

# **Multi-sensory integration in orientation and self-motion perception**

Ivar Clemens

ISBN: rep-la-cethi-s!-!

Printed by REPLACE-ME, Nijmegen, The Netherlands

©2015 Ivar Clemens

# MULTI-SENSORY INTEGRATION IN ORIENTATION AND SELF-MOTION PERCEPTION

Proefschrift

ter verkrijging van de graad van doctor  
aan de Radboud Universiteit Nijmegen  
op gezag van de rector magnificus prof. dr. Th.L.M. Engelen  
volgens besluit van het college van decanen  
in het openbaar te verdedigen op [dow] [dow] [month] [year]  
om [hour].[min] precies

door

**Ivar A.H. Clemens**  
geboren op 2 juni 1984  
te Boxtel, Nederland

**Promotoren:**

Prof. dr. W.P. Medendorp

**Co-promotoren:**

Dr. ing. L.P.J. Selen

**Manuscriptcommissie:**

Prof. dr. A.V. van den Berg

Prof. dr. J. van der Steen

Dr. S. Nooij

---

# *Contents*

<b>1 General introduction</b>	<b>1</b>
1.1 Sensory signals for navigation . . . . .	3
1.1.1 Vestibular . . . . .	3
1.1.2 Somatosensory . . . . .	7
1.1.3 Visual . . . . .	8
1.2 Vestibular reflexes for the self-motion compensation . . . . .	12
1.3 Multisensory integration . . . . .	13
1.4 Studying the link between sensory signals and perception . . . . .	15
1.4.1 Body-tilt perception . . . . .	16
1.4.2 Visual vertical perception . . . . .	17
1.4.3 Translation perception . . . . .	18
1.5 Outline of this thesis . . . . .	18
<b>2 Multisensory processing in spatial orientation:</b>	
<b>An inverse probabilistic approach</b>	<b>21</b>
2.1 Introduction . . . . .	23
2.2 Materials and Methods . . . . .	25
2.2.1 Subjects . . . . .	25
2.2.2 Setup . . . . .	26
2.2.3 Experiments . . . . .	26
2.2.4 Data analysis . . . . .	28
2.2.5 Sensory integration model . . . . .	29
2.2.6 Model fitting . . . . .	35
2.2.7 Model evaluation . . . . .	36
2.2.8 Model validation: independent test of neck noise . . . . .	37
2.2.9 Model simulation of patient data . . . . .	38
2.3 Results . . . . .	38
2.3.1 Psychometric results . . . . .	38
2.3.2 Model predictions . . . . .	42
2.3.3 Sensory weights . . . . .	43
2.3.4 Model evaluation . . . . .	45
2.3.5 Model validation . . . . .	46
2.4 Discussion . . . . .	46
2.4.1 Comparison with previous work . . . . .	47

2.4.2	Implications of the model . . . . .	48
2.4.3	Model evaluation . . . . .	49
2.4.4	Clinical implications . . . . .	50
2.5	Appendix . . . . .	52
<b>3</b>	<b>Integration of ocular and vestibular signals for self-motion perception in darkness</b>	<b>55</b>
3.1	Introduction . . . . .	57
3.2	Materials and methods . . . . .	58
3.2.1	Participants . . . . .	58
3.2.2	Experimental setup . . . . .	58
3.2.3	Paradigm . . . . .	59
3.2.4	Data analysis . . . . .	61
3.3	Results . . . . .	64
3.3.1	Eye movement contributions to self-motion perception . .	66
3.3.2	Precision depends on PSE . . . . .	68
3.4	Discussion . . . . .	69
3.4.1	Alternative interpretations . . . . .	71
3.4.2	Implications for other studies . . . . .	72
3.4.3	Possible neural substrate . . . . .	73
<b>4</b>	<b>Translation perception is modulated by eye movements that are partially scaled by fixation depth</b>	<b>75</b>
4.1	Introduction . . . . .	77
4.2	Materials and methods . . . . .	78
4.2.1	Participants . . . . .	78
4.2.2	Experimental setup . . . . .	78
4.2.3	Paradigm . . . . .	79
4.2.4	Data analysis . . . . .	81
4.2.5	Model . . . . .	82
4.3	Results . . . . .	83
4.4	Discussion . . . . .	87
4.4.1	Relation to other studies . . . . .	87
4.4.2	Alternative explanations . . . . .	89
<b>5</b>	<b>Visual stability across combined eye and body motion</b>	<b>91</b>
5.1	Introduction . . . . .	93
5.2	Methods . . . . .	94
5.2.1	Participants . . . . .	94
5.2.2	Setup . . . . .	95
5.2.3	Visual stimuli . . . . .	95
5.2.4	Paradigm . . . . .	96
5.2.5	Data analysis . . . . .	98
5.2.6	Model . . . . .	99
5.3	Results . . . . .	101
5.4	Discussion . . . . .	105
5.4.1	Relation to previous studies . . . . .	106

5.4.2	Modelling implications . . . . .	107
<b>6</b>	<b>Summary and discussion</b>	<b>111</b>
6.1	Multisensory processing in spatial orientation . . . . .	113
6.2	Eye movements influence self-motion perception . . . . .	114
6.3	Partial compensation for fixation depth in self-motion perception	114
6.4	Gaze-dependent effects in spatial updating . . . . .	115
6.5	Concluding remarks . . . . .	116
	<b>Bibliography</b>	<b>119</b>



---

*CHAPTER 1*

---

*General introduction*

Life began about four billion years ago with the encapsulation of self-replicating RNA in a lipidic membrane (e.g. Orgel, 1968). Simple as these probionts were, they did not have any means of locomotion and thus relied in full on the currents to obtain the nutrients required to replicate. As organisms evolved, they began to assert an increasing amount of control over their environment. A primitive example can be found in chemotaxis, where ciliae and flagellae, coupled with simple biochemical sensors allow the organism to follow biochemical gradients and thus actively gather nutrients.

Because flagellar motion is limited to either tumbles – random reorientation movements – or runs, the pathway linking these flagella to the biochemical sensors consists of only a few steps. With the advent of more advanced sensory and motoric systems, both the amount of information available to drive behaviour as well as the amount of possible behaviours available increased tremendously. As direct biochemical links between perception and action were no longer sufficient, simple neural networks evolved to process the sensory signals allowing for abstract decisions about movement. Because organisms became more complex, these networks evolved into a complete (central) nervous system, culminating in the cerebral cortex.

In most higher organisms, including humans, the amount of data collected by the sensory organs is enormous, the visual system alone generating between  $10^7$  and  $10^9$  bits per second (Koch et al., 2006; Kelly, 1962). The nervous system, and the cerebral cortex in particular, seem to use probabilistic models to compress the vast amount of the incoming sensory data (Zhaoping, 2006). The principal idea behind these models being that they reduce redundancy in the input by knowledge about the statistics of the natural world (Olshausen et al., 1996).

Predictive models, or maps, on the spatial organization of the environment are of particular interest to an organism, since they provide for example probable locations of food and predators. To use these maps, the brain needs to know

both the location and orientation of the body within the environment. In this thesis, we will investigate how the brain processes the available sensory signals in the perception of gravity as well as in the internal estimation of self-motion. The goal is to build computational models and perform thorough psychometric testing in order to examine the constraints that physics and biology impose on the interaction between the vestibular and other sensory systems.

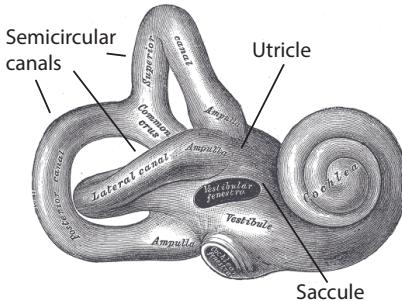
In the following section, sensory sources for gravity and self-motion perception are described in more detail. We then elaborate on how these signals may be used and integrated in the brain, and conclude the chapter with methods for studying the processing of these signals.

## 1.1 Sensory signals for navigation

The signals used for self-motion and orientation perception, can be split in two broad categories: absolute signals and relative signals. Absolute signals, such as landmarks, can be used to directly estimate the location and orientation of an animal while relative signals, such as acceleration, first need to be (mathematically) integrated (i.e. dead reckoning or path integration) and then used to update previous estimates of position and orientation. In the following sections the sensory system involved in position and orientation estimation will be introduced.

### 1.1.1 Vestibular

While many sensory organs supply information that can be used to estimate position and orientation, there is one sensory system that evolved specifically for this purpose: the vestibular system ([Figure 1.1](#)). It contains two sensory components, the semicircular canals and the otoliths. These sensory organs are sensitive to angular velocity and linear acceleration, respectively. Both are located in the labyrinth of the temporal bone in the inner ear.

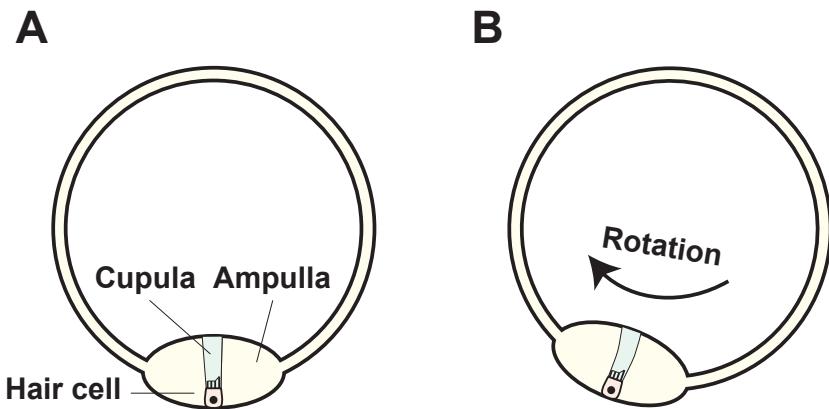


**Figure 1.1:** Schematic view of the inner ear complex consisting of the auditory system (cochlea) and the vestibular system. The vestibular system has three angular velocity (semicircular canals) and two linear acceleration (saccule and utricle) sensors. Image adapted from Gray & Lewis, 1918.

### Semicircular canals

The semicircular canals measure the three-dimensional rotational velocity of the head. Each side of the head contains three orthogonally oriented canals allowing for rotation to be perceived in all three dimensions. Each of the six canals consists of a circular tube filled with a fluid known as endolymph (see [Figure 1.2A](#)). One part of the tube, the ampulla, is a bit thicker than the rest and contains a membrane, the cupula, which separates the fluid. When the head rotates, the fluid stays behind because of its inertia which in turn causes the membrane to deflect (see [Figure 1.2B](#)). While the inertial fluid motion suggests that the cupula should be sensitive to angular acceleration, reactive forces resulting from the fluid motion, such as endolymph viscosity and cupular elasticity, cause the cupular deflection to reflect angular velocity instead (Goldberg, 2012).

The deflection of the cupula is transduced by special hair cells which are partially embedded in the cupula. These hair cells contain bundles of small protrusions, or hairs, on their apical surface. Each bundle contains several smaller stereocilia mechanically linked to one larger kinocilium (Pickles, Comis & Osborne, 1984). When the stereocilia are stretched towards the kinocilium the links cause cation channels (mechanoelectric transducers or METs) to open and the membrane to depolarise. Glutamate is then released into the synapse and causes the afferent neuron to depolarise and, after sufficient depolarisations, to generate action



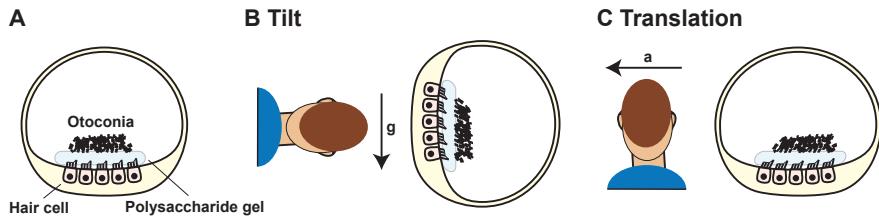
**Figure 1.2:** Detailed view of a semicircular canal **A** in rest and **B** during rotation. Inertia of the endolymph in the canal causes the cupula to deflect during rotation.

potentials (Purves, 2012). These action potentials transmit the angular velocity sensed by the canals to the central nervous system through the vestibulocochlear nerve.

Initially, the afferent signal closely follows rotational velocity, as opposed to acceleration. During sustained rotation, the elasticity of the cupula returns it to its resting position with causes the rotational signal to subside slowly. This is partly compensated for through velocity storage in the central nervous system (Goldberg, 2012). The canals are only sensitive to changes in orientation, as their circular nature makes them insensitive to the effects of linear acceleration and gravity (Goldberg, 2012).

### Otoliths

The linear acceleration and gravitational forces are measured by the two otolith organs on either side of the head: the saccule and the utricle. Each otolith consists of an endolymph filled compartment containing calcium carbonate crystals known as otoconia (Figure 1.3A). Due to their high inertia, these otoconia fall behind during linear acceleration (Figure 1.3B), and move ahead during deceleration. In addition, they are pulled downward as a result of the gravitational



**Figure 1.3:** Detailed view of an otolith **A** in rest, **B** during head tilt, and **C** during linear acceleration. The high inertia of the otoconia keeps them from moving rapidly during both head tilt as well as linear acceleration. As the effect is physically identical in either case, head tilt and linear acceleration cannot be disambiguated based on the otolith signal alone.

field (Figure 1.3C). The otoconia are mounted on top of a flexible polysaccharide gel, in which hair cells, similar to those found in the semicircular canals, are partly embedded (Goldberg, 2012). These hair cells transform the shearing of the polysaccharide gel into a neural signal.

Einstein's equivalence principle states that it is not possible to measure forces caused by linear acceleration and by gravity independently. The signal coming from the otoliths is therefore proportional to the combination of these forces (Fernández & Goldberg, 1976a), which is commonly referred to as the gravito-inertial force (GIF). Compared to the three semicircular canals, which are able to sense rotation in three dimensions due to their orthogonal organization, either side of the head only contains two otolith organs. It is still possible to sense the gravito-inertial force in three dimensions because the otolith organs are curved, and the orientation of the hair cells determines the direction of sensitivity (Goldberg, 2012).

For many actions, the brain needs to disentangle the contributions of gravity and linear acceleration to the gravito-inertial force. For example, the linear vestibulo-ocular reflex (LVOR) that stabilizes gaze during translation should be sensitive to translation while ignoring gravity. Under normal circumstances, the brain is able to perform this task well (Merfeld & Young, 1995), but during extreme conditions such as in airplanes or space flight errors might occur. The somatogravitational illusion (Glasauer, 1995) is an example of a disambiguation error. In this illusion an airplane accelerates forward which causes an inertial force

in the opposite direction. The direction of the resulting GIF will therefore be in between gravity (downward) and the inertial force (backward). The brain erroneously interprets the (majority of the) GIF as being caused by gravity, leading to the perception of "nose-up" pitch tilt.

Various theories have been proposed as to how the brain might solve the tilt-translation ambiguity. The frequency segregation hypothesis, for example, makes use of the fact that when stationary, the only force we experience is gravity. In this case, sustained (low frequency) accelerations should be attributed to gravity, while the high frequency components should be attributed to translation (Paige & Tomko, 1991; Telford, Seidman & Paige, 1997). Other models keep track of the expected direction of gravity by integrating over the vector product of gravity and angular velocity from the semi-circular canals, and subtracting that signal from the otolith signal to obtain an estimate of linear acceleration. Because the semi-circular canals play a crucial role in these models, they are known as multisensory integration models (Mayne, 1974; Ormsby & Young, 1977). Merfeld and Zupan (1995; 2002) further refined the multisensory integration model by explicitly stating that the brain uses an (internal) model of the physical world to resolve the tilt-translation ambiguity.

The disambiguation of linear acceleration and gravity is not solely based on vestibular inputs, but also takes the range of possible movements into account. For example when moved on a linear sled, the probability of a participant perceiving tilt is greatly reduced (Wertheim, Mesland & Bles, 2001), indicating that cognitive processes also influence disambiguation.

### 1.1.2 Somatosensory

The gravito-inertial force (GIF) caused by a combination of gravity and linear acceleration is not only detected by the vestibular system, but also sensed by other sensory systems. Early evidence that non-vestibular sources were used by the brain came from DeKleyn and Versteegh (1933) who showed that inertial

reflexes still occurred after removal of the otolith organs.

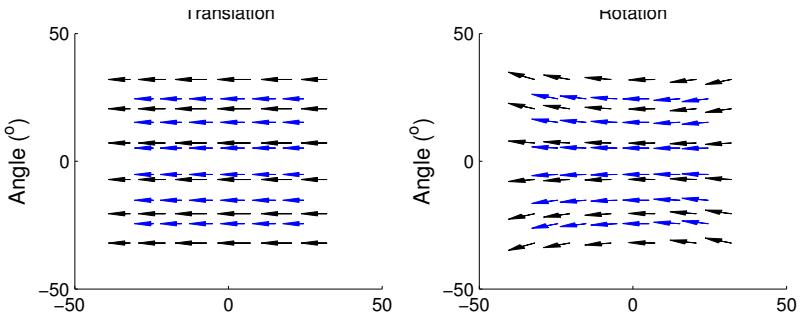
Since then, the contribution of specific organs to GIF perception has been demonstrated. In 1992, Mittelstaedt rotated supine participants along their naso-occipital axis, causing the centrifugal force to act on somatic GIF sensors but not on the otoliths. In nephrectomised participants, the perceived direction of gravity relied less on the centrifugal force compared to controls, suggesting that the kidneys play a crucial role in perception of the gravito-inertial force. Further evidence came from Trousselard (2004), who showed that the perception of gravity in a tilted position depends on whether the stomach is full or empty. In addition, reducing somatosensory cues, by applying a body cast, also affects the perception of gravity (Trousselard et al., 2004). Similar results have been obtained for many other visceral factors, such as the blood vessels (Vaitl, Mittelstaedt, Saborowski, Stark & Baisch, 2002), and spinal axis fluid (Vaitl, Mittelstaedt & Baisch, 1997).

### 1.1.3 Visual

Even though the vestibular and somatosensory systems directly measure the gravito-inertial force, orientation and navigational information can also be extracted from the visual system. In many cases, especially when the low latency of the vestibular signal is not required, the visual signal even overshadows the vestibular one (Wright, DiZio & Lackner, 2005; Grace Gaerlan, Alpert, Cross, Louis & Kowalski, 2012).

#### Vection

When we move through the environment, the image of the world on our retina shifts. This large-field shift pattern, also known as optic flow, depends on the movement being made. Lateral translation (Figure 1.4A) for example causes a different pattern than roll rotation (Figure 1.4B). At the turn of the 19th



**Figure 1.4:** Vection pattern during **A** lateral translation and **B** rotation. During translation nearby targets (red) have a larger retinal displacement compared to far away ones (blue). Retinal displacement across rotation does not depend on target distance.

century von Helmholtz (1867) recognized the importance of these flow signals for self-motion perception. In some cases the optic flow signal is so strong that it causes a percept of self-motion in stationary participants, calledvection (Dichgans & Brandt, 1978). This is experienced, for example, when sitting on the train and a neighbouring train starts to move. This effect is much less likely to occur when on the platform, suggesting that thevection signal is integrated with prior knowledge about the environment before causing self-motion perception (Andersen & Braunstein, 1985; Lepecq, Giannopulu & Baudonnier, 1995)

Similar retinal shifts are also observed during movement of the eyes or head. When interpreting optic flow the brain needs to distinguish object- from self-motion, this process is called optic flow parsing. One strategy is to use extra-retinal cues such as the vestibular, and somatosensory signals to subtract out the retinal stimulation due to self-motion (Wertheim, 1994; Wexler, Panerai, Lamoure & Droulez, 2001; MacNeilage, Zhang, DeAngelis & Angelaki, 2012).

While the extra-retinal information does contribute to optic flow parsing, the existence ofvection (Dichgans & Brandt, 1978) suggests that the brain can parse optic flow using a purely visual approach (Rushton & Warren, 2005; Warren & Rushton, 2007). Warren and Rushton (2009) have shown that the brain indeed uses the global pattern of retinal motion caused by self-motion to parse optic flow. Even though global retinal flow patterns are used to disambiguate



**Figure 1.5:** **A** Trees are aligned with gravity, making them ideal cues for orientation. **B** The rod-and-frame illusion; even though the frame is rotated, participants might perceive the line as being slanted.

between object- and self-motion, it does not mean that a percept of self-motion, i.e.vection, exists. Research has shown thatvection can take up to 30s to establish, but that purely visual optic flow parsing occurs within 1s (Warren & Rushton, 2009).

### Landmarks

In addition to relative cues, the brain seems to use absolute cues in both the perception of gravity and self-motion. For the perception of gravity, it makes use of the fact that many lines within the world are aligned with either the horizon or gravity (Figure 1.5A). Straight lines therefore act as a priors attracting the perceived direction of gravity towards them. A special case is the rod-and-frame illusion (Figure 1.5B), where the perceived angle of a rod relative to gravity is affected by the orientation of the frame that contains it (Witkin & Asch, 1948).

