigms
REM	“Rapid” eye movement	Jerky main sequence oculomotor event (saccade)
SEM	“Slow” eye movement	Smooth non-main sequence oculomotor event (includes VOR, OKR, SPEM, OKN)
FEM	Fixational eye movement	Movement of the eye during fixation, includes REM-type microsaccade and SEM-type drift
CEM	Compensatory eye movement	OE responses that compensate for head rotation and translation (stabilize gaze)
VOR	Vestibulo-ocular response	SEM type CEM dependent on vestibular signals
OKR	Optokinetic response	Mainly bottom-up ocular following tracking response, dependent on retinal image motion. A SEM type CEM
SPEM	Smooth pursuit eye movement	Largely top-down tracking behavior.
OKN	Optokinetic nystagmus	Complex CEM that contains alternating REM-type quick phases and SEM-type slow phases in opposite directions.
OKN SP	Slow phase of OKN	Ocular following or tracking phase of OKN (cf. SPEM, OKR)
OKN QP	Quick phase of OKN	Re-setting phase of OKN
GB	Gaze behavior	Rotation or translation of gaze in 3D space
POV	Point of vantage	Point of origin of gaze (line of sight along the visual axis)
POF	Point of fixation	Endpoint of gaze in the scene
POR	Point of regard	Intersection of gaze with frontoparallel 2D visual field (or field of view of eye-tracker scene camera)
AOI	Area of interest	A region in the 2D visual field (or field of view of eye-tracker scene camera) around a putative visual target