Similarly, our position within the world can be established using world-fixed landmarks. With the exception of animals and vehicles, most items in the world rarely change position. By comparing our current visual scene (e.g. Sequoia trees in Figure 1.5A) with our knowledge about the environment we can

determine our location (in this case northern California). Because the experiments outlined in this thesis were performed in near darkness, these absolute navigational cues can be considered less relevant.

## Oculomotor

Self-motion is typically accompanied by eye movements that help to improve dynamic visual acuity and reduce retinal slip. In [Chapter 3](#) of this thesis, we explore whether these eye movements may have a reversed role, providing cues about self-motion perception beyond optic flow parsing.

In [1963](#), Guedry and Harris reported a substantial underestimation of displacement when their observers watched a small body-fixed target compared to displacements in the dark. Because of the VOR, eye movement in darkness are larger than those in the body-fixed condition. The underestimation in the body-fixed condition can therefore be interpreted as the involvement of eye movements, although the authors did not suggest this. Studies on postural sway also indicate a role of eye movement in self-motion perception. Making eye movements causes postural sway to increase, suggesting that eye movements influence self-motion perception (Glasauer, Schneider, Jahn, Strupp & Brandt, [2005](#); Rodrigues et al., [2015](#)).

One complication is that the magnitude,  $e$ , of these eye movements not only depend on the translation amplitude,  $T$ , but also on the fixation depth,  $d$ . Oculomotor cues to depth include binocular vergence angle and accommodation, which are both most robust for fixations less than a meter or so away. When fixating at Cartesian position  $(x, d)$ , the eye angle is  $e_0 = \arctan(d/x) \approx d/x$ . When translating the body, this angle becomes,  $e_1 = \arctan(d/(x + T)) \approx d/(x + T)$ . The eye movement amplitude is then,  $e = e_1 - e_0 \approx d/T$ . In [Chapter 4](#) we explore whether the brain takes this depth dependent geometry into account and compensates for fixation depth when using eye movements in self-motion perception. Fixation depth plays also a crucial role in [Chapter](#)

[5](#), in which the effects of eye and head movements on spatial updating are investigated.

## 1.2 Vestibular reflexes for the self-motion compensation

Whilevectionisavaluablecue to self-motion perception, the brain wants to minimizevectionbecauseitblurs theimageonthefovea, whichhampersvisualperception. Severalreflexivemechanismsattempttokeepstablefixationtoavoidimageblurringduringself-motion. Twosuchmechanismsarethevestibulo-ocularreflex, theVOR(Goldberg,[2012](#)), and the optokinetic reflex, theOKR(Purves,[2012](#)).

Thevestibulo-ocularreflex(VOR)isdrivenbytheshort-latencyvestibularsignalandcompensatesforbothrotationandtranslationofthehead. TherotationalandtranslationalpartoftheVORhavetheirowndedicatedreflexarcs. ThetranslationalVOR, orTVOR, isdrivenbytheotolithsignalwhiletherotationalVOR, orRVOR, ismainlydrivenbytheseemicircularcanals. TheRVORoperatesduringheadrotationbycounterrotatingtheeyesintheoppositedirectiontothehead. Becauseitsgainisaboutone, eyevelocityisaboutequal, but of oppositesign, totheheadvelocity(Goldberg,[2012](#)). TheTVORoperatesduringheadtranslation, whentheheadmovesorthogonallytothelineofsight. IncontrasttotheRVOR, theTVORneedstotakefixationdepthintoaccount. SimplegeometryshowsthattheidealTVORresponseisinverselyproportionaltofixationdistance,  $e_{vor} = \arctan(d/T)$ . Asaresult, nocompensatoryeyemovementsare requiredwhenfixatingtargetsatinfinitytomaintainastableretinalimage.

Theoptokineticreflexisdrivenbyoptic-flow, andalsocompensatesforboth translationalandrotationalmovement. BecausethecellsdrivingtheOKRare

more sensitive to slow motion, the OKR compensates predominantly for the low frequency components of movements. While the VOR dies away during sustained rotation, the OKR remains because it is sensitive to constant velocity stimuli (Soodak & Simpson, 1988).

The eye movements that are induced by the OKR and the VOR consist of two parts; a slow-phase pursuit-like movement which keep the eyes on target and quick-phase saccades which quickly move the eyes back after they lost track due to physical constraints of the oculomotor system (Goldberg, 2012).

### 1.3 Multisensory integration

While reflexes by definition depend on minimally processed sensory signals, higher animal functions can rely on more elaborate processing. One such processing step is determining the underlying physical cause of a sensory signal. Multiple sensory systems provide information about the same physical quantity. For example, both the visual and the vestibular system provide information about self-motion. Two very naive approaches to derive a self-motion percept from these signal would be to solely rely on the most reliable cue and ignore the others or to just average the available signals. A better solution, however, is to weigh all available signals by their relative reliabilities and combine them with a-priori knowledge about the probability of specific self-motion states. This approach is known as statistically optimal, or Bayesian, integration.

We will now briefly explain the mathematical foundations of optimal integration. Suppose there is a true physical stimulus in the world,  $x$ , which is observed by multiple sensory systems,  $x_i$ . We assume that these observations are corrupted by independent Gaussian noise,  $\sigma_i^2$ . The probability of the observations,  $x_i$ , given the stimulus  $x$ , is therefore:

$$P(x_i|x) = \mathcal{N}(x_i, \sigma_i^2) \quad (1.1)$$

From the point of view of the brain, the observations,  $x_i$ , are given while the stimulus,  $x$ , has to be inferred. In this case, probability  $P(x_i|x)$  is referred to as the likelihood of the stimulus given the observations, or  $\mathcal{L}(x|x_i)$ . Given multiple sensory observations,  $x_1, \dots, x_n$ , we can now compute the likelihood of the stimulus:

$$\mathcal{L}(x|x_1, \dots, x_n) = P(x_1, \dots, x_n|x) = \prod_i P(x_i|x) \quad (1.2)$$

Making use of Bayes' rule, the brain can infer the probability of the stimulus given its observations:

$$P(x|x_1, \dots, x_n) = \frac{P(x_1, \dots, x_n)P(x)}{P(x_1, \dots, x_n)} \quad (1.3)$$

In this equation,  $P(x)$  represents the prior probability of the stimulus, that is the probability of the stimulus occurring without taking the sensory input into account. The left-hand side of the equation,  $P(x|x_1, \dots, x_n)$ , is commonly referred to as the posterior probability.

The most likely stimulus given the observations is the one for which the posterior probability is largest. This method of finding the most optimal estimate ([Equation 1.4](#)) is known maximum a posteriori estimation (MAP).

$$\hat{x} = \arg \max_x P(x_1, \dots, x_n|x)P(x) \quad (1.4)$$

When using a flat prior, that is when all stimuli are equally likely to occur,  $P(x) = c$ , [Equation 1.4](#) can be simplified to

$$\hat{x} = \arg \max_x P(x_1, \dots, x_n|x) \quad (1.5)$$

In this case, we refer to it as maximum likelihood estimation (MLE). The

solution for the most likely stimulus,  $\hat{x}$ , is a weighted sum of the sensory inputs,

$$\hat{x} = \sum_i w_i x_i \quad (1.6)$$

with weight,  $w_i$ ,

$$w_i = \frac{1/\sigma_i^2}{\sum_j 1/\sigma_j^2} \quad (1.7)$$

As the posterior is a distribution, it also provides us with an estimate of uncertainty in the most likely stimulus value,

$$\hat{\sigma}_i^2 = \frac{\prod_i \sigma_i^2}{\sum_i \sigma_i^2} \quad (1.8)$$

In many cases the brain does not use a flat prior, but a Gaussian one. For example, in [Chapter 2](#), we use a prior that encodes the assumption that the head is mostly upright in space. In this case, the Gaussian prior can be seen as an additional sensory signal that is weighted into the posterior estimate.

Statistically optimal integration has been shown in many situations, for example in sensorimotor learning (Kording & Wolpert, [2004](#)), spatial localization (Battaglia, Jacobs & Aslin, [2003](#)) and updating (Vaziri, Diedrichsen & Shadmehr, [2006](#)).

## 1.4 Studying the link between sensory signals and perception

A large part of our knowledge on the relation between perception and sensory processing comes from psychophysical experiments. In this type of experiments

the relation between physical stimuli, e.g. the true angle of the body relative to gravity, and perception, e.g. perceived body orientation, is quantified.

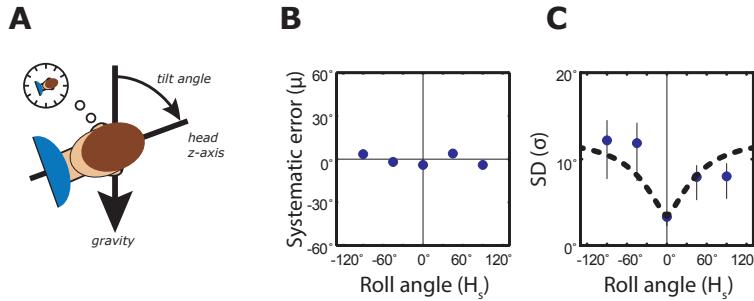
A powerful tool in psychophysical research is the two-alternative forced choice (2AFC) paradigm. In this paradigm, participants are presented with two stimuli, e.g. two translation distances, and are then forced to make a choice, e.g. on which of the two translations was longer. By systematically varying one of the choice alternatives (i.e., the probe stimulus), the point of subjective equality, or PSE, can be determined. The PSE is the point at which the participant is completely unsure about which stimulus to pick, i.e. gives a random response, because he perceives the two stimuli as being equal. If the PSE does not line up with the actual reference stimulus, there is a bias in the percept, represented by the Greek letter mu ( $\mu$ ).

In addition to establishing perceptual equality, the 2AFC paradigm also allows for quantifying the uncertainty in perception by looking at the slope of the psychometric curve around the PSE. This uncertainty is commonly quantified by the standard deviation of the cumulative Gaussian probability distribution that underlies the responses of the subject, represented by the Greek letter sigma ( $\sigma$ ).

The next three sections will introduce the three 2AFC tasks that are central in the present work.

#### 1.4.1 Body-tilt perception

In the subjective body tilt, or SBT, task the perceived body orientation with respect to a given body tilt angle is probed. Participants are first given a reference angle, e.g.  $45^\circ$ , and are then rotated to an angle close to the reference angle, e.g.  $46^\circ$  (see [Figure 1.6A](#)). They then have to indicate whether their current orientation is clockwise or counter-clockwise with respect to the reference angle. By systematically probing rotation angles around the reference angle we



**Figure 1.6:** Subjective body tilt task (SBT); **A** Graphical representation of the SBT task. **B** Bias and **C** uncertainty as a function of roll angle in a typical participant (circles). The dashed line represents the typical pattern.

can obtain both the bias and uncertainty on the percept of body tilt.

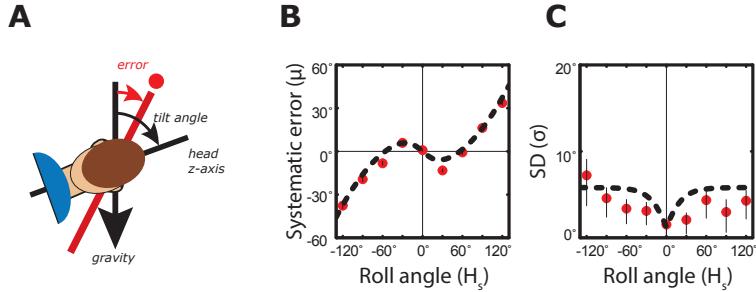
Most participants are able to do this perfectly, regardless of the reference angle (see Figure 1.6B). Their uncertainty has been shown to increase as a function of reference angle though (see Figure 1.6C).

As the SBT probes body orientation, the somatosensory signal originating from the torso can be used without any reference transformation and thus provides a direct contribution. Other sensory signals such as the vestibular signal can also be used, but only after a reference frame transformation.

#### 1.4.2 Visual vertical perception

The subjective visual vertical, or SVV, task is a similar task in which participants have to judge the orientation of a line with respect to gravity. The PSE, that is the angle at which the line is perceived to be aligned with gravity, can be found by presenting lines at different angles and asking the participant whether the line is rotated clockwise or counter-clockwise relative to gravity (see Figure 1.7A).

When seated straight this task, participants do not make any static errors and are very certain about their responses (see Figure 1.7B and C). This changes when tilting the participant before the task. In general, the static error increases



**Figure 1.7:** Subjective visual vertical task (SVV); **A** Graphical representation of the SVV task. **B** Bias and **C** uncertainty as a function of roll angle in a typical participant (circles). The dashed line represents the typical pattern.

with tilt angle. When the direction of the error is in the direction of the body mid-line, this effect is known as the Aubert or A-effect (Aubert, 1861). At smaller angles overcompensation occurs and the static error is in the opposite direction, which is away from the body mid-line. This latter effect is known as the E-effect and could be due to the effects of ocular counter-roll (OCR).

### 1.4.3 Translation perception

In the lateral translation task participants had to judge which of two subsequent lateral translations was longest in magnitude. By manipulating one of these translations while keeping the other constant, the effects of these manipulations on self-motion perception can be studied. Because we ask participants to report on the difference between two translations (i.e. longer versus shorter), only relative effects can be assessed. We used the lateral translation task in Chapters 3 and 4 to assess the effects of fixation type and distance on self-motion perception.

## 1.5 Outline of this thesis

Both the estimation of body orientation and self-motion require the integration of multiple sensory signals. This thesis explores how the brain weights and

integrates the different sensory signals to form dynamic but coherent percepts of self-motion and orientation.

We start with [Chapter 2](#) in which we investigate sensory noise levels of the visual, somatosensory, and vestibular cues to spatial orientation. Because the contributions of these signals cannot be studied in isolation, we adopted an inverse approach where we assumed statistical optimality and attempted to compute the statistical properties of individual sensors. To this end, we fitted an optimal integration model to the behaviour from two psychophysical tasks. The first gauged head-in-space orientation (SVV) and the second body-in-space orientation (SBT). The resulting estimates of the noise levels in the involved sensory modalities were used to predict the responses of patients with somatosensory and vestibular deficits. These predictions were consistent with previously published deficits in these patient groups, strengthening the idea that human spatial orientation is statistically optimal.

In [Chapter 3](#) we examine the dynamic aspects of the integration of vestibular and non-vestibular cues for self-motion perception. We specifically focus on the contributions of eye movements to the perception of whole-body translation. Using a psychophysical task, we first show that translations in which the eyes fixate a body-fixed target are perceived as shorter than those in which the eyes fixate a world-stationary target and thus make pursuit movements. We further demonstrate that this result does not depend on the presence of the fixation point *per se*, but that the magnitude of unconstrained eye movements in complete darkness directly influences the self-motion percept.

Because fixation distance influences the magnitude of pursuit eye movements, we investigated whether the brain takes fixation distance into account in [Chapter 4](#). We compared self-motion perception during gaze fixation on near and far world-fixed targets. Results suggest that fixation depth is only partially taken into account in the integration of vestibular and non-vestibular information for the calculation of translation distance. It seems that only raw eye movements

augment self-motion perception, resulting in a biased percept.

An accurate internal estimate of self-motion perception is not only crucial for navigation, but can also be used to updated locations of previously seen objects. This process is known as spatial updating. [Chapter 5](#) presents the final experimental study of this thesis, where we investigate how self-motion signals are used in spatial updating. We show that the errors made during spatial updating do not only depend on the location of the previously seen object, but also on the location of gaze. This is consistent with spatial updating in a gaze-centred reference frame. We further show that the underestimation of self-motion is a possible cause of the errors found.

---

*CHAPTER 2*

---

*Multisensory processing in  
spatial orientation:  
An inverse probabilistic  
approach*

**Abstract** Most evidence that the brain uses Bayesian inference to integrate noisy sensory signals optimally has been obtained by showing that the noise levels in each modality separately can predict performance in combined conditions. Such a forward approach is difficult to implement when the various signals cannot be measured in isolation, as in spatial orientation, which involves the processing of visual, somatosensory, and vestibular cues. Instead, we applied an inverse probabilistic approach, based on optimal observer theory. Our goal was to investigate whether the perceptual differences found when probing two different states – body-in-space and head-in-space orientation – can be reconciled by a shared scheme using all available sensory signals. Using a psychometric approach, seven human subjects were tested on two orientation estimates at tilts  $<120^\circ$ : perception of body tilt [subjective body tilt (SBT)] and perception of visual vertical [subjective visual vertical (SVV)]. In all subjects, the SBT was more accurate than the SVV, which showed substantial systematic errors for tilt angles beyond  $60^\circ$ . Variability increased with tilt angle in both tasks, but was consistently lower in the SVV. The sensory integration model fitted both datasets very nicely. A further experiment, in which supine subjects judged their head orientation relative to the body, independently confirmed the predicted head-on-body noise by the model. Model predictions based on the derived noise properties from the various modalities were also consistent with previously published deficits in vestibular and somatosensory patients. We conclude that Bayesian computations can account for the typical differences in spatial orientation judgments associated with different task requirements.

---

This chapter has been published as

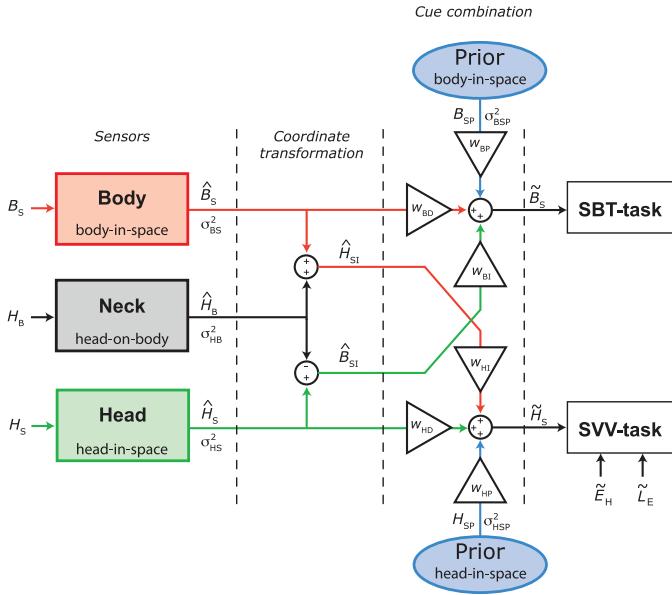
Clemens, I.A.H., De Vrijer, M., Selen, L.P.J., Van Gisbergen, J.A.M. and Medendorp, W.P. (2011). Multisensory processing in spatial orientation: An inverse probabilistic approach. *Journal of Neuroscience*, 31(14): 5365-5377.

## 2.1 Introduction

To infer the current state of the body in space, the brain must rely on noisy sensory inputs. Thus, a degree of uncertainty in the reconstructed physical state is unavoidable. However, according to the rules of Bayesian inference, perceptual uncertainty can be reduced by combining overlapping information from different sensory modalities, weighting each signal in proportion to its reliability (Knill & Pouget, 2004; Kording & Wolpert, 2004; Angelaki & Cullen, 2008). For example, psychophysical studies have shown that human observers behave optimally when integrating visual–proprioceptive (van Beers, Sittig & Denier van der Gon, 1999), visual–haptic (Ernst & Banks, 2002), or visual–auditory (Alais & Burr, 2004) cues. In these studies, the approach was to estimate noise levels of the two sensory modalities in separate unimodal experiments that were then used to predict performance in the bimodal case. Unfortunately, such a forward approach is difficult to implement when the involved sensory modalities cannot be studied in isolation, as in spatial orientation, which involves visual, somatosensory, and vestibular cues. Here the visual contribution can be eliminated easily, but to test whether somatosensory and vestibular cues are combined optimally, one cannot simply “switch off” one system to assess the noise level of the other.

Instead, we took optimality as a starting point and implemented an inverse probabilistic approach to deduce noise levels of the various individual sensors. We probed two spatial orientation estimates—body-in-space and head-in-space orientation—which, according to optimal theory, will use all available sensory signals that can be obtained by various reference frame transformations. This transformation and integration scheme, shown in [Figure 2.1](#), involves at least three sensory systems: (1) head sensors, supplying information about the orientation of the head with respect to gravity (vestibular system); (2) body sensors, providing an estimate of the orientation of the body in space (“somatic graviceptors”) (Mittelstaedt, 1997); and (3) neck sensors, providing an estimate of the angle between head and body (neck proprioceptors).

In this scheme, an estimate of body orientation in space can be obtained directly from the body sensors, but also indirectly from the head-sensor signal, by subtracting



**Figure 2.1:** Schematic representation of the sensory integration model. Sensory signals, denoted by a hat symbol (^), are assumed to be calibrated accurately, but contaminated by Gaussian noise. Optimal estimates are denoted by a tilde (^~). Body sensors, neck sensors, and otoliths provide information about orientation of body in space ( $B_S$ ), head on body ( $H_B$ ), and head in space ( $H_S$ ), respectively. Neck signal ( $\hat{H}_B$ ) is used for a reference frame transformation of otolith information into a body-in-space signal ( $\hat{H}_S - \hat{H}_B = \hat{B}_{SI}$ ), and for a transformation of body-tilt information into a head-in-space signal ( $\hat{B}_S + \hat{H}_B = \hat{H}_{SI}$ ). For an optimal estimate of body-in-space orientation,  $\hat{B}_S$  (SBT task), the model combines the body-sensor signal ( $\hat{B}_S$ , red pathway) with a reference-frame-transformed otolith signal ( $\hat{B}_{SI}$ , green pathway). Relative contributions of the two pathways ( $w_{BD}$  and  $w_{BI}$ ) depend on their relative precision (Equation 2.2). The scheme shows a symmetrical arrangement with two priors, but there is ample reason to believe that their effects are not identical. The simplest explanation of current and previous SBT data (see Materials and Methods, SBT computation) indicates that the associated prior in this task is uniform, which implies that  $w_{BP}$  can be ignored. In the SVV task, an optimal estimate of head-in-space ( $\tilde{H}_S$ ) is obtained by integration of otolith information ( $\hat{H}_S$ , green pathway), reference-frame-transformed information from body sensors ( $\hat{B}_{SI}$ , red pathway), and a significant contribution from prior information ( $H_{SP}$ , blue pathway). Relative weights are denoted by  $w_{HD}$ ,  $w_{HI}$ , and  $w_{HP}$ , respectively. Estimate of line-in-space orientation is obtained by combining  $\tilde{H}_S$  and estimates of eye-in-head ( $\tilde{E}_H$ ) and line-on-eye ( $\tilde{L}_E$ ) orientation. Noise variance in body sensors ( $\sigma_{BS}^2$ ), neck sensors ( $\sigma_{HB}^2$ ), otoliths ( $\sigma_{HS}^2$ ), and width of prior ( $\sigma_{HSP}^2$ ) defines their relative weights (see Materials and Methods). Otolith noise may depend on tilt angle (Equation 2.11). Note that the process of sensory integration, denoted here by summation of weighted sensory signals, is equivalent to multiplication of the underlying probability distributions (Equations 2.2 and 2.6 and Appendix).

the neck signal. Likewise, the estimate of head-in-space orientation can be obtained from the head sensors, but also through an indirect pathway, by combining the body-sensor signal with neck information. Importantly, as the two state estimates require different transformations, Bayesian theory predicts that the relative contribution of the sensory signals will differ as well (McGuire & Sabes, 2009). Apart from the crucial role of sensory information, the scheme allows for the possibility that the estimates of the two orientation states can be further influenced by prior beliefs about sensory states.

Here, we used two psychophysical tasks – subjective body tilt (SBT) and subjective visual vertical (SVV) to quantify the two orientation estimates in a group of healthy subjects. Using an inverse probabilistic approach, we obtained stable solutions for the noise properties of the involved sensor systems. Independent measurements of neck noise confirmed the levels predicted by the model. Forward model predictions based on these noise properties were consistent with previously published deficits of bilateral vestibular and paraplegic patients, which would be difficult to explain otherwise. Our results suggest that Bayesian integration of multisensory information explains the major task-dependent features in spatial orientation perception.

## 2.2 Materials and Methods

### 2.2.1 Subjects

Seven subjects (6 male, 1 female) provided written informed consent to participate in the experiments. Ages ranged from 23 to 65 years. Subjects were free of any known vestibular or other neurological disorder and had normal or corrected-to-normal visual acuity. All subjects took part in SBT and SVV experiments (see below, [Experiments](#)) and returned to the laboratory for an independent measurement of neck proprioception. Before each experiment started, subjects received careful instructions and performed a few practice runs to get used to the task. Participants never received any feedback about their performance, not even in the training trials. Each subject participated in 20 experimental sessions of ~45 min each, yielding >15h recording time per participant.

### 2.2.2 Setup

Body tilt was controlled by a computer-controlled vestibular chair, which was configured to allow subject rotation in the roll axis. A digital position encoder measured roll position with an angular resolution of  $0.04^\circ$ . The subject's body was tightly fixated using a five-point seat belt and adjustable shoulder and hip supports. Velcro straps restrained both legs and feet, and a padded helmet firmly fixated the head in a natural upright position for looking straight ahead. Subject-specific seat adjustments ensured that the naso-occipital axis, midway between the eyes, coincided with the roll axis of the chair. Experiments took place in complete darkness.

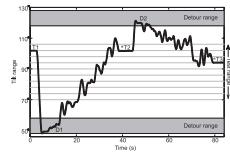
### 2.2.3 Experiments

#### SBT

The SBT experiment served to obtain a psychometric measure of each subject's accuracy and precision of body-tilt perception at five body-tilt angles: upright ( $0^\circ$ , SBT<sub>0</sub> task) and  $45^\circ$  and  $90^\circ$  right side and left side down (SBT $\pm 45$  task and SBT $\pm 90$  task). Negative angles indicated left side down. We applied the method of constant stimuli, using a set of 10 equidistant body-tilt angles, centred on tentative estimates of the subject's  $0^\circ$  (SBT<sub>0</sub>),  $45^\circ$  (SBT<sub>45</sub>),  $-45^\circ$  (SBT<sub>-45</sub>),  $90^\circ$  (SBT<sub>90</sub>), and  $-90^\circ$  (SBT<sub>-90</sub>) body-tilt percept. The latter were determined in a few pilot trials that also served to familiarise the subject with the task, without providing a reference of the five respective orientations to be tested. Relative to the test angle, we used test angle intervals of  $3^\circ$ ,  $4^\circ$ , and  $4^\circ$  in the SBT<sub>0</sub>, SBT $\pm 45$ , and SBT $\pm 90$  tasks, respectively. Body-tilt angles were tested 14 times in random order, yielding 140 responses for each psychometric curve.

To perform the psychophysical SBT experiments, two methodological problems had to be solved. The first relates to the number of experimental sessions that we could reasonably ask subjects to perform. We realised that returning the subject to upright for reorientation after each trial would require too large a number of experimental sessions. Our overriding concern was that starting each trial from upright would

confound the SBT<sub>0</sub> task in the sense that subjects could then simply notice the change in chair position. To prevent this, we always inserted a detour rotation before moving the subject to the test angle in a given trial. The detour, always to a tilt position clearly outside the psychometric test range, served to reset the subject's memory of the previous tilt position. These detour angles were chosen randomly from a range at 30-40° clockwise (CW) and counterclockwise (CCW) from the presumed threshold. As an illustration, Figure 2.2 shows how the subject was moved from one trial to the next in the course of an SBT<sub>90</sub> experiment. Detour angles preceding each test angle were taken from the CW and the CCW detour range in equal proportions. An analysis of trial history effects indicated that detour angles did not affect the judgment in the subsequent trial ( $p > 0.05$ ).



**Figure 2.2:** Tilt paradigm in SBT<sub>90</sub> task. T1, T2, T3, Test angles at which the subject was prompted with a beep signal (\*) to indicate whether body orientation was CW or CCW from the instructed reference orientation (i.e., 90° in this example). D1, D2, Detour angles randomly drawn from detour range (30-40° CW and CCW from centre of test range). Rotations from detour (D) to test (T) angle were performed in a noisy fashion (see Materials and Methods, SBT).

Each experimental run started in upright position with the room lights on. After the lights were turned off, subjects were first rotated at a constant angular velocity of 30°/s to a random detour angle, outside of the test angle range, where they remained for 3 s. The chair then moved to a randomly chosen position within the test range with a very slow and noisy profile, defined by the sum of a ramp of 0.4 -

2°/s and Gaussian white noise (bandwidth, 0 - 0.7 Hz; RMS amplitude, 3.4°). Ramp speed was chosen such that the trajectory between detour angle and test angle was reached in 30 s (Figure 2.2). These precautions were taken to enforce independent absolute tilt judgments and to deter reliance on sensed changes in tilt position that had occurred since the previous trial. Three seconds after arrival at the test angle, a beep signal prompted the subject to indicate whether body orientation was CW or CCW from the instructed reference orientation (0° in the SBT<sub>0</sub> task, ±45° in the SBT<sub>±45</sub>, or ±90° in the SBT<sub>±90</sub> task), using a toggle switch. The subject was then rotated at a constant velocity to a new randomly drawn detour angle, and the above procedure was repeated. Each run, comprising seven test angles, lasted ~5 min, after which the subject was rotated back to upright, and room lights were turned on. Between runs,

there was a 60 s rest interval before the next run started. Each SBT task was tested in separate sessions of ~45 min each, thus amounting to a total of 15 sessions per subject (i.e., ~11 h of recording time).

## SVV

The same subjects were also tested in a series of SVV experiments. Part of this dataset (four subjects) has been published previously as part of a larger dataset on visual verticality perception (de Vrijer, Medendorp & van Gisbergen, 2009). Data in the other three subjects were collected anew. Here we provide a brief summary of the paradigm. SVV was tested at nine roll-tilt angles, ranging from -120 to 120° at 30° intervals. A luminous line (angular subtend, 20°), polarised with a bright dot at one end, was mounted in front of the subject. The line's rotation axis coincided with the chair rotation axis. In each experimental run, the subject was rotated from upright to the chosen test angle at a constant angular velocity of 30°/s. After a 30 s waiting period that allowed canal effects to subside, a luminous line was briefly flashed (20 ms), and the subject indicated whether its orientation in space was CW or CCW from the perceived direction of gravity. The line orientation was selected randomly from a set of 11 line orientations. After all line orientations had been tested, the subject was rotated back to upright, and room lights were turned on. Positive and negative body-tilt angles were alternated regularly. As in the SBT experiment, we used the method of constant stimuli. The set of 11 line orientations was centred on a coarse estimate of the SVV threshold at each tilt angle. We used orientation intervals of 3°, except for upright, where intervals of 2° were taken. Each set of line orientations was tested in random order in 12 experimental runs, thus yielding a total of 132 responses for each psychometric curve. SVV data were collected in a total of five 45 min sessions per subject.

### 2.2.4 Data analysis

CW tilt angles of the body and the luminous line were defined positive. We quantified performance, for each roll-tilt angle (5 in the SBT and 9 in the SVV) independently, by examining the proportion of CW responses as a function of body orientation (SBT)

and the proportion of CCW responses as a function of line orientation with respect to the body (SVV). Psychometric data were quantified by fitting a cumulative Gaussian function ([Figure 2.3](#)):

$$P(x) = \lambda + (1 - 2\lambda) \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(y-\mu)^2/2\sigma^2} dy, \quad (2.1)$$

in which  $x$  represents body orientation in space (SBT experiment) or line orientation with respect to the body (SVV experiment). The mean of the Gaussian  $\mu$  represents the subjective perception of the reference orientation in the SBT task, or the SVV compensation angle (the angle between the apparent visual vertical line and the body axis) in the SVV task. The width of the curve,  $\sigma^2$ , inversely related to precision, serves as a measure of the subject's variability in the SBT or SVV task. Parameter  $\lambda$ , representing the lapse rate, accounts for stimulus-independent errors caused by subject lapses or mistakes and was restricted to small values ( $\lambda < 0.06$ ). Fits were performed using Matlab software (MathWorks) with the “psignifit” (Wichmann & Hill, [2001b](#)) routine.

### 2.2.5 Sensory integration model

To provide a theoretical framework that explains the observed responses, we designed a sensory integration model for visual verticality and body-tilt perception that assumes optimal processing of all potentially relevant sensory signals, including body, head, and neck sensors. The model links accuracy and variability in the two spatial tasks to the properties of the underlying sensors. For simplicity, the scheme is limited to SBT and SVV signal processing in darkness.

In the scheme ([Figure 2.1](#)), we use the following conventions: physical variables are denoted by a capital with a subscript indicating the frame of reference. For example,  $H_S$  represents the physical orientation of the head in space. Sensory signals and their reference-frame-transformed counterparts are denoted by a hat symbol (^), as in  $\hat{H}_S$ , which represents the orientation of the head in space as measured by the head-in-space sensors. The optimal estimate of a variable, obtained by integration of all available information, is indicated by a tilde (~), as in  $\tilde{H}_S$ , representing the final head-in-space

estimate.

It is assumed that all sensory signals are accurately calibrated (i.e., unbiased) but corrupted by independent Gaussian noise with a given variance ( $\sigma^2$ ), with subscripts to indicate the sensory modality (e.g.,  $\sigma_{BS}^2$  represents noise variance in the body-in-space sensors).

### SBT computation

To obtain an estimate of the orientation of the body in space, the brain can use “direct” sensory information from body sensors ( $\hat{B}_S$ ), such as tactile receptors in the skin or so-called graviceptors in the trunk (Mittelstaedt, 1997, 1998; Vaitl et al., 2002). Alternatively, an “indirect” pathway, involving a reference frame transformation, can also provide a body-in-space estimate. For this purpose, sensory head-in-space information, provided by the otoliths, must be combined with information about head-on-body orientation, provided by proprioceptive signals from the neck ( $\hat{B}_{SI} = \hat{H}_S - \hat{H}_B$ ). Because the sensors are contaminated by noise, the direct and indirect signals can be represented as Gaussian probability distributions with mean values of  $\hat{B}_S$  and  $\hat{H}_S - \hat{H}_B$ , and variance levels of  $\sigma_{BS}^2$  and  $\sigma_{HS}^2 + \sigma_{HB}^2$ , respectively. Theoretically, as shown in Figure 2.1, the brain could also use prior information about body-in-space orientation in the computation of the body-in-space estimate. The effect of including a prior on the SBT (centred on upright) would be a systematic error of underestimation at larger tilt angles. However, neither previous findings (Mittelstaedt, 1983; Mast & Jarchow, 1996; Jarchow & Mast, 1999; van Beuzekom, Medendorp & van Gisbergen, 2001) nor the present results (Figures 2.3 and 2.4), showed such systematic errors across subjects. In modelling terms, this indicates a uniform (uninformative) prior, which corresponds to a weight of 0. Accordingly, a statistically optimal estimate of body-in-space orientation ( $\tilde{B}_S$ ) is then given by the peak of the Gaussian distribution that results from the multiplication of the two distributions representing the direct and indirect sensory pathways. It follows that

$$\tilde{B}_S = w_{BD} \cdot \hat{B}_S + w_{BI} \cdot (\hat{H}_S - \hat{H}_B), \quad (2.2)$$

with

$$w_{BD} = \frac{1/\sigma_{BS}^2}{1/(\sigma_{HS}^2 + \sigma_{HB}^2) + 1/\sigma_{BS}^2} \quad (2.3)$$

and

$$w_{BI} = \frac{1/(\sigma_{HS}^2 + \sigma_{HB}^2)}{1/(\sigma_{HS}^2 + \sigma_{HB}^2) + 1/\sigma_{BS}^2} \quad (2.4)$$

in which  $w_{BD}$  and  $w_{BI}$  (Figure 2.1) represent the respective weights of the direct and indirect pathways, which add up to 1 (Landy, Maloney, Johnston & Young, 1995; Jacobs, 1999; Ernst & Banks, 2002; Bays & Wolpert, 2007). Note that the weight of each pathway depends on its reciprocal noise variance (also known as precision), so that precise signals have a stronger influence on the final estimate than noisy signals. Furthermore, because both sensory pathways are supposed to carry unbiased signals, the mean estimate of body in space in multiple trials,  $\mu(\tilde{B}_S)$ , will also be accurate.

It can further be shown that the variance in  $\tilde{B}_S$  in multiple trials, denoted as  $\sigma^2(\tilde{B}_S)$ , equals

$$\sigma^2(\tilde{B}_S) = \frac{\sigma_{BS}^2 \cdot (\sigma_{HS}^2 + \sigma_{HB}^2)}{\sigma_{BS}^2 + (\sigma_{HS}^2 + \sigma_{HB}^2)} \quad (2.5)$$

which implies that the final estimate has a lower variance than the signal provided by either the direct or the indirect pathway (Ernst & Banks, 2002; Ernst & Bülthoff, 2004). Because we assume that sensory signals are accurate and that there is no prior information about body in space, the model predicts that there are no systematic errors in the SBT, so that  $\mu(\tilde{B}_S) = B_S$ . The variance in the SBT task is represented by  $\sigma^2(\tilde{B}_S)$ .

## SVV computation

The scheme applies a similar sensory signal processing strategy to estimate the orientation of head in space,  $\tilde{H}_S$ , used in the SVV. A direct estimate of head-in-space orientation is provided by the head-in-space sensors ( $\hat{H}_S$ ), and an indirect estimate is obtained

by a reference frame transformation of the body-in-space signal ( $\hat{B}_S$ ) by adding the head-on-body estimate ( $\hat{H}_B$ ), provided by the neck sensors ( $\hat{H}_{SI} = \hat{B}_S + \hat{H}_B$ ). Again, direct and indirect pathway signals are represented by two Gaussian probability distributions, with mean values of  $\hat{H}_S$  and  $\hat{B}_S + \hat{H}_B$ , respectively, and corresponding variances of  $\sigma_{HS}^2$  and  $\sigma_{BS}^2 + \sigma_{HB}^2$ . In the computation of the head-in-space estimate, to account for systematic errors (MacNeilage, Banks, Berger & Bühlhoff, 2007; de Vrijer, Medendorp & van Gisbergen, 2008), it is further assumed that the brain uses prior knowledge about head-in-space orientation, which entails that small head-tilt angles are considered more probable than large tilts. Mathematically, the prior is represented by a Gaussian distribution that is centred at  $0^\circ$  head tilt ( $H_{SP} = 0^\circ$ ) with a variance of  $\sigma_{HSP}^2$ . Note that, in our scheme, the head-in-space prior, which contributes to the SVV computations, does not affect the body-in-space estimate. Integration of the direct and indirect sensory pathways and prior knowledge is performed by multiplication of the three Gaussian distributions. The peak of the resulting posterior distribution represents the optimal estimate of head-in-space orientation ( $\tilde{H}_S$ ), which is given by the following:

$$\tilde{H}_S = w_{HD} \cdot \hat{H}_S + w_{HI} \cdot (\hat{B}_S + \hat{H}_B) + w_{HP} \cdot H_{SP} \quad (2.6)$$

with

$$w_{HD} = \frac{1/\sigma_{HS}^2}{1/(\sigma_{BS}^2 + \sigma_{HB}^2) + 1/\sigma_{HS}^2 + 1/\sigma_{HSP}^2}, \quad (2.7)$$

$$w_{HI} = \frac{1/(\sigma_{BS}^2 + 1/\sigma_{HB}^2)}{1/(\sigma_{BS}^2 + \sigma_{HB}^2) + 1/\sigma_{HS}^2 + 1/\sigma_{HSP}^2} \quad (2.8)$$

and

$$w_{HP} = \frac{1/\sigma_{HSP}^2}{1/(\sigma_{BS}^2 + \sigma_{HB}^2) + 1/\sigma_{HS}^2 + 1/\sigma_{HSP}^2} \quad (2.9)$$

In this equation,  $w_{HD}$ ,  $w_{HI}$ , and  $w_{HP}$  (which add up to one) represent the weights of the direct and indirect pathways and the prior, respectively, which are proportional to the relative precision of the sensory signals and the width of the prior. [Equation](#)

[2.6](#) would result in an accurate estimate of  $\tilde{H}_S$ , if all three pathways were accurate by themselves. However, because the prior is centred on zero ( $H_{SP} = 0^\circ$ ), it introduces more and more bias toward upright, as head tilt increases further. Thus, optimization in terms of variance has a downside by causing underestimation of the actual head tilt. The amount of underestimation depends on the width of the prior and the reliability of the sensory inputs.

The variance in the head-in-space estimates, measured across many trials,  $\sigma^2(\tilde{H}_S)$ , can be derived directly from [Equation 2.6](#) by applying the rules of error propagation (see [Appendix](#) for complete derivation). From these calculations, it follows that

$$\sigma^2(\tilde{H}_S) = w_{HD}^2 \cdot \sigma_{HS}^2 + w_{HI}^2 \cdot (\sigma_{BS}^2 + \sigma_{HB}^2), \quad (2.10)$$

in which the variance contributions of the direct and indirect pathways are represented by their squared weights. Although it does not appear explicitly in [Equation 2.10](#), it is important to notice that the prior has a noise-reducing effect by downscaling the sensory-related weighting terms ( $w_{HD}$  and  $w_{HI}$ ). The narrower the prior, the larger its relative weight ( $w_{HP}$ ) and the smaller the sensory weights, because  $w_{HD} + w_{HI} + w_{HP} = 1$ . Thus, the effect of the head-in-space prior is twofold: it reduces the variance, but as noticed above, this occurs at the cost of a bias in the final estimate of head-in-space orientation, which becomes pronounced at large tilts (see [Appendix](#) for further details).

Previously, we have shown that to account for the typical nonlinear increase of the systematic SVV errors with tilt, the variability of the head-tilt signal in the model must increase with tilt angle (de Vrijer et al., [2008, 2009](#)). In line with this conclusion, decreasing effectiveness of the otoliths with increasing tilt has been suggested by various other reports (Schöne & Udo de Haes, [1968](#); Tarnutzer, Bockisch, Straumann & Olasagasti, [2009](#); Tarnutzer, Bockisch & Straumann, [2010](#)) and may reflect the geometry of otolith organs, the nonuniform distribution of otolith afferents in the roll-plane and nonlinear firing rates (Tarnutzer et al., [2010](#)). This feature was incorporated by allowing the noise in the sensory head-tilt signal,  $\sigma_{HS}$ , to increase rectilinearly with tilt angle:

$$\sigma_{HS} = a_{HS}|H_S| + b_{HS} \quad (2.11)$$

in which  $a_{HS}$  reflects the proportional increase of noise with tilt angle and  $b_{HS}$  represents the noise at  $H_S = 0^\circ$ . Note that, in the data fits, parameter  $a_{HS}$  was allowed to be zero, so that the present model did not force  $\sigma_{HS}$  to depend on head tilt.

To compute the SVV, the brain not only requires an estimate of head orientation in space ( $\tilde{H}_S$ ), but also needs estimates of eye-in-head orientation ( $\tilde{E}_H$ ) and retinal line orientation ( $\tilde{L}_E$ ). Together, these signals determine the orientation of a visual line in space ( $\tilde{L}_S$ ) according to  $\tilde{L}_S = \tilde{H}_S + \tilde{E}_H + \tilde{L}_E$ . The systematic error in the SVV experiment ( $\Delta SVV$ ) corresponds to the error in  $\tilde{L}_S$  and is thus given by  $\Delta SVV = \Delta H_S + \Delta E_H + \Delta L_E$ , in which  $\Delta$  denotes the bias in each estimate. For simplicity, we assumed that the visual signal representing retinal line orientation is accurate, so that  $\Delta L_E = 0^\circ$ . As explained in a previous study (de Vrijer et al., 2009), underestimation of eye torsion causes errors in the eye-in-head estimate ( $\Delta E_H$ ), which can be represented by  $\Delta E_H = -A_{OCR} \cdot \sin(\hat{H}_S)$ , where parameter  $A_{OCR}$  denotes uncompensated ocular counterroll. Finally, the error in the head-in-space estimate is obtained by subtracting  $\hat{H}_S$  (see Equation 2.6) from the actual head tilt  $H_S$ , which ultimately leads to the following relation for the mean SVV error,  $\mu(\Delta SVV)$ , in multiple trials:

$$\mu(\Delta SVV) = (1 - w_{HD} - w_{HI}) \cdot H_S - A_{OCR} \cdot \sin(H_S) \quad (2.12)$$

In Equation 2.12, the influence of the prior works through the weight factors  $w_{HD}$  and  $w_{HI}$ . Because these weights do not add up to 1 (see above,  $w_{HD} + w_{HI} = 1 - w_{HP} < 1$ ), the result is a systematic error in the head-in-space estimate, which becomes more pronounced at large tilt. The noise level in the eye-in-head and line-on-eye estimates is probably relatively small compared with the noise in the head-in-space estimate considering results from Vandebussche et al. (1986), who reported just-noticeable difference levels for orientation discrimination of  $<1^\circ$ . Given this low value, SVV variance is determined mainly by the variance in the latter estimate, so that  $\sigma^2(\Delta SVV) \sim \sigma^2(\tilde{H}_S)$ .

## 2.2.6 Model fitting

The model contains seven fit parameters ( $a_{HS}$ ,  $b_{HS}$ ,  $\sigma_{HSP}$ ,  $\sigma_{BS}$ ,  $\sigma_{HB}$ ,  $A_{OCR}$ , and  $\lambda$ ) that were fitted to all data (SBT and SVV) simultaneously for each subject. As stated earlier, parameters  $a_{HS}$  and  $b_{HS}$  represent the increase and offset of sensory noise in the head-in-space estimate, respectively. The parameter  $\sigma_{HSP}$  denotes the width of the prior distribution, reflecting a priori knowledge about head-in-space. Noise levels in the body and neck sensors are represented by parameters  $\sigma_{BS}$  and  $\sigma_{HB}$ . Finally, the amplitude of uncompensated ocular counterroll is denoted by  $A_{OCR}$ . In addition to these six parameters related to sensory processing, there is a seventh parameter to account for lapses ( $\lambda$ ).

In addition to these “parameters of interest”, the data were preprocessed before model fitting by applying mean correction (McGuire & Sabes, 2009). Mean correction was performed to remove systematic errors in the SBT and the asymmetries in the SVV between CW and CCW tilt angles. Because the model is inherently left-right symmetric, it would try to account for differences in SVV bias between equal but opposite tilt angles by falsely increasing the variance. Likewise, because the model assumes that there is no bias in the SBT, it would try to explain any slight deviation from zero by excessively increasing the variance. The asymmetry in the SVV, if any, and a nonzero SBT bias, if any, are captured by fixed parameters of non-interest ( $n = 9$ ) in the model fits. Thus, for the SBT data, one bias correction term was needed for each tilt angle (yielding five parameters of non-interest), and for the SVV data, one correction parameter was needed for each pair of equal but opposite tilt angles (yielding four parameters of non-interest). We emphasise that the nine parameters of non-interest are not free-fit parameters because they are not optimised by the model. So, although technically our number of free parameters amounts to a total of 16, only seven were determined by fitting the model.

In total, the seven free parameters of the model had to account for 149 data points, spread across various tilt angles, with each data point reflecting a proportion of CW responses based on either 14 (for the SBT) or 11 (for the SVV) experimental forced-choice CW/CCW responses. We fitted the model by maximizing the likelihood of the data [maximum likelihood estimation (MLE)], in relation to the set of six model

parameters ( $a_{HS}$ ,  $b_{HS}$ ,  $\sigma_{HSP}$ ,  $\sigma_{BS}$ ,  $\sigma_{HB}$ , and  $AOCR$ ) and lapse rate ( $\lambda$ ). Optimal parameter values were obtained by minimizing the negative likelihood function using the Matlab function “fmincon” (de Vrijer et al., 2008; McGuire & Sabes, 2009). Simulations confirmed that the inverse modelling approach was not sensitive to overfitting. SDs of the best-fit parameters were obtained by performing 1999 bootstrap runs. For each run, we constructed 149 data points (reflecting the size of the dataset), each of which was obtained by random sampling with replacement from the original dataset. The model was fit to this new dataset. The distribution of model parameters across all runs was used to derive the 68.2% confidence interval of each parameter.

We emphasise that the model fit provided an estimate of the proprioceptive variance of the neck ( $\sigma_{HB}$ ), even though the head-on-body signal was not directly manipulated during the experiment. Nevertheless, this signal, as sensed by the neck proprioceptors, is essential to implement the reference frame transformation from the body-in-space to the head-in-space signal, and vice versa. Because the neck signal is noisy, these reference frame transformations induce neck-related noise in the original  $B_S$  and  $H_S$  signals, even when the head and body are aligned. Because the SBT and SVV tasks require different reference frame transformations, they depend differently on the noise properties of the three sensory systems (body and neck sensors and otoliths). By solving the inverse problem, the noise properties of the three sensory systems, as well as the other fit parameters, can be determined. Finally, we note that the inverse problem can only be solved using both tasks at multiple tilt angles; just using a single task (SVV or SBT, not both) would have made this problem intractable.

### 2.2.7 Model evaluation

To assess the importance of cross-modal sensory integration, we also fitted our model without the indirect, cross-modal pathways by setting the head-on-body noise to infinity, which effectively eliminates the indirect pathways and removes one degree of freedom. To compare the maximum-likelihood estimates from the full and the reduced model, we used a log-likelihood ratio test. The test statistic is two times the difference between the negative log-likelihoods of the data, given the reduced and the full model. A  $\chi^2$  test with one free parameter (the difference in degrees of freedom between the

two models) is used to calculate the  $p$  value (Dobson, 2001).

Furthermore, we evaluated our mechanistic model in comparison with a pure descriptive model of the same dataset based on separate psychometric accounts, each with three free parameters, at the five SBT and nine SVV angles. We used the Bayesian information criterion (BIC) for model comparison. BIC provides a measure of the adequacy of the model fit and corrects for the number of parameters. The BIC is defined as  $BIC = -2 \log(L) + k \cdot \log(n)$ , in which  $L$  is the total likelihood of the data given the model,  $k$  the number of free parameters, and  $n$  the number of data points to be explained. The number of free parameters is 42 [14 psychometric curves  $\times$  3 parameters ( $\mu$ ,  $\sigma$ , and  $\lambda$ )] for the psychometric curves, whereas for the mechanistic Bayesian model, the number of free parameters is seven. A more appropriate model is characterised by a lower BIC value.

### 2.2.8 Model validation: independent test of neck noise

The SBT and SVV measurements to test the model proposed in Figure 2.1 have yielded solutions for the noise properties of the involved sensor systems. To validate the model structure and the noise predictions that were obtained, we also devised an experiment that independently measured the noise in the neck sensors (head-on-body sensors), in a psychometric fashion. In this experiment, subjects were lying on a bed, in supine position, with their head fixed on a rotating platform. The platform was constructed such that it could passively rotate the head relative to the body, in the roll plane, while accounting for the shifting rotation axis in the neck vertebrae. The rotation of the platform was computer controlled, keeping the speed below  $0.2^\circ/\text{s}$ , which is far below detection threshold of the canals ( $> 0.5^\circ/\text{s}$ ) (Benson, Hutt & Brown, 1989). In the supine condition, there is no gravity modulation of the otolith signal, so we excluded the contribution of the vestibular system in detecting head-on-body orientation and were only probing the role of the neck afferents. We applied the method of constant stimuli, using a set of 11 head angles relative to body midline. Test angles ranged from  $-6^\circ$  to  $6^\circ$ .

In complete darkness, subjects were first rotated at a constant angular velocity of  $\leq 15^\circ/\text{s}$  to a random detour angle similar to the idea shown in Figure 2.2. The

head then moved to a randomly chosen position within the test range with a very slow speed ( $< 0.2^\circ/\text{s}$ ) such that the test angle was reached within 20 s. Meanwhile, auditory white noise was presented to the subjects through earphones to mask any auditory cues generated by the moving platform. After arrival at the test angle, the auditory noise was interrupted, signalling the subject to indicate whether head-on-body orientation was CW or CCW relative to the body midline, using a toggle switch. The subject was then rotated at a constant velocity to a new randomly chosen detour angle, and the above procedure was repeated. Each test angle was repeated 10 times, yielding a total number of 110 responses in each subject. Psychometric data were quantified by fitting a cumulative Gaussian function (see above, [Equation 2.1](#)). The width of the curve,  $\sigma^2$ , inversely related to precision, serves as an independent measure of the subject's variability of the head-on-body estimate and was compared with the model prediction.

### 2.2.9 Model simulation of patient data

Based on the average parameter values of the model established in normal, healthy subjects, the model was also used to make predictions about SVV and SBT performance in two patient groups: bilateral vestibular patients and patients with somatosensory loss. The model simulated SVV and SBT in these patient groups by raising the variance values of the lost signals to infinity.

## 2.3 Results

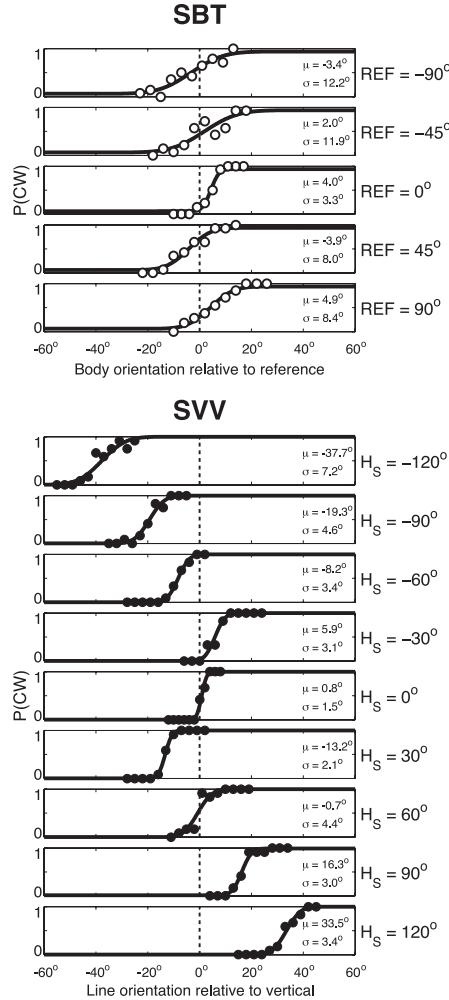
### 2.3.1 Psychometric results

The SBT experiment, performed in seven subjects, tested the accuracy and precision of SBT percepts, near upright (SBT<sub>0</sub>), at  $45^\circ$  and  $90^\circ$  right side down (SBT<sub>45</sub> and SBT<sub>90</sub>), and at  $45^\circ$  and  $90^\circ$  left side down (SBT<sub>-45</sub> and SBT<sub>-90</sub>). The same subjects also performed the SVV experiment, tested at nine roll-tilt angles, ranging from  $-120^\circ$  to  $120^\circ$  in  $30^\circ$  intervals. [Figure 2.3](#) shows the results of a typical subject (S1) in both tasks. The top panels show the proportion of CW responses for the five SBT

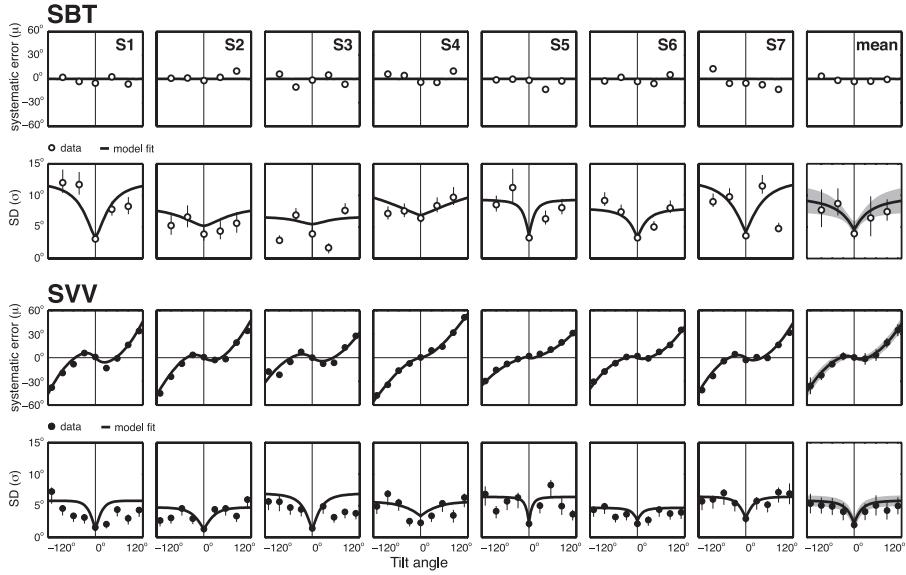
tasks, relative to the reference orientation. For an ideal observer, all psychometric functions would resemble a step centred at zero. Across the five reference orientations ( $0^\circ$ ,  $\pm 45^\circ$ , or  $\pm 90^\circ$ ), the psychometric data indicate underestimations and overestimations of perceived body angle, but no consistent bias, which resembles previous reports (Mittelstaedt, 1983; Mast & Jarchow, 1996; Jarchow & Mast, 1999; van Beuzekom et al., 2001) that body-tilt perception is accurate on average. We fitted psychometric curves through these data (see Materials and Methods, Equation 2.1), to obtain estimates for the mean ( $\mu$ ), SD ( $\sigma$ ), and lapse rate ( $\lambda$ ). Parameter  $\mu$  is a measure for the accuracy of the subject's body-tilt percept. Perceptual variability, inversely related to precision, is reflected by  $\sigma^2$ , whereas the lapse rate ( $\lambda = 0.06$ ) accounts for stimulus-independent errors (Wichmann & Hill, 2001a). In all five SBT tasks, the  $\mu$  values are relatively close to the veridical reference orientation ( $0^\circ$ ,  $45^\circ$ , or  $90^\circ$ ), i.e., errors are  $<5^\circ$ . The psychometric fits further show that variability is lower in the  $SBT_0$  task, with  $\sigma \approx 4^\circ$ , than in the  $SBT_{\pm 45}$  and  $SBT_{\pm 90}$  task, where  $\sigma \approx 10^\circ$ .

The bottom section of Figure 2.3 illustrates the performance of the same subject in the SVV task. Each panel demonstrates how the fraction of CW responses changes as a function of line orientation relative to the perceived vertical, for each tilt angle tested. Performance is very accurate in the upright condition. For moderate body tilts, i.e.,  $30^\circ$ , this subject shows small systematic errors, indicating that the line must be set in a direction opposite to the head tilt to be perceived vertical in space. For larger tilts ( $\geq 60^\circ$ ), systematic errors occur with increasing tilt angle, with amplitudes up to  $40^\circ$ , as if tilt is underestimated. This response pattern is consistent with previous literature (Aubert, 1861; Udo de Haes, 1970; Mittelstaedt, 1983; van Beuzekom & van Gisbergen, 2000). Close scrutiny also reveals that the precision in the vertical percept deteriorates away from the upright position.

Psychometric fits capture these observations. In the upright position, the percept of visual vertical is virtually unbiased, as indicated by a  $\mu$  value of  $0.8^\circ$ . At large tilts, e.g., at  $-120^\circ$  and  $120^\circ$ ,  $\mu = -37.7^\circ$  and  $\mu = 33.5^\circ$ , respectively, which means that the line must be tilted away from true vertical to be perceived as vertical in space. Furthermore, the fitted psychometric curves are steepest at  $0^\circ$  tilt, reflected by  $\sigma = 1.5^\circ$ . With larger tilt angles,  $\sigma$  increases, reaching maximum values of  $7.2^\circ$  and  $4.3^\circ$  at tilts of  $-120^\circ$  and  $+120^\circ$ , respectively.



**Figure 2.3:** SBT versus SVV performance in one subject (S1). Top, SBT. Proportion of CW responses is plotted against body orientation relative to the reference orientation ( $0^\circ$ ,  $\pm 45^\circ$ , or  $\pm 90^\circ$ ).  $\mu > 0^\circ$  indicates tilt underestimation. Bottom, SVV. Proportion of CW responses is plotted against line orientation relative to vertical. Solid lines, Best-fit cumulative Gaussians, typified by  $\mu$  and  $\sigma$ .



**Figure 2.4:** Model predictions superimposed on parameters from the psychometric fits to the SBT (two top rows) and SVV data (two bottom rows). Accuracy and variability characteristics as a function of roll-tilt angle are shown; values are psychometric fits ( $\mu$  and  $\sigma$  values,  $\circ$ ) and model predictions (line) from all subjects. Mean data and mean predictions across subjects are plotted in the rightmost column.

The results of this subject are exemplary for all subjects, as shown by the bias and SD data points in Figure 2.4. The mean results across the seven subjects are shown in the rightmost column. The bold lines in Figure 2.4 represent the fits from our Bayesian model, which will be discussed later in this section.

The two top rows in Figure 2.4 show the accuracy ( $\mu$ ) and precision ( $\sigma$ ) of SBT percepts, now plotted against body orientation. For each subject, these values ( $\mu$  and  $\sigma$ ) were derived from the fitted cumulative Gaussian curves (Figure 2.3). Biases for SBT, shown in the top row of Figure 2.4, indicate moderate deviations in either direction from perfect performance, but no systematic pattern emerges. Across the seven subjects, the  $\mu$  values ranged from  $-14.2^\circ$  to  $+11.7^\circ$  across the five SBT tasks. On average, however, there was no systematic bias for the five body orientations (ANOVA;  $F(4,24) = 1.4$ ;  $p = 0.25$ ), as also indicated by the rightmost panels. The data further show that, in all subjects, variability is statistically lower ( $p < 0.05$ ) at the upright orientation, with  $\sigma$  values  $< 4^\circ$ , than in the tilted conditions ( $45^\circ$  and  $90^\circ$ ), with  $\sigma$  values ranging up to  $12^\circ$ .

The two bottom rows of Figure 2.4 summarise our SVV data across the entire tilt range. Accuracy is close to perfect at upright orientation in all subjects, with mean

values ranging between  $0.1^\circ$  and  $2.8^\circ$ . For tilts  $\geq 60^\circ$ , all subjects show systematic SVV errors (biases) of undercompensation, ranging up to maximum values close to  $60^\circ$ . Three subjects (S1, S2, and S3) also show slight errors of overcompensation in the smallest tilt range ( $<60^\circ$ ). The variability in the SVV is  $<3.0^\circ$  for upright, which is consistently lower than in the tilted conditions, where variability reaches values ranging up to  $8^\circ$ .

Together, the results in [Figure 2.4](#) show that SVV and SBT have different accuracy and precision characteristics. Subjects perceive their body-tilt angle more accurately than the spatial orientation of the visual line. However, when it comes to variability, performance is reversed: SVV curves are narrower than the SBT curves in all subjects, at both tilt angles, meaning that they are consistently less variable in the SVV task than in the SBT task.

### 2.3.2 Model predictions

The bold lines in [Figure 2.4](#) present the predictions of the Bayesian integration model, fitted simultaneously to the original responses from the SBT and SVV tasks. The right-most column of [Figure 2.4](#) shows the mean predictions from this model superimposed on the averaged parameters from the psychometric fits, indicating that the sensory integration model can account very well for all the characteristics of the data.

By design (see [Materials and Methods](#), the sensory integration model fits a horizontal fit line through  $\mu_{SBT} = 0^\circ$  because it cannot account for the small systematic SBT errors. As to SBT precision, the model predictions show an increase of noise with tilt angle similar to the actual increase of noise between  $0^\circ$  and  $90^\circ$  tilt, for all subjects. These model fits further suggest that the increase of SBT noise is steepest at small tilt angles and levels off at larger tilts. According to the model, the increase of SBT noise with tilt angle is attributable to the corresponding increase of noise in the head sensors (parameter  $a_{HS}$ ), but levels off by the constant noise level in the body sensors. The third row in [Figure 2.4](#) depicts the model predictions of the systematic SVV errors, which show a very good match. Also with respect to SVV variability, fits and data show similar trends, suggesting an increase of SVV noise with tilt angle, which levels off at larger tilts.

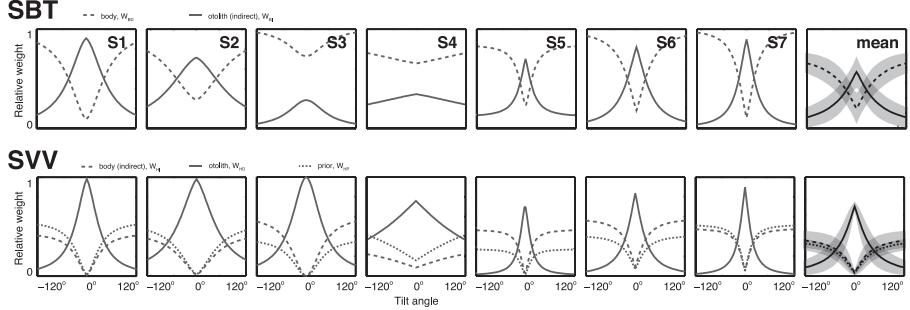
Subject	$a_{HS}$ ( $^{\circ}/^{\circ}$ )	$b_{HS}$ ( $^{\circ}$ )	$\sigma_{HSP}$ ( $^{\circ}$ )	$\sigma_{BS}$ ( $^{\circ}$ )	$\sigma_{HB}$ ( $^{\circ}$ )	$A_{OCR}$ ( $^{\circ}$ )
S1	0.23 $\pm$ 0.02	1.2 $\pm$ 0.32	11.6 $\pm$ 1.0	12.3 $\pm$ 1.1	3.3 $\pm$ 1.2	27.0 $\pm$ 2.2
S2	0.12 $\pm$ 0.02	1.2 $\pm$ 0.52	9.4 $\pm$ 1.1	8.4 $\pm$ 2.9	6.4 $\pm$ 4.1	17.0 $\pm$ 3.8
S3	0.20 $\pm$ 0.03	1.1 $\pm$ 0.42	14.4 $\pm$ 1.7	6.7 $\pm$ 1.9	9.3 $\pm$ 2.4	17.5 $\pm$ 2.1
S4	0.07 $\pm$ 0.50	3.9 $\pm$ n/a	11.2 $\pm$ 1.3	12.6 $\pm$ 2.3	7.1 $\pm$ 3.5	0 $\pm$ n/a
S5	0.11 $\pm$ n/a	3.3 $\pm$ 1.0	18.7 $\pm$ 4.8	15.0 $\pm$ n/a	3.6 $\pm$ 2.1	1.06 $\pm$ n/a
S6	0.23 $\pm$ 0.09	3.0 $\pm$ 1.5	9.5 $\pm$ 1.1	8.0 $\pm$ 0.83	1.8 $\pm$ n/a	18.8 $\pm$ 4.1
S7	0.20 $\pm$ 0.14	3.2 $\pm$ 1.0	12.8 $\pm$ 2.4	12.7 $\pm$ 6.1	3.0 $\pm$ n/a	20.8 $\pm$ 9.0
Mean	0.16 $\pm$ 0.06	2.4 $\pm$ 1.2	12.5 $\pm$ 3.2	10.8 $\pm$ 3.1	4.9 $\pm$ 2.7	14.6 $\pm$ 10.2

**Table 2.1:** Best-fit parameter and bootstrap-based SD values. Imposed fit limits were as follows:  $a_{HS}$ :  $0.5^{\circ}/^{\circ}$ ;  $b_{HS}$ ,  $\sigma_{HSP}$ ,  $\sigma_{BS}$ ,  $\sigma_{HB}$ ,  $50^{\circ}$ ;  $A_{OCR}$ ,  $30^{\circ}$ . SD values are not shown (n/a) when bootstrapped values formed a skewed distribution.  $a_{HS}$ , Tilt-related increase in otolith noise;  $b_{HS}$ , otolith noise in upright position;  $\sigma_{HSP}$ , width of head-tilt prior;  $\sigma_{BS}$ , noise in body-in-space sensors;  $\sigma_{HB}$ , noise in neck sensors;  $A_{OCR}$ , uncompensated ocular counterroll.

For each subject, best-fit parameter values and their bootstrap-based SD levels are listed in [Table 2.1](#). Parameter  $b_{HS}$ , representing the noise ( $\sigma_{HS}$ ) in the otolith signal in the upright subject, ranges between  $1.1^{\circ}$  and  $3.9^{\circ}$ . Best-fit values of parameter  $a_{HS}$  are significantly positive ( $p < 0.05$ ) for all subjects, ranging from  $0.07^{\circ}/^{\circ}$  (S4) to  $0.23^{\circ}/^{\circ}$  (S1). This implies that the noise in the otoliths increases with tilt angle. The width of the head-in-space prior ( $\sigma_{HSP}$ ) ranges from  $9.4^{\circ}$  (S2) to  $18.7^{\circ}$  (S5), with a mean of  $12.5 \pm 3.2^{\circ}$ , consistent with our previous report (de Vrijer et al., [2009](#)). Best-fit values of parameter  $\sigma_{BS}$ , reflecting the noise in the sensory body-in-space signal, range from  $6.7^{\circ}$  (S3) to  $15.0^{\circ}$  (S5), with a mean of  $10.8 \pm 3.1^{\circ}$ , which is about twice as large as the best-fit values of parameter  $\sigma_{HB}$ , reflecting noise in the head-on-body signal, ranging from  $1.8^{\circ}$  (S6) to  $9.3^{\circ}$  (S3), with a mean of  $4.9 \pm 2.7^{\circ}$ . Thus, the parameter fits imply that the neck sensors are more precise than the body-tilt sensors. As has been discussed extensively in our previous paper (de Vrijer et al., [2009](#)), the amplitude of uncompensated ocular counterroll ( $A_{OCR}$ ) shows large inter-subject variability.

### 2.3.3 Sensory weights

To obtain the model fits in [Figure 2.4](#), we made the assumption (see Introduction) that information from both direct and indirect pathways ([Figure 2.1](#)) is used to estimate body and head orientation in space. The sensory weights, indicating the relative contribution of both pathways, can be computed from the fit results in [Table 2.1](#). To obtain the body-in-space estimate, necessary for the SBT, the model uses both direct information from the body sensors and indirect information from the combination of otolith and neck information. Because the variability of the otolith signal ( $\sigma_{HS}$ ) increases with tilt angle ( $a_{HS} > 0$ ), as shown in [Table 2.1](#), the relative importance of



**Figure 2.5:** Tilt dependence of weight factors in SBT (top) and SVV (bottom) for each subject. Trends with tilt angle are similar for all subjects. Head-in-space prior is only involved in SVV computations. Means across subjects are plotted in the rightmost column.

direct and indirect pathways becomes dependent on tilt angle. This can be seen in Figure 2.5 (top row), which shows the relative weights of these signals for each subject, derived from Equation 2.2 and the best-fit parameter values in Table 2.1. The mean ( $\pm SD$ ) pattern across subjects is shown in the rightmost pattern.

Instead of an overall dominance of body receptors in the direct pathway, the model implies that it is actually the indirect pathway, carrying the otolith signal, that dominates the behaviourally important range near upright. In most of our subjects (S1, S3, S5, S6, and S7), it was only when the otoliths became less reliable, at larger tilts, that the body sensors (direct pathway) got the upper hand ( $w_{BD} > 0.5$ ).

For the SVV task, the model assumes that both information from the otoliths (direct) and the combination of body and neck information (indirect) is used. Figure 2.5 (bottom row) illustrates the relative contributions from these sensors as well as from the prior, based on the model fits (Table 2.1) and Equation 2.6. The SVV pattern looks similar to the SBT pattern (Figure 2.5, top row): in all subjects, the otoliths are very dominant near upright, with weights close to 1, but their contribution declines when tilt angle increases. As we saw for the SBT signal weights, this decline reflects increasing otolith noise levels. In the SVV, the decline is steeper than in the SBT, where the reference frame transformation leads to an enhanced noise level with a less pronounced tilt dependence. As the otolith contribution decays, the contributions of the prior and indirect pathway become more manifest. According to our model fits, the weight of the body sensors in the SVV task ( $w_{HI}$ ) at 90° tilt ranges between 0.19 (S4) and 0.53 (S6).

Subject	MLE full model	MLE reduced model	<i>p</i>	BIC full model	BIC psychometric fits
S1	231.5	312.4	<0.001	498.1	540.3
S2	197.8	216.9	<0.001	430.6	539.3
S3	267.5	313.4	<0.001	570.0	548.2
S4	207.5	218.2	<0.001	450.0	569.1
S5	195.6	216.3	<0.001	426.2	571.0
S6	163.5	183.2	<0.001	362.0	502.5
S7	248.0	260.0	<0.001	531.1	633.5

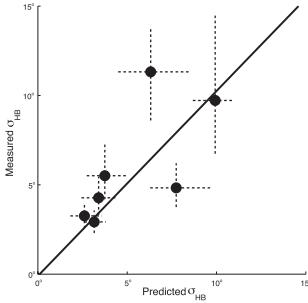
**Table 2.2:** Validation of the model. Log-likelihood ratio test of the full model (with indirect pathways) and reduced model (without indirect pathways). For all subjects, the full model outperforms the reduced model lacking multisensory integration through the indirect pathways. BIC values are much lower for the Bayesian integration model compared with separate psychometric fits in six of seven subjects.

### 2.3.4 Model evaluation

To test whether the assumption of indirect pathways in the model is warranted, we compared its performance with a reduced version with only direct pathways (see [Materials and Methods](#)). With this in mind, we performed a likelihood ratio test of the complete model fit (with direct and indirect pathways) versus the fit of a model with direct pathways only [i.e., SVV just based on head sensors (the otoliths), the SBT just based on body sensors]. The results are shown in [Table 2.2](#). For each subject, the complete model provided a significantly better account of the data than the reduced model without multisensory integration through the indirect pathways. In other words, head, neck and body sensors all contribute to both SBT and SVV.

We also compared our model, which provides a mechanistic explanation of the full dataset, with the pure descriptive account of the data as obtained by fitting separate psychometric curves to the data for the five SBT angles and nine SVV angles ([Equation 2.1](#), [Figure 2.3](#), and [Materials and Methods](#), Model evaluation). Maximum-likelihood estimates were calculated and corrected for the number of free parameters using the BIC. As shown in [Table 2.2](#), we found the lowest BIC values, indicating a better model, for the Bayesian model in all subjects, except S3. One might argue that the mean corrections before fitting the Bayesian model added another nine parameters that should be taken into account when comparing the models, even though these parameters were not fitted by the model. However, even for the worst-case scenario of 16 parameters, our Bayesian model still outperformed the individual psychometric fits ( $\text{BIC}_{\text{mechanistic}} = 3583 < \text{BIC}_{\text{psychometric}} = 3910$ ).

**Figure 2.6:** Model validation. Independent measurement of neck (head-on-body) noise versus the values predicted by the model. The dots represent the median values and the dashed lines the 95% confidence interval determined from a bootstrap. Note that the variance of the estimates increases with the mean value. The solid line shows the regression based on log-transformed data (slope, 1.03;  $p = 0.04$ ).



### 2.3.5 Model validation

To further validate the model, we independently tested one of its predictions that can be assessed experimentally in isolation: head-on-body variance. In supine position, subjects judged their head orientation (CW/CCW) relative to the body midline after it had been passively roll-rotated with speeds subthreshold for the canals to various angles (see [Model validation: independent test of neck noise](#)). Psychometric fits indicate no systematic bias in these head-on-body percepts (data not shown). Figure 2.6 depicts the experimental noise levels derived from these psychometric fits and the predicted values provided by the model, including their 95% confidence intervals. Because the variance of the estimates increases with the average head-on-body percept, we performed a regression on the log-transformed data (Hopkins, [2011](#)). The significant correlation between predicted and measured neck noise levels (slope, 1.03; intercept, -0.1;  $p = 0.04$ ) provides an independent confirmation of the proposed model.

## 2.4 Discussion

In this study, we made intra-subject comparisons of the accuracy (bias) and precision (inverse variance) characteristics in two spatial orientation tasks: SBT and SVV. The main experimental findings were as follows: (1) the SBT is more accurate than the SVV, (2) the SBT is less precise than the SVV, and (3) both SBT and SVV precision are smaller in tilted conditions than near upright. Under the assumption of optimality, a Bayesian model of sensory integration could account very well for these findings. Independent measurements, in supine subjects, of head-on-body variance confirmed

the predicted value.

### 2.4.1 Comparison with previous work

A world-vertical visual line appears tilted in space when the head is tilted in a darkened room (Aubert, 1861). Mittelstaedt (1983) was the first to emphasise that this phenomenon cannot be explained by errors in the body-tilt percept. He showed that subjects could accurately adjust themselves to a horizontal position, but, once in this position, made substantial systematic errors in the perception of visual verticality. Later, combined tests confirmed the discrepancy between SVV and SBT accuracy (Mast & Jarchow, 1996; Jarchow & Mast, 1999; van Beuzekom & van Gisbergen, 2000; van Beuzekom et al., 2001; Kaptein & van Gisbergen, 2004; Vingerhoets, Medendorp & van Gisbergen, 2008). The present study is consistent with these findings, showing substantial systematic SVV errors at tilts  $\geq 60^\circ$  and fairly accurate SBT performance.

Compared with the abundant literature on SBT and SVV accuracy, data on their perceptual variability are scarce. In contrast to Mittelstaedt's observation (1983), Mast and Jarchow (1996) found that the SBT was much more variable than the SVV. The present study, the first to measure both SVV and SBT precision using an extensive psychometric approach, has clearly established that SBT variability is consistently higher than SVV variability, both in the upright and in the horizontal ( $90^\circ$ ) tilt position.

Furthermore, although various studies have noted that SVV variability increased at larger tilts (Schöne, 1964; Schöne & Udo de Haes, 1968; Udo de Haes, 1970; van Beuzekom et al., 2001; de Vrijer et al., 2008), little is known about SBT variability as a function of tilt angle. Nelson (1968) showed that subjects were more variable when adjusting themselves to a horizontal position than to a vertical (upright) position. The present findings are consistent with these early observations.

## 2.4.2 Implications of the model

After earlier indications that both the otoliths and body sensors contribute to the SBT (Clark & Graybiel, 1963, 1964; Nelson, 1968), Mittelstaedt (1997) made a quantitative assessment of their impact, using an ingeniously designed experiment. Subjects lay on their side in a horizontal centrifuge. The crux of the experiment was to vary the distance between the rotation axis and the interaural axis to equalise the opposite contributions from the otoliths and the body sensors so that the subject felt horizontal. By testing normal, paraplegic, and nephrectomised subjects, Mittelstaedt inferred how much each sensory system contributes to body-tilt perception. It was shown that, apart from the otoliths, also internal “graviceptors” in the trunk (such as the viscera) participate in the computation of the SBT. Later, some related studies provided evidence that the distribution of blood in the body also affects postural perception (Vaitl et al., 1997, 2002). According to Mittelstaedt (1998), the weight of the somatic graviceptors to estimate horizontal body orientation in healthy subjects is ~0.6 on average, with considerable inter-subject variability. His estimate seems quite compatible with our  $w_{BD}$  values at 90° tilt, which range between 0.35 (S5) and 0.93 (S6) (Figure 2.5).

Previous attempts to identify the separate contributions of the otoliths, neck, and body sensors on the SVV have yielded mixed results. Whereas Mittelstaedt (1998) found no evidence that the SVV was affected by the body sensors in his centrifuge experiment, other studies indicate that neck- and trunk-tilt aftereffects (Wade, 1968), neck muscle vibration (McKenna, Peng & Zee, 2004), and manipulation of tactile and interoceptive body cues (Trousselard et al., 2004) can affect the SVV. In other words, even in the absence of direct head-in-space information from the otoliths, the brain can still obtain an estimate of head orientation in space through the indirect sensory pathway. These findings suggest that these modalities operate together with the otoliths in the computation of the SVV, consistent with our model.

### 2.4.3 Model evaluation

The architecture of the model, as far as the reference frame transformations and the sensory integration is concerned, follows entirely from the principles of Bayesian inference. However, to account for our major findings and inter-subject differences, we made two less straightforward assumptions. First, to explain the increased variability in both tasks at 90° tilt, we allowed for the possibility that the otoliths become more noisy with increases in tilt. Second, we hypothesised that prior knowledge is used in the visual vertical but not in body-tilt perception. Can these assumptions be justified on physiological and rational grounds?

One reason to assume that otolith noise depends on tilt angle is based on the fact that the utricle contains considerably more hair cells than the saccule (Rosenhall, 1972, 1974). Because the utricle is most sensitive to tilts of ~0°, whereas the saccule is most sensitive at ~90° tilt (Jaeger, Kondrachuk & Haslwanter, 2008), this may well cause the proposed increase of otolith noise with tilt angle (Tarnutzer et al., 2010). A tilt-dependent noise level of the otoliths would also help to explain why the perturbing effect of roll-optokinetic stimulation on the SVV (Dichgans, Diener & Brandt, 1974; Fernández & Goldberg, 1976b) and on the SBT (Young, Oman & Dichgans, 1975) is more pronounced at larger tilt angles and why the SVV is more strongly influenced by residual canal signals at larger tilt angles, after prolonged roll rotations (Lorincz & Hess, 2008).

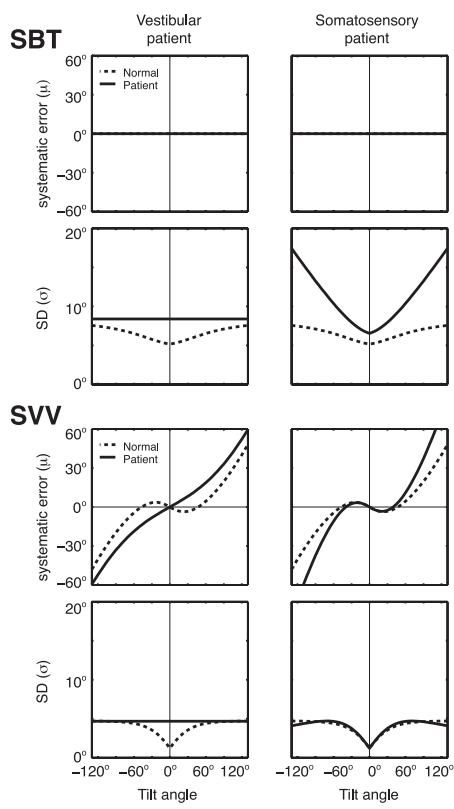
In the SVV literature, it is widely assumed that the visual vertical is determined by a weighted combination of a sensory head-tilt signal and a head-fixed reference, denoted as the idiotropic vector (Mittelstaedt, 1983). Recently, this idiotropic vector has been reinterpreted in terms of a Bayesian prior (Eggert, 1998; MacNeilage et al., 2007; de Vrijer et al., 2008), with which it is mathematically equivalent. Interestingly, when tested in gravity-free conditions, subjects still retain a sense of visual vertical, always aligned with their long-body axis, compatible with the idea of head-fixed prior (Mittelstaedt, 1983). Vingerhoets et al. (2008) recently found a similar phenomenon in the SVV during multiple-cycle dynamic roll rotation in normal gravity. Remarkably, when the same subjects were tested in a comparable dynamic SBT experiment, their responses showed very little bias on average, indicating that a prior is used only in the

SVV and not for body-tilt estimation. To explain how this difference in computational approach might make sense, Vingerhoets et al. (2008) speculated that precision is more important than accuracy for the visual system, for reasons of visual stability. Combining the sensory tilt signal with prior knowledge yields a more stable percept of visual space than can be derived from the sensory signal alone. In a recent study, Bortolami et al. (2006) report virtually no bias in the haptically indicated vertical, which would be consistent with the suggestion that the prior plays primarily a role in visual processing. Likewise, for body-tilt perception, for which it might be less important to be precise and more useful to be accurate, the prior does not take part in this process.

#### 2.4.4 Clinical implications

According to our model, statistically optimal performance requires the use of information from both direct and indirect pathways to estimate body and head orientation in space (Figure 2.1). Thus, if one of the sensory inputs is lost or severely disrupted, SBT and SVV performance will deteriorate, but should not completely break down due to their multimodality dependence. By setting the appropriate parameter values of the model to infinity, we simulated the model to predict SVV and SBT performance in two patient groups: bilateral vestibular patients (noise level of the otoliths set to infinity) and patients with somatosensory loss (noise level of the body sensors set to infinity). Figure 2.7 shows the results of these simulations.

When the otolith signal is lost (vestibular patient), the model predicts increased but constant noise levels in both the SBT and SVV, regardless of tilt angle. From the perspective of our model, the increased SBT variability can be attributed to the loss of the otolith contribution through the indirect body-in-space pathway. The increased bias in the SVV, predicted by the model, can also be understood: as the sensory-derived head-tilt estimate becomes noisier, the effect of the prior becomes more noticeable. Although there are no accuracy and variance measurements across the entire tilt range in these patients, the few previously published deficits are consistent with these predictions. Bisdorff et al. (1996) have shown that bilateral vestibular patients performed quite accurately in the SBT at upright, but were ~40% more variable than normal



**Figure 2.7:** Clinical implications of the model. The model simulates the SBT and SVV in a vestibular patient by raising the level of the otolith noise to infinity, keeping the other parameters at the mean values of [Table 2.1](#). A somatosensory patient is modelled by setting the noise level of the body sensors to infinity. Solid lines, Patient predictions. Dashed lines, Prediction for normals.

subjects. Bronstein et al. (1996) showed that vestibular patients still compensated for their tilt angle when testing the SVV at 90°, but with a bias about twice as large as in normal subjects, consistent with our simulations.

The right column of Figure 2.7 depicts a simulation of the SBT and SVV in a patient with loss of somatosensory information (somatosensory patient). In this case, the SBT depends solely on otolith information mediated through the indirect pathway. Although this signal is still accurate, it is spoiled by the larger variability of the neck signal, needed to perform the appropriate reference frame transformation. The model also predicts larger errors in the SVV in these patients than in normal subjects. Although there are no reports containing measurements of bias and variance in the SBT and SVV, paraplegic patients show that along with the otoliths, internal body sensors also contribute to the SBT if lesions are below the 12th thoracic segment (Mittelstaedt, 1997). This evidence supports the design of our model.

In conclusion, we have tested the performance of healthy subjects in two psychophysical tasks that probe two spatial orientation estimates (SBT and SVV) and show that perceptual accuracy and precision in these tasks can be linked to the reference-frame-dependent weighting of sensory signals. We verified our theoretical framework by independent measurements of neck noise levels and by showing that it can account for the stereotypical performance of two patient groups. In this respect, our reverse-engineering approach also provides a new tool to establish diagnostic and prognostic markers of the quality of the signals involved in spatial orientation in neurological disease.

## 2.5 Appendix

Here we provide further explanation about the Bayesian computations underlying the SVV as expressed in Equations 2.6 and 2.10 in Materials and Methods. Figure 2.8 illustrates graphically that the variance of the posterior distribution in a single trial ( $\sigma_{HS}^2$ ) is not simply the same as the variance in its peak location in multiple trials,  $\sigma^2(\tilde{H}_S)$ . In a single trial (Figure 2.8A-C), the optimal estimate of head tilt is based on the likelihood (Figure 2.8B, green curve) associated with the combined sensory input

from the direct and the indirect pathway ([Figure 2.8A](#), green line,  $\hat{H}_S$ ) and the prior ([Figure 2.8B](#), blue curve), by multiplication of the two probability distributions. The prior distribution is a Gaussian with mean  $H_{SP}$  and variance  $\sigma_{HSP}^2$ . The peak of the resulting posterior distribution ([Figure 2.8B-C](#), orange curve) is used as the optimal estimate of head tilt ( $\tilde{H}_S$ ), given by the following:

$$\tilde{H}_S = w_{HS} \cdot \hat{H}_S + w_{HP} \cdot H_{SP}, \quad (2.13)$$

with

$$w_{HS} = \frac{1/\sigma_{HS}^2}{1/\sigma_{HS}^2 + 1/\sigma_{HSP}^2}, \quad (2.14)$$

and

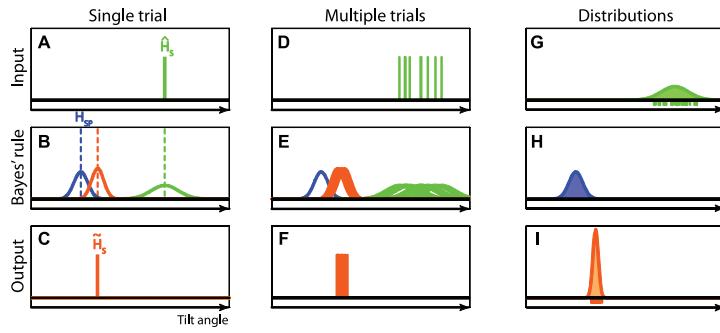
$$w_{HP} = \frac{1/\sigma_{HSP}^2}{1/\sigma_{HS}^2 + 1/\sigma_{HSP}^2}, \quad (2.15)$$

in which  $\sigma_{HS}$  denotes the noise in the sensory signal, known to the observer, and  $w_{HS}$  and  $w_{HP}$  represent the relative weights of the sensory signal and the prior, respectively. Note that [Equation 2.13](#) is equivalent to [Equation 2.1](#) in [Materials and Methods](#). The variance of the posterior distribution in a single trial is given by the following:

$$\sigma_{HS}^2 = w_{HS} \cdot \sigma_{HS}^2 = \frac{\sigma_{HSP}^2}{\sigma_{HS}^2 + \sigma_{HSP}^2} \cdot \sigma_{HS}^2 \quad (2.16)$$

and is reflected by the width of the orange curve in [Figure 2.8B](#). [Figure 2.8D-F](#) illustrates performance in multiple trials, in which the posterior distributions vary due to sensory noise ( $\sigma_{HS}$ ), whereas the prior remains fixed. The variance of each posterior distribution is fixed and is given by [Equation 2.16](#).

That the variance of the peak locations across multiple trials,  $\sigma^2(\tilde{H}_S)$ , is smaller can be shown by applying the rules of noise propagation to [Equation 2.13](#).



**Figure 2.8:** Bayesian computations in single and multiple trials. **A-C**, Single trial. **D-F**, Multiple trials. **G-I**, Resulting distributions. For further explanation, see [Appendix](#).

$$\sigma^2(\tilde{H}_S) = \left( \frac{\delta \tilde{H}_S}{\delta \hat{H}_S^2} \right) \cdot \sigma^2(\hat{H}_S) + \left( \frac{\delta \tilde{H}_S}{\delta H_{SP}^2} \right) \cdot \sigma^2(H_{SP}) = w_{HS}^2 \cdot \sigma_{HS}^2 = \frac{\sigma_{HSP}^2}{\sigma_{HS}^2 + \sigma_{HSP}^2} \cdot \sigma_{HS}^2 \quad (2.17)$$

which is equivalent to [Equation 2.10](#) in Materials and Methods. Corresponding panels G-I ([Figure 2.8](#)) illustrate the distribution of the sensory signals for a given tilt angle (green-shaded curve), the prior distribution (blue-shaded curve), and the optimal estimates (orange-shaded curve), respectively. [Figure 2.8I](#) illustrates that the distribution of the optimal estimates across many trials has a lower variance than the posterior distribution in each single trial ([Figure 2.8B](#)), which follows from the comparison of Equations [2.16](#) and [2.17](#), respectively.

---

---

***CHAPTER 3***

---

*Integration of ocular and  
vestibular signals for  
self-motion perception in  
darkness*

**Abstract** Self-motion is typically accompanied by compensatory eye movements that help minimise retinal slip and maximise dynamic visual acuity. To date, it is unknown whether these eye movements also have a reversed role, serving as a cue for self-motion perception. To address this question, we had participants ( $n = 8$ ) judge self-motion during different eye movement conditions in the absence of full-field optic flow. In a 2AFC task, participants indicated whether the second of two successive passive lateral whole-body translations was longer or shorter than the first. Eye movements during each translation were world-stationary, or body-stationary in an otherwise dark room. Results of these two conditions show that the perceived translations were shorter with body-fixed gaze compared to world-fixed gaze. Using a linear model, we estimated the relative contributions of vestibular and eye movement signals to self-motion perception and found that eye movement signals contribute approximately 25 percent. The model was independently validated by successfully predicting the effects of eye movements on self-motion in a third condition without any visual fixation, i.e. when the eyes were free to move. We conclude that eye movement signals influence self-motion perception, even in the absence of visual stimulation, and even when oculomotor and vestibular estimates are in conflict, e.g. during body-fixed gaze. We hypothesise that adverse consequences of this seemingly inflexible arrangement are minimal under natural conditions because eye movements and self-motion are highly correlated, and because eye movements are most often accompanied by veridical optic flow cues to self-motion.

---

This chapter is being revised for publication

Clemens, I.A.H., Selen, L.P.J., MacNeilage, P.R. and Medendorp, W.P. (2015a).

### 3.1 Introduction

An accurate estimate of self-motion is important to guide interactions with the environment. During passive self-motion both vestibular and optic flow signals provide information about self-motion (Gibson, Olum & Rosenblatt, 1955; Benson, Kass & Vogel, 1986; Harris, Jenkin & Zikovitz, 2000; Israël & Berthoz, 1989; Angelaki & Hess, 2005; Carriot, Brooks & Cullen, 2013; Chen, DeAngelis & Angelaki, 2010). However, also compensatory eye movements that maintain fixation on world-fixed objects carry self-motion information. These eye movements are driven by retinal slip or vestibular signals. For example, the linear vestibulo-ocular reflex (LVOR) stabilises gaze during head translations, even in complete darkness (Paige, 1989; Medendorp, van Gisbergen & Gielen, 2002; Angelaki, 2004). Many studies have shown that the brain uses oculomotor signals to extract the optic flow component related to self-motion (Warren & Hannon, 1988; Royden, Banks & Crowell, 1992; Freeman & Banks, 1998; Lappe, Bremmer & van den Berg, 1999), but to our knowledge a direct influence of eye movements on self-motion perception has not been investigated. Here we study whether these oculomotor signals are also used to estimate self-motion directly.

When gaze is world-stable during whole-body translation, the eye displacement correlates with translation size and is modulated by fixation depth (Schwarz, Busettini & Miles, 1989; Paige, Telford, Seidman & Barnes, 1998; McHenry & Angelaki, 2000; Medendorp et al., 2002). When properly scaled this eye movement signal could serve as a self-motion cue. In contrast, when fixation is body-fixed the eyes remain stationary in their orbits (Paige et al., 1998; Ramat, Straumann & Zee, 2005) making them no longer informative about self-motion. If, however, the brain assumes that eye movements are always made to maintain world-stable gaze, as in the LVOR, it would equate the absence of eye movements with the absence of self-motion. As a result, self-motion with body-fixed gaze should be underestimated compared to self-motion with world-fixed gaze, despite identical vestibular cues.

This hypothesis implies that oculomotor signals are always combined with vestibular signals to estimate self-motion, even in complete darkness. In this case, the size of the unconstrained eye movements should resemble a VOR movement that is interme-

diate between body- and world-fixed fixation, and should parametrically relate to the perceived self-motion.

To test whether eye movements are used in self-motion perception, we employed a two-alternative forced choice (2AFC) paradigm in which participants were presented with two consecutive lateral translations. They had to indicate whether the second translation was longer or shorter than the first. Eye movements during each interval were either constrained using a body or world stationary fixation point or not constrained at all (i.e. free). We show that identical translations were perceived shorter when gaze was body stationary compared to world stationary. Furthermore, using a linear model we predicted perceived displacement during the free gaze condition based on vestibular signals and unconstrained eye movements. We conclude that eye movements influence self-motion perception even in the absence of optic flow or other visual stimulation.

## 3.2 Materials and methods

### 3.2.1 Participants

Eight naive participants (three male, five female), aged between 22 and 29 years, provided written informed consent to participate in the experiment. All participants were free of any known vestibular or neurological disorder and had normal or corrected-to-normal visual acuity. Participants never received any feedback about their performance.

### 3.2.2 Experimental setup

A motorised linear sled (see Clemens, Selen, Koppen & Medendorp, 2012, for details) was used to laterally translate participants following a minimum jerk profile (Flash & Hogan, 1985) of fixed duration (1 s) and amplitudes ranging from 1 to 27 cm. Participants were seated on the sled such that the inter-aural axis aligned with the motion axis. They were restrained using a five-point seat belt and a chin rest. In

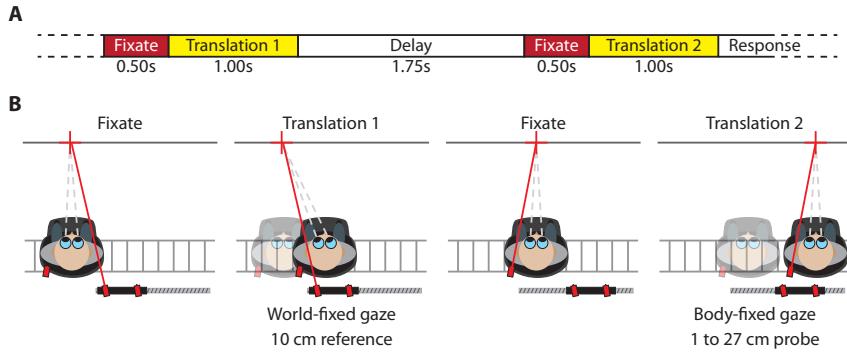
addition, the head was held in place using a sled-fixed mold which resembled headphones and pressed down on the head surrounding the pinnae. Auditory cues were suppressed using white noise presented through in-ear headphones. Experiments were conducted in complete darkness except for visual fixation points, projected by a laser pointer on a black bar 50 cm in front of the participant at eye level. Laser pointers used to project body-fixed targets were attached to the sled. Those used to project world-fixed targets were mounted on the wall behind the sled.

Eye movements were recorded at 500Hz using an EyeLink II system (SR Research, Kanata, Canada) whose cameras were mounted to the sled and therefore remained stable with respect to the head during the entire experiment. Because the head and body positions were fixed during the experiment, the orientation of the eyes within the head, as measured by the tracker, was equivalent to the orientation of the eyes in space. The eye tracking system was calibrated before each session using 11 evenly spaced calibration points ranging from -22 to 22° degrees. We used linear regression to link EyeLink measurements to gaze angles.

### 3.2.3 Paradigm

We used a two-alternative forced choice (2AFC) task to measure perceived linear self-motion across three different eye fixation types: world-fixed, body-fixed, and unconstrained (free) fixation. We refer to these as world, body and free, respectively. A trial contained two sequential translation intervals of equal duration (1 s) and in the same direction (either leftward or rightward). Different fixation types were presented in the two translation intervals. Participants were instructed to judge whether the translation during the second interval was longer or shorter compared to the first interval. They were additionally instructed to always look at the fixation point when it was visible; no instructions were given for when the fixation point was switched off (i.e. during free fixation).

The time evolution of a single trial is shown in [Figure 3.1](#). Each trial started with the onset of a central fixation point (i.e. aligned between the eyes) for 0.5 s. Subsequently, the first translation interval commenced. Depending on the fixation type, the fixation point remained visible (world and body) or was extinguished (free) during the trans-



**Figure 3.1:** **A** Time course of key events within a single trial. In each of the two intervals, a 0.50 s fixation period (red) precedes the lateral translation (yellow). A 1.75 s long delay period (shown in white) separates the two intervals. After the second translation, the participant responded whether this second translation was longer or shorter than the first. **B** Top-view illustrating key events during a body vs. world trial. First panel: participant fixates the world-fixed target (red cross) at the start of the first interval. Second panel: translation with world-fixed fixation target. Third panel: body-fixed fixation at start of second fixation interval. Fourth panel: translation with body-fixed fixation in second interval.

lation interval. The trial shown in the figure depicts the 10 cm reference translation with world fixation. After this first interval, a delay followed in which the participant was kept in complete darkness for 1.75 s. Then, the central fixation point reappeared, followed 0.5 s later by the second interval, in which the probe translation was presented. The set of possible probe translations ranged from 1 to 27 cm in equidistant steps of 0.4 mm. The fixation type in the probe interval was always different than in the associated reference interval (the trial in Figure 3.1 illustrates body fixation). After the second interval, the participant had to indicate whether he or she perceived the second translation as longer or shorter than the first using a 1-dimensional joystick. Moving the poke away from the body indicated that the second movement was longer, while moving it towards the body indicated that the second movement was shorter.

Thus, a trial consists of two translations with different fixation types; in the three main conditions we compare the body versus world, world versus free, and body versus free fixation types. For each main condition, we varied which fixation type served as the reference stimulus and the order in which reference and probe were presented, which gives a total of four variations per main condition (see Table 3.1). In addition we varied translation direction (either leftward or rightward on consecutive trials). The amplitude of the probe translation was adaptively chosen using the Psi method. This method picks the amplitude for the next trial which maximises the expected decrease in entropy based on participants' responses to earlier trials (Kontsevich & Tyler, 1999).

Comparison	Reference	1st interval	Direction
Body vs. world	Body	Reference	Right Left
		Probe	Right Left
	Body	Reference	Right Left
		Probe	Right Left
Body vs. free	Body	Reference	Right Left
		Probe	Right Left
	Free	Reference	Right Left
		Probe	Right Left
World vs. free	World	Reference	Right Left
		Probe	Right Left
	Free	Reference	Right Left
		Probe	Right Left

**Table 3.1:** List of the three main comparisons that we tested. The (10 cm) reference movement was presented in either the first or second movement interval. We also manipulated movement direction (leftwards or rightwards), yielding a total of 24 trial types.

This was done separately for all 24 trial types (3 main conditions x 2 reference stimuli x 2 reference/probe orders x 2 translation directions; see [Table 3.1](#)). A total of 25 trials were collected per trial type yielding a total of 200 trials for each of the three main conditions.

Trials were presented in three one-hour sessions. To prevent dark adaptation, we turned on the lights for 5 s after every block of 6 trials, and for at least 30 s every 4 blocks. We made sure that each of the 24 unique trial types were presented once every 4 blocks. After each block, the adaptive procedure determined which translation amplitudes to test in the following block. To increase the number of data-points available to the adaptive psychometric procedure at the beginning of the experiment, we collapsed across translation direction and reference order for the first 10 trials of every condition. After those collapsed trials, the procedure ran separately for each of the 24 distinct trial types.

### 3.2.4 Data analysis

For each combination of the three main conditions, and the two reference/probe orders (see [Table 3.1](#)), we quantified the perceived probe translation by calculating the

probability of the probe translation judged longer compared to the 10 cm reference translation as a function of actual probe translation, given by  $x$ . We used a maximum likelihood fit of a cumulative Gaussian function to summarise the psychometric data:

$$P(x) = \lambda + (1 - 2\lambda) \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(y-\mu)^2/2\sigma^2} dy, \quad (3.1)$$

in which  $|x|$  represents the size of the absolute probe displacement. The mean of the Gaussian represents the point of subjective equality (PSE). The slope of the curve reflects the precision ( $1/\sigma$ ) of reference-probe discrimination performance. Parameter  $\lambda$ , representing the lapse rate, accounts for stimulus-independent errors caused by subject lapses or mistakes and was restricted to small values ( $\lambda < 0.06$ ). Fits were performed using the Psignifit toolbox (Wichmann & Hill, 2001a, 2001b).

For each trial type (see [Table 3.1](#)), we also quantified eye movements, corrected for drift, based on initial fixation. The main source of drift were tiny lateral movements of the eye tracking cameras due to sled motion. We discarded trials containing blinks as well as trials in which the final eye position exceeded two standard deviations from the condition's average. Based on these criteria, 6.1%, 3.6% and 1.6% of all trials were rejected based on errors in body, world, and free fixation respectively. In addition we rejected 1.2% of all trials because participants blinked within the movement interval.

For the remaining trials, we computed the average ratio between the measured eye excursion,  $\varphi_i$ , and the angle that would be needed were the trial testing the world-fixed condition. The latter is computed by taking the arc-tangent of the actual translation distance,  $m_i$ , divided by the fixation depth,  $d_i$ , which for small  $\varphi$  can be approximated by  $g = \varphi m/d$ . We computed this ratio,  $g$ , for every fixation type and interval (see [Table 3.1](#)). Ideally, for body-fixed trials  $g = 0$ , and for world-fixed trials  $g = 1$ . Using this ratio, we are able to compute the expected eye excursion,  $\hat{\varphi} = gd/m$ , for any given translation distance even those we did not explicitly measure.

## Model

Using a simple cue integration model, we investigated whether inter-subject and inter-condition differences in the observed PSEs in conditions containing a translation under free fixation depend on actual eye movement behaviour. We modelled perceived distance,  $p$ , as a weighted linear combination of a vestibular and an oculomotor estimate of translation ([Equation 3.2](#)). We assumed that the vestibular estimate is equal to the actual translation,  $m$ , and that the oculomotor estimate is equal to expected eye movement given the actual,  $\hat{\phi}_m$ . As the weights represent the relative contributions of the oculomotor and vestibular systems, they can sum to any arbitrary value; in [Equation 3.2](#) their sum is fixed to 1. Thus, the weighting parameter  $\alpha$  regulates the eye movement contribution and  $1 - \alpha$  the vestibular contribution:

$$p = \alpha\hat{\phi}_m d + (1 - \alpha)m = \alpha g_m + (1 - \alpha)m \quad (3.2)$$

By definition, the probe displacement is perceived as equal in length to the 10 cm reference displacement at the PSE. By substituting both sides by the right hand side of [Equation 3.3](#) and using subscripts for reference ( $r$ ) and probe intervals ( $p$ ), we obtain:

$$\alpha g_r m_r + (1 - \alpha)m_r = \alpha g_p m_p \alpha + (1 - \alpha)m_p + \epsilon \quad (3.3)$$

In the present experiment, the reference displacement,  $m_r$ , was always 10 cm and the probe displacement,  $m_p$ , was equal to the measured PSE for the presented combination of fixation types (i.e.,  $PSE$  in [Equation 3.1](#)). This model (i.e. [Equation 3.3](#)) was then fit to data from the body and world conditions using linear regression, finding weight  $\hat{\alpha}$  that minimises the sum of squared errors ( $\sum \epsilon^2$ ).

$$m_r - m_p = \alpha(g_{f_p} m_p - g_{f_r} m_r + m_r - m_p) + \epsilon \quad (3.4)$$

By only using data from conditions where a visual fixation point was present during both translations (i.e. body versus world) to fit the model, we could examine whether the same weight  $\alpha$  can also explain the PSEs found in the conditions containing a

free fixation interval. To this end, we solved Equation 3 for  $m_p$  and computed PSE estimates,  $P\hat{S}E$ , for the body versus free and world versus free conditions ([Equation 3.5](#)).

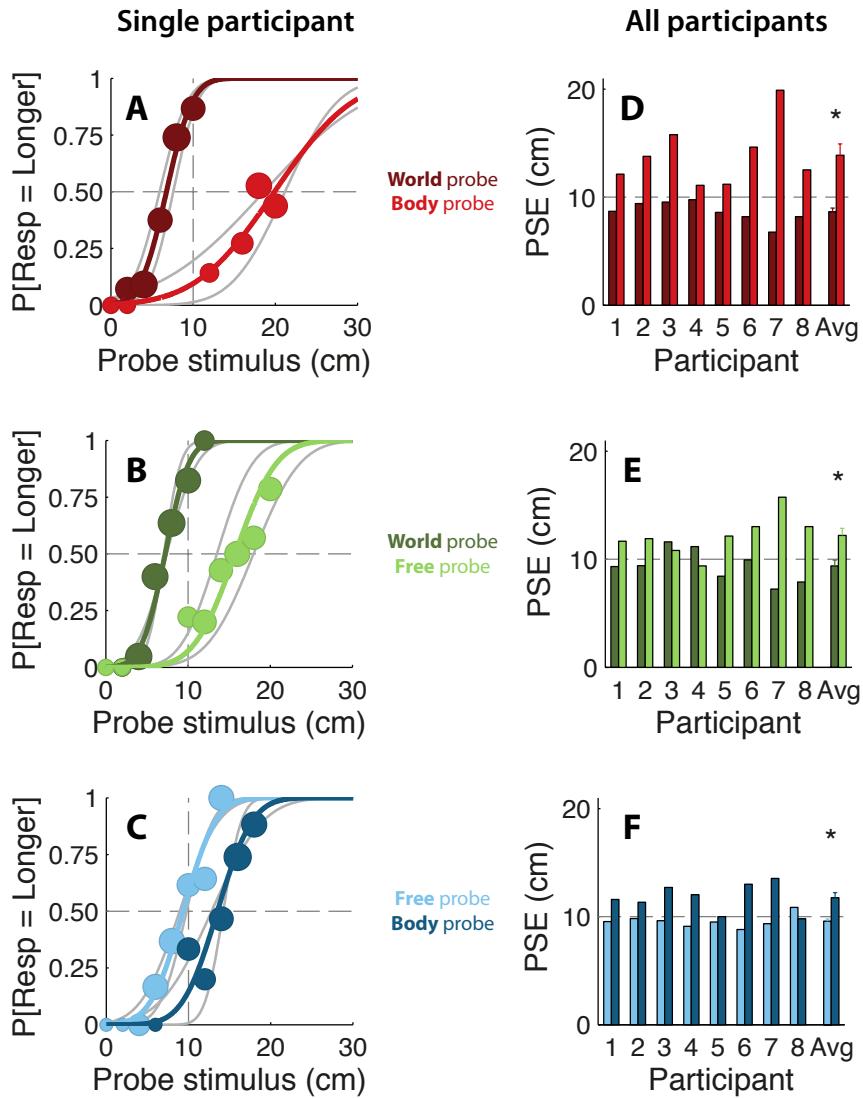
$$\hat{m}_p = P\hat{S}E = \frac{\alpha g_r + (1 - \alpha)}{\alpha g_p + (1 - \alpha)} m_r \quad (3.5)$$

In addition to minimizing the sum of squared errors in [Equation 3.4](#), we also fit [Equation 3.5](#) to the data in order to see if weight  $\alpha$  depends on the way the model is formulated. Parameters obtained by fitting [Equation 3.5](#) fell well within the standard deviation reported in [Table 3.2](#) for all participants, suggesting that they did not depend on the way the model was formulated.

### 3.3 Results

The current experiments investigate the influence of fixation type and associated eye movements on the perception of self-motion. Participants were presented with two subsequent lateral translations ([Figure 3.1](#)) and they had to judge whether the second was longer or shorter than the first. During each interval participants fixated a body- or world-fixed target (body and world fixation) or were moved in absence of a fixation point (free fixation).

The performance of one participant is illustrated in the left column of [Figure 3.2](#). Each row shows one main condition: body versus world fixation (top/red), world versus free fixation (middle/green), and body versus free fixation (bottom/blue). The lighter and darker colours in each panel indicate which fixation type was the reference movement (see Legend). The shift of the psychometric functions relative to the 10 cm reference (i.e. the PSE) quantifies the influence of fixation type. For example, the rightward shift of the light red curve in [Figure 3.2A](#) means that for a body fixation a longer translation ( $\approx 19$  cm) was required for that translation to be perceived equivalent to a 10cm reference translation with world fixation. On the other hand, the leftward shift of the dark red curve means that a shorter translation with world fixation ( $\approx 7$  cm) was required for that translation to be perceived equivalent to the 10 cm reference translation with body fixation. Together, these oppositely directed shifts demonstrate



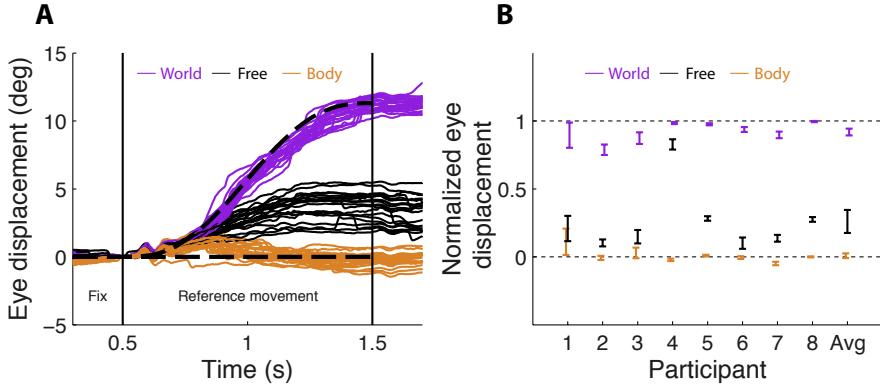
**Figure 3.2:** Psychometric curves (coloured lines) and associated binned data (circles) for one participant (top row). Circle size represents the number of trials within each 2 cm bin. Binning was only done in order to visualise this participant's responses and was not used otherwise. Gray lines show psychometric curves before collapsing across reference order. **A** Body-world comparison; body reference, dark red; world reference, light red. **B** World-free comparison; world reference, light green; free reference, dark green. **C** Body-free comparison; body reference, dark blue; free reference, light blue. PSEs for all participants and the average  $\pm$ SE (bottom row). **D** Body-world (dark red) and world-body (light red) conditions. **E** World-free (light green) and free-world (dark green) conditions. **F** Body-free (dark blue) and free-body (light blue) conditions. Because a t-test revealed a main effect of reference order,  $t(47) = -5.2$ ,  $p < 0.01$ , we used the mean PSE across reference order (e.g. [Figure 3.3](#), gray lines) instead of the PSE collapsing across reference order (e.g. [Figure 3.3](#), coloured lines); these values were not significantly different.

that translations with world fixation were perceived longer than equivalent translations with body fixation, regardless of which translation was the reference. Similarly, the shifts in [Figure 3.2B](#) shows that world fixation translations were also perceived to be longer than free fixation movements and [Figure 3.2C](#) shows that free fixation translations were perceived to be longer than body fixation translations. Note that [Figure 3.2](#) also shows effects on slope, which will be further discussed in the paragraph [Precision depends on PSE](#).

Similar results were obtained for all subjects, as shown by the individual PSEs for all participants (right column of [Figure 3.2](#)). Statistical significance of the fixation-induced effects for each main condition (world versus body, [Figure 3.2D](#); world versus free, [Figure 3.2E](#); and free versus body, [Figure 3.2F](#)) was evaluated by comparing PSEs between the two reference conditions using a paired t-test. These PSEs were significantly different in all cases (world versus body,  $t(7) = -4.09$ ,  $p < 0.05$ ; world versus free,  $t(7) = -2.48$ ,  $p < 0.05$ ; free versus body,  $t(7) = -3.38$ ,  $p < 0.05$ ). As for the example subject, these results indicate that translations made with body fixation are perceived shorter than with world fixation, suggesting that self-motion perception is modulated by eye movements even in absence of full-field optic flow. The free fixation translations, which control for confounds of the small fixation point, were perceived to be longer than body and shorter than world fixation translation intervals, which could be expected if their gains were smaller than 1 but larger than 0.

### 3.3.1 Eye movement contributions to self-motion perception

In order to relate psychophysical performance to eye movement behaviour we recorded and analysed eye movements during both intervals of every trial for all subjects. Exemplar eye traces for the 10 cm reference translation for the three fixation types are depicted in [Figure 3.3A](#). Fixation behaviour was quite accurate for both body fixations, where no eye movements were expected, and world fixation, where eye movement excursions of 11° were expected, seemingly supported by catch-up saccades. Under free fixation, the amount of eye movement was intermediate between body and world fixation and behaviour was more variable. A similar pattern was observed in all



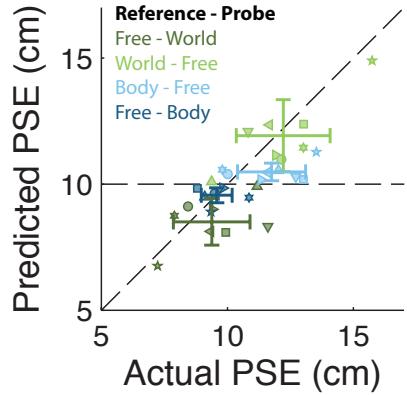
**Figure 3.3:** **A** Actual (solid lines) and ideal (dashed lines) eye movement traces of one participant during world fixation (purple), body fixation (brown), and free fixation (black). All traces shown are for 10 cm reference movements. **B** Normalised eye position for each participant ( $\pm 95\%$  confidence interval) at the end of translation interval (error bars) for world fixation (purple), body fixation (brown) and free fixation (blue). In addition, the average  $\pm \text{SE}$  across all participants is shown. Zero indicates that the eyes remained stationary relative to the body, and one indicates that eye position was perfectly world-fixed.

Participant	Parameter ( $\alpha$ ) $\pm \text{SD}$
1	0.27 ( $\pm 0.04$ )
2	0.27 ( $\pm 0.05$ )
3	0.35 ( $\pm 0.04$ )
4	0.06 ( $\pm 0.04$ )
5	0.13 ( $\pm 0.03$ )
6	0.33 ( $\pm 0.04$ )
7	0.58 ( $\pm 0.02$ )
8	0.21 ( $\pm 0.02$ )

**Table 3.2:** Estimated eye movement contribution ( $\alpha$ ) to the perception of self-motion (see [Equation 3.5](#)). Standard deviations are based on a bootstrap for each participant.

participants, as illustrated by the normalised eye movement data (see [Materials and methods](#), and [Figure 3.3B](#)).

To quantify the role of eye movements in self-motion perception, we tested a linear model (see [Model](#)) in which perceived translation is a weighted average of a vestibular estimate (equal to the actual translation) and an oculomotor estimate (equal to the normalised eye movement times the actual translation; [Equation 3.2](#)). This model contains a single free parameter ( $\alpha$ ), which corresponds to the relative weight given to the oculomotor estimate. We fit this model to the two body versus world conditions and obtained the value of the oculomotor weight for every subject ([Table 3.2](#)). The average oculomotor weight is  $0.25 \pm 0.12$  (SD), indicating that the contribution of the eye movement signal to the self-motion estimate is about 25 percent. Note that participant 4, whose oculomotor weight is furthest from this mean ( $\alpha = 0.06$ ), also shows a radically different eye movement gain during the free-fixation (see [Figure](#)

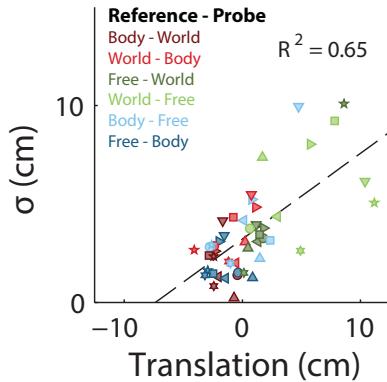


**Figure 3.4:** Eye movement based prediction for the PSE plotted against the actual PSE. A data point (symbol) is shown for each participant (symbol shape) and condition (symbol colour) pair, following the same colour scheme as in [Figure 3.2](#). The identity line, corresponding to a perfect prediction, is shown in black.

[3.3B](#)). We then used these oculomotor weights along with the normalised eye movement values to predict the PSEs in the remaining four conditions according to [Equation 3.5](#). The predicted PSEs are plotted against the actually observed PSEs in [Figure 3.4](#). The positive correlation ( $\rho = 0.78, p < 0.01$ ) between observed and predicted PSEs suggests that eye movements are indeed used in self-motion perception, even in the absence of a fixation point (i.e. during free fixation). Furthermore, the fact that data points generally cluster near the unity line shows that our simple model does reasonably well in predicting perceptual performance across subjects and conditions based on oculomotor weight and normalised eye movement magnitude only. This holds true even for subject 7 whose oculomotor weight ([Table 3.2](#)) was approximately double the average, yet whose data points remain close to the unity line.

### 3.3.2 Precision depends on PSE

The psychometric curves of the example participant in [Figure 3.2](#) show that precision ( $\sigma^{-2}$  in [Equation 3.1](#)) decreases as the difference between translated distance in the reference and probe intervals (i.e. the bias) increases. In [Figure 3.5](#) precision is plotted as a function of bias for all participants and all conditions, showing a significant linear relationship ( $R^2 = 0.64$ ) between the two. This effect, which follows Weber's perceptual law (Fechner, 1860) is consistent with the signal-dependence of (discrimination)



**Figure 3.5:** Effect of difference movement amplitude between reference and probe interval (i.e. the bias) on response uncertainty ( $\sigma$ ). A data point is shown for every participant and condition (symbol colour) pair, following the same colour scheme as in [Figure 3.2](#). The dashed black line is the linear regression trend line.

precision that has been shown recently for vertical self-motion (Nesti, Barnett-Cowan, MacNeilage & Bülthoff, [2014](#)).

### 3.4 Discussion

We investigated the contribution of eye movements to the perception of passively-induced self-motion. Experiments were performed in the absence of full-field optic flow to eliminate the contribution of this visual motion signal. Perception of self-motion was compared across three fixation types: during free fixation the fixation target was extinguished before the movement, while during world and body fixation, targets remained stable relative to the world and body, respectively. Our results show that self-motion is underestimated during body fixation (in which the eyes remain stationary) compared to world fixation (in which the eyes move to maintain fixation). The eye movements during free fixation, which are driven by the VOR, show a non-unity gain with excursions in-between body- and world-fixation conditions. Self-motion perception reflects this pattern of eye movements, suggesting an important contribution of this extraretinal signal to the perception of self-motion.

To quantitatively characterise the separate vestibular and eye movement contributions, we fit a single parameter model to the perceptual responses for the body versus world comparison conditions and validated this model independently by predicting the effects

of eye movements on self-motion perception during free fixation conditions. This model takes into account subject specific oculomotor weight and eye movement patterns. Based on these inputs it accurately predicts the responses in the free fixation conditions. This demonstrates that extra-retinal eye movement signals are used as a cue in the perception of self-motion, contributing significantly to the self-motion percept with a weight of approximately 25 percent, even in the absence of optic flow.

It is surprising that an influence of eye movements can be observed even for body-stationary fixations, during which the stationary eye movement signal is clearly in conflict with the non-zero vestibular signal. While this demonstrates the strength of the assumption that fixation targets are world-stationary, it raises the question how reliable this assumption is. Simultaneous recording of angular head and eye movements during natural behaviour reveals that approximately 80 percent of eye movements can be classified as compensatory, i.e. eye movements directed opposite to head movement and therefore consistent with maintenance of world-fixed fixation (Einhäuser et al., 2007). Similarly, other studies have shown that world-stationary fixations are common for many every day activities, ranging from making a cup of tea (Hayhoe & Ballard, 2014) to driving a car (Land & Lee, 1994), to walking (Foulsham, Walker & Kingstone, 2011) and even reaching, where people tend to look at the source and destination of the object, but not at the hand (Flanagan & Johansson, 2003). Because world-stationary fixations are so common, the natural world statistics imply that self-motion and eye movements are highly correlated, thus making eye movements a fairly reliable cue for self-motion.

Even when fixation is not world-fixed, eye movement signals are combined with optic flow signals to yield realistic self-motion estimates (e.g Royden et al., 1992; van den Berg & Beintema, 2000). During world-fixed fixation, the eyes move to compensate for body translation, thereby reducing the optic flow component in the retinal signal. The self-motion estimate will therefore be driven predominantly by the eye movement signal. On the other hand, in the body-fixed condition, eye movements are minimal and optic flow maximal such that perceived self-motion will be driven predominantly by the optic flow signal itself. Because our experiment was performed in darkness, this optic flow signal was absent in the body-fixed condition which can explain why self-motion was underestimated.

During body and world fixation, eye movements are driven by retinal slip of the fixation target. However, in the free fixation condition, retinal slip is not available and resulting eye movements resemble the linear vestibulo-ocular reflex (LVOR), in that the gain relative to world fixation was  $\approx 0.4$  (see [Figure 3.3B](#); Ramat & Zee, [2003](#)). This reflex is thought to be driven by a double integration of the vestibular signal, converting the head acceleration signal from the otoliths to eye position (Green, Meng & Angelaki, [2007](#); Walker, Tian, Tamargo, Ying & Zee, [2010](#)). If eye movements during free fixation are in fact vestibularly driven, then combination of this eye movement signal with the vestibular signal itself seems redundant. However such combination could reflect a strategy to reduce noise. Both the direct (vestibular) and indirect (LVOR) signals depend on integration of the linear acceleration signal and may be corrupted by independent noise sources. Combining them in a statistically optimal fashion will decrease the noise level towards the noise level of the original source signal (Faisal, Selen & Wolpert, [2008](#); Clemens et al., [2011](#); Fetsch, DeAngelis & Angelaki, [2013](#)). The consequence of this integration will be a reduced self-motion estimate when the gain of the LVOR is less than 1, as we observed in the free condition.

### 3.4.1 Alternative interpretations

In the above, we suggest that eye movements themselves drive perception of self-motion. However, it is conceivable that a common correlate of eye movements, such as attention or visual motion influenced our results. In 1963, Guedry and Harris reported a substantial underestimation of displacement when their observers watched a small body-fixed target compared to displacements in the dark. They attributed their findings to an attentional shift from judgements of body displacement in the dark to judgements of target displacement in the fixation condition. We favor an explanation based on eye movement characteristics. In their study, it is likely that the VOR caused eye movements to occur during the translations in darkness. If these movements were used to augment self-motion perception, then the perception of such translations would be overestimated compared to translations made without eye movements, e.g. when fixating a body-fixed target. Because Guedry and Harris ([1963](#)) did neither record nor explicitly manipulate eye movements, they were not able to unveil their explicit role. Conversely, we did not manipulate attentional processes (Kitazaki & Sato, [2003](#)), so

we cannot completely exclude the possibility they play a role.

Others have reported errors in the disambiguation of self and object-motion. Examples include the perceived motion of body-fixed visual targets during angular acceleration (the oculogyral illusion; Carriot, Bryan, DiZio & Lackner, 2011), the apparent displacement of body-fixed stimuli during linear acceleration (the oculogravic illusion; Graybiel, 1952) and the apparent movement of world-stationary targets during self-motion in darkness (Dyde & Harris, 2008). Similar disambiguation errors could cause the effects we observed. More specifically, if movement of the fixation point relative to the observer were always attributed to self-motion, then self-motion would be underestimated during body relative to world fixation, as we observed. However, such attribution errors cannot account for the effects in the free condition, because no fixation point was visible and no attribution was required. In the free condition, we demonstrate that eye movements by themselves, occurring in the absence of visual tracking and other external cues, influence the perception of self-motion.

### 3.4.2 Implications for other studies

Many previous self-motion studies have used a body-fixed fixation point to control for eye movement related effects. Our results suggest, however, that using a body-fixed fixation point causes underestimation of self-motion. For example, Li, Wei, and Angelaki (2005) investigated spatial updating across lateral translation and found that saccades to updated targets undershot the actual target location. As self-motion perception drives this update, the effects of eye movements on self-motion perception should also influence the updating process. In other words, the observed undershoot could be due to the underestimation of self-motion caused by the body-fixed fixation point. Another example is a study on the perception of vertical object-motion during lateral translation (Dokka, MacNeilage, DeAngelis & Angelaki, 2013). This study reports incomplete compensation for self-motion when judging the deviation from vertical motion of a moving object. This observation could also be due to underestimation of self-motion induced by the fixation of the body-fixed target.

A moving fixation point is also known to influence self-motion perception, as in the Slalom Illusion (Freeman, Banks & Crowell, 2000); observers viewing expanding optic

flow while fixating a target that oscillates from left to right perceive slaloming motion which is inconsistent with the purely forward motion specified by the expanding optic flow display. However, this observation is consistent with the idea that oculomotor signals are used in estimating self-motion. Additionally, it has been shown that eye movements affect postural sway (Glasauer et al., 2005). Participants performed smooth pursuit eye movements in complete darkness and displayed lateral sway consistent with the stabilization of posture using a self-motion estimate influenced by pursuit eye movements.

Studies conducted to characterise vestibular-only sensitivity are often performed in complete darkness or with closed eyes (Grabbherr, Nicoucar, W. & Merfeld, 2008; MacNeilage, Banks, DeAngelis & Angelaki, 2010; MacNeilage, Turner & Angelaki, 2010; Roditi & Crane, 2012; Valko, Lewis, Prieson & Merfeld, 2012; Nesti et al., 2014). However, the results of our free-fixation condition suggest that even under these circumstances, results could easily be influenced by vestibularly driven eye movements. Overall, we suggest that any study concerned with self-motion processing must consider the possible influence of eye movements.

### 3.4.3 Possible neural substrate

This leaves us with the question of where in the brain these effects originate. The locus of our effect is likely to carry both eye movement and vestibular signals. Prime candidate areas known to carry both vestibular and eye movement signals are the vestibular nuclei (Henn, Young & Finley, 1974; Daunton & Thomsen, 1979) and the cerebellum (Waespe, Büttner & Henn, 1981). On the other hand, eye movements could influence self-motion perception indirectly via optic flow processing. In particular, cortical areas that carry both vestibular and optic flow signals (which can be modulated by eye movements) include the ventral intraparietal area (VIP; Bremmer, Klam, Duhamel, Ben Hamad & Graf, 2002; Chen, DeAngelis & Angelaki, 2011), and the dorsal medial superior temporal area (MSTd; Gu, Angelaki & DeAngelis, 2008). Future work should reveal how such brain areas, directly or indirectly, merge both vestibular and oculomotor signals into a coherent percept of self-motion.



---

---

*CHAPTER 4*

---

*Translation perception is  
modulated by eye  
movements that are  
partially scaled by fixation  
depth*

**Abstract** It has been shown that the compensatory eye movements, that minimise retinal slip during self-motion, also serve as a cue for translation perception. However, to provide an adequate translation estimate, the brain must internally scale these ensuing eye movements by fixation distance. Using a 2AFC approach, we investigated whether the brain applies this scaling. Participants ( $n = 8$ ) were translated sideways in the absence of full-field optic flow but with gaze maintained on either a nearby or faraway target, that was either fixed in the world (world-fixed) or moved along with the body (body-fixed). Results show that translations were perceived shorter with gaze on nearby than faraway world-fixed targets, indicating that eye movements are not properly scaled in translation perception. Translation perception was not affected by the depth of body-fixed targets. Taken together, our results suggest that eye movements are merely a rudimentary cue to self-motion, with a compensation for fixation depth that is partial at best.

---

This chapter is being prepared for publication

**Clemens, I.A.H.**, Selen, L.P.J., MacNeilage, P.R. and Medendorp, W.P. (2015b).

## 4.1 Introduction

An accurate internal estimate of self-motion is required to navigate effectively through a complex three-dimensional environment. The vestibular system as well as optic flow provide essential information about self-motion (Gibson et al., 1955; Benson et al., 1986; Harris et al., 2000; Israël & Berthoz, 1989; Angelaki & Hess, 2005; Carriot et al., 2013; Chen et al., 2010). During navigation, however, the eyes typically move to maintain visual acuity on important objects. These eye movements disturb the optic flow patterns. Using the oculomotor signal, the brain is able to account for these disturbances by internally separating optic flow into two components, one caused by self-motion and the other by eye movement (Warren & Hannon, 1988; Royden et al., 1992; Freeman & Banks, 1998; Lappe et al., 1999).

When the eyes track world-centred objects, their angular displacement is directly related to the size of the motion of the observer (Schwarz et al., 1989; Paige et al., 1998; McHenry & Angelaki, 2000; Medendorp et al., 2002). Because the majority of fixations are on world-stationary objects, we recently proposed that these tracking eye movements could also be used as a self-motion cue, in addition to optic flow and vestibular signals. To test this hypothesis, we compared self-motion perception in the absence of full-field optic flow during passively induced whole-body translations (Clemens et al., 2015a). Our results showed that self-motion is underestimated during body-centred fixations (in which the eyes remain stationary in their orbits) compared to fixations on world-stationary objects (in which the eyes must move to maintain fixation).

Geometrically, eye movements that keep fixation on a world-centred target during lateral whole body translation (i.e. the linear vestibulo-ocular reflex; LVOR), must scale with fixation depth (Angelaki, 2004). When fixating body-centred targets these eye movements must be suppressed irrespective of fixation distance (Angelaki, 2004). Conversely, when fixating world-centred targets, the brain must internally scale the ensuing eye movement by fixation distance to serve as an adequate self-motion cue. Because we did not manipulate fixation distance in our previous study, we could not dissociate whether eye movements are used as a rudimentary cue for self-motion (i.e. without taking fixation depth into account), or are properly scaled in the mechanisms

for self-motion perception.

In the present study, we investigate how fixation distance influences perception of self-motion during passive side-to-side translations. Using a psychophysical approach, participants had to indicate whether the second body displacement of two one-second translation intervals was smaller or longer than the first. We show that translation amplitude is perceived smaller when fixating a far compared to a nearby world-centred target, indicating that eye movements are not properly scaled in self-motion perception. Together with the observation that self-motion perception is not affected by the depth of a body-centred fixation target, we conclude that eye movements are merely a rudimentary cue to self-motion, with a compensation for fixation depth that is partial at best.

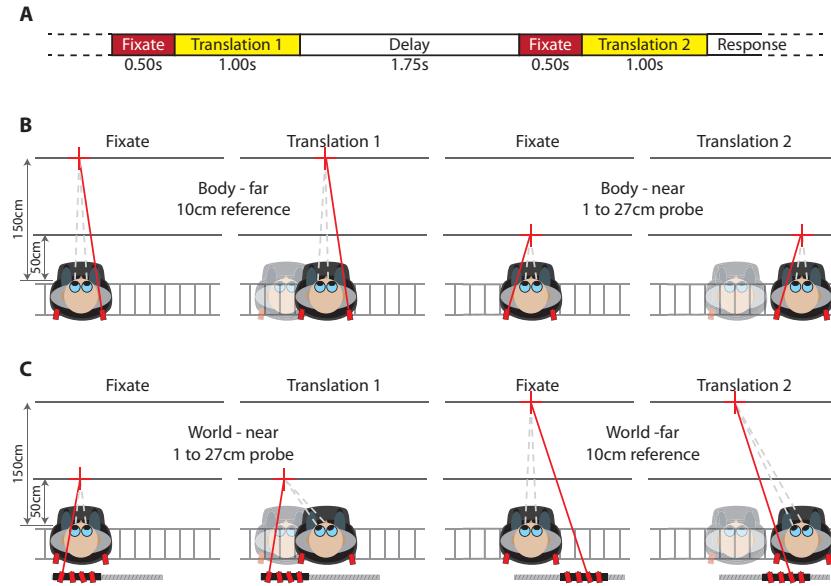
## 4.2 Materials and methods

### 4.2.1 Participants

Eight naive participants (three male, five female), aged between 22 and 29 years, gave written informed consent to participate in the study. They were all free of any known vestibular or neurological disorder and had normal or corrected-to-normal visual acuity. Participants never received any feedback about their performance. The experimental setup and methods used here are similar to those used in our previous paper on the influence of eye movement type on self-motion perception (Clemens et al., 2015a). We only provide a brief summary here, and refer to our previous paper for further details.

### 4.2.2 Experimental setup

Participants were seated on a motorised linear sled with their body and head restrained such that the inter-aural axis aligned with the motion axis. The sled laterally translated participants following a minimum jerk profile of fixed duration (1 s) and amplitudes ranging from 1 to 27 cm. Auditory cues were suppressed using white noise



**Figure 4.1:** **A** Time course of key events within a single trial. In each of the two intervals, a 0.50 s fixation period (red) precedes the lateral translation (yellow). A 1.75 s long delay period (shown in white) separates the two intervals. After the second translation, the participant responded whether this second translation was longer or shorter than the first. **B** and **C** Top-view illustrating key events in a far versus near body-fixed fixation trial (**B**) and a near versus far world-fixed fixation trial (**C**). The first panel shows the initial fixation to the target (red cross) followed by a translation in the second panel. The third panel shows the initial fixation before the second movement interval followed by a translation in the fourth panel.

presented through in-ear head-phones. Experiments were conducted in complete darkness except for visual fixation points, projected by body- or world-fixed laser pointers on a black bar, either 50 or 200 cm in front of the participant and at eye level. Eye movements were recorded at 500Hz using an EyeLink II system (SR Research, Kanata, Canada). Eye position was calibrated before every session using 11 evenly spaced calibration points ranging from -22 to 22 degrees.

#### 4.2.3 Paradigm

We used a two-alternative forced choice (2AFC) task to study the influence of fixation depth on the perception of linear translation. We tested two fixation depths: near (50 cm) and far (200 cm) and two different eye fixation conditions: world-, and body-fixed fixation. A trial contained two sequential motion intervals of equal duration (1 s), with the motion in the same direction (either leftward or rightward).

The timing of a single trial is shown in [Figure 4.1A](#). Every trial started with the onset of a central fixation point (i.e. aligned between the eyes), at a depth of 50 or 200cm, for

Comparison	Reference	1st interval	Direction
Body Near vs. far	Near	Reference	Right
		Probe	Left
	Far	Reference	Right
		Probe	Left
World Near vs. far	Near	Reference	Right
		Probe	Left
	Far	Reference	Right
		Probe	Left

**Table 4.1:** List of the 2 main comparisons that we tested. The (10cm) reference movement was presented in either the first or the second interval. We also manipulated movement direction (leftward vs rightward) yielding a total of 16 trial types.

0.5 s. Subsequently, the first 1s motion interval commenced. During this translation, the fixation point either remained world stationary (world condition), or moved along with the participant (body condition). Subsequently a 1.75 s delay followed in which the sled was stationary and no fixation light was shown. Next, a central fixation point reappeared at the other depth (50 or 200cm) than was used in the first interval. After 0.5 s, the second 1 s translation interval started, with the same fixation condition as in the first interval. After this second interval, the participant responded whether the second displacement was perceived longer or shorter than the first by moving a 1-dimensional joystick away from (longer) or towards (shorter) the body. Top-view illustrations of example body-centred and world-centred trials are shown in [Figure 4.1B](#) and [C](#) respectively.

Across trials, we varied the order of the fixation depths and the order of the reference and probe interval, resulting in four variations of both the world and body condition (see [Table 4.1](#)).

Leftward and rightward motion alternated between trials, but were not considered as variations of the condition. To determine the point of subjective quality (PSE), the size of the probe translation was adaptively chosen based on the participants' earlier responses (Psi method; Kontsevich & Tyler, [1999](#)). This was done separately for all 16 trial types (2 main conditions x 2 depth orders x 2 reference/probe orders x 2 movement directions; see Table 1). A total of 25 trials was collected per trial type yielding a total of 200 trials for each of the two main conditions.

These trials were presented in two one-hour sessions. To prevent dark adaptation, we turned on the lights for 5 s after every 8 trials, and for at least 30 s after every 16 trials. Each of the 16 unique trial types were presented once in every block of 16 trials. After each block, the adaptive procedure determined which translation size to test in the following block. To increase the number of data-points available to the adaptive psychometric procedure at the beginning of the experiment, we collapsed across movement direction and reference order for the first 10 trials. After that, the procedure ran separately for the 16 distinct trial types.

#### 4.2.4 Data analysis

For each combination of the two main conditions and the two reference/probe orders we computed the probability  $P(x)$  of probe translation  $x$  being judged as longer than the reference translation. To summarise these data, we fit cumulative Gaussian functions to these probabilities, resulting in a total of four Gaussian functions per participant (see [Table 4.1](#)):

$$P(x) = \lambda + (1 - 2\lambda) \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(y-\mu)^2/2\sigma^2} dy, \quad (4.1)$$

The mean of the Gaussian,  $\mu$ , represents the point of subjective equality (PSE). The slope of the curve reflects the precision ( $1/\sigma$ ) of reference-probe discrimination performance. Parameter  $\lambda$ , representing the lapse rate, accounts for stimulus-independent errors caused by subject lapses or mistakes and was restricted to small values ( $\lambda < 0.06$ ). Fits were performed using the Psignifit toolbox (Wichmann & Hill, [2001a](#), [2001b](#)).

For each trial type (see [Table 4.1](#)), we also quantified the eye movements, corrected for drift based on initial fixation. We discarded trials containing blinks as well as trials in which final eye position exceeded two standard deviations from the condition's average. Based on these criteria, 12% of all trials were discarded. Of the remaining trials, we computed the average ratio between the measured eye excursion,  $\varphi_i$ , and the angle that the eyes would have moved through had they perfectly tracked a world-stationary fixation target at the same fixation depth. The latter is computed by taking the arc-

tangent of the actual translation distance,  $m_i$ , divided by the fixation depth,  $d_i$ , which for small  $\varphi$  can be approximated by  $g_c = \frac{\varphi_i m_i}{d_i}$ . We computed this ratio,  $g_c$ , for every trial type  $c$  (see [Table 4.1](#)). Ideally, for body-fixed trials,  $g_c = 0$ ; and for world-fixed trials,  $g_c = 1$ .

#### 4.2.5 Model

Using a straightforward model, we investigate to what extent fixation depth is taken into account in the contribution of eye movements to self-motion perception. As in Clemens et al. ([2015a](#)), we model the perceived translation,  $p_i$ , as a weighted combination of a vestibular,  $m_i$ , and oculomotor estimate of translation,  $\hat{\varphi}_i$  ([Equation 4.2](#)).

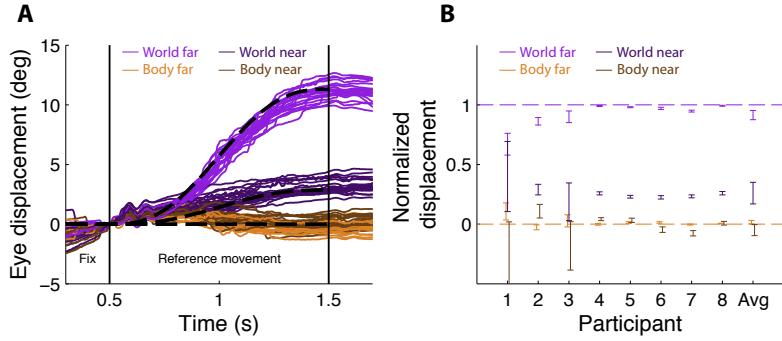
$$p_i = \alpha_{d_i} \hat{\varphi}_i + (1 - \alpha_{d_i}) m_i \quad (4.2)$$

Variable  $i$  represents either the reference,  $r$ , or probe,  $p$ , interval. To serve as a veridical cue for self-motion, eye movements need to be scaled by the depth of fixation,  $d$ . This scaling is reflected by parameter,  $\alpha_{d_i}$ . If this parameter is the same across the two fixation depths,  $d$ , then there is no depth-dependent modulation of the oculomotor estimate of translation.

By definition, at the PSE, the probe translation is perceived as equal in length to the 10 cm reference translation,  $p_r = p_p$ . By substituting both sides by the right hand side of [Equation 4.2](#), we obtain:

$$\alpha_{d_r} \hat{\varphi}_r + (1 - \alpha_{d_r}) m_r = \alpha_{d_p} \hat{\varphi}_p + (1 - \alpha_{d_p}) m_p + \epsilon \quad (4.3)$$

We fit [Equation 4.3](#) to the data using linear regression, finding one weight for each of the two fixation depths (that is,  $\alpha_{50}$  and  $\alpha_{200}$ ) that minimises the sum of squared errors ( $\sum \epsilon^2$ ). Because these parameters can, in theory, contain both a depth-dependent and a depth-independent scaling, we compute their ratio,  $\alpha_{200}/\alpha_{50}$ , to remove any depth independent components. In case of perfect compensation, the expected ratio is  $200/50 = 4$ , while in case of depth-independent scaling it is 1.



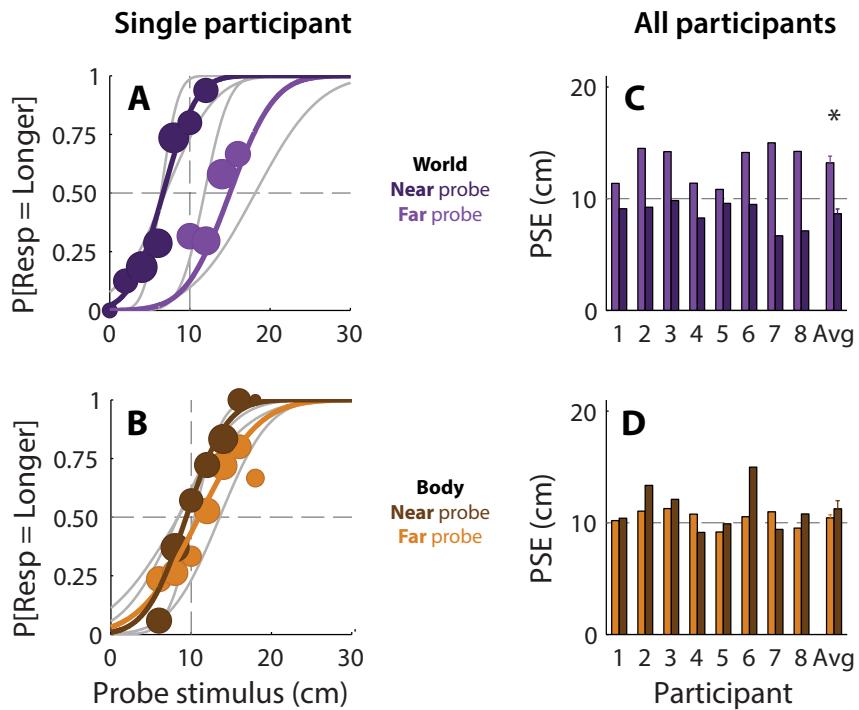
**Figure 4.2:** **A** Actual (solid lines) and ideal (dashed lines) eye movement traces of one participant in the body-fixed (brown and orange) and world-fixed conditions (purple and pink). Gaze was directed at a near (brown and purple) or far (orange and pink) target. All traces shown are for 10cm reference displacements. **B** Normalised eye position for each participant ( $\pm 95\%$  confidence interval) at the end of the translation interval for the near and far world fixed targets (purple and pink respectively) as well as the near and far body fixed fixation targets (brown and orange respectively). In addition, the average  $\pm \text{SE}$  across all participants is shown. Zero indicates the eyes remained stationary relative to the body, and one indicates the eyes followed the near world fixed target perfectly.

### 4.3 Results

In a recent study we have shown that eye movement signals contribute to the perception of body translation, even in the absence of optic flow or a visual fixation point (Clemens et al., 2015a). Because eye rotations must be scaled by target depth ( $\varphi d = T$ ) to serve as an adequate translation cue, we tested self-motion perception for near (50 cm) and far (200 cm) fixations. Participants were presented with two subsequent translations Figure 4.1 while they kept fixation on a world- or body-stationary target that was presented either nearby or far away. After the two translation intervals, participants had to judge whether the second translation was longer or shorter than the first.

We first investigated the ability of participants to fixate body and world stationary targets. Figure 4.2A depicts exemplar eye traces for the 10cm reference translation with nearby and far fixation points in both the body and world condition. Changes in gaze are largely absent in the body near and body far conditions (brown and orange traces respectively), as required. During the world conditions, the eye excursions were large when fixating nearby targets and small when fixating far away ones (purple and pink traces respectively), which reflects the geometrical constraints.

We normalised the eye movement data by taking the average ratio between the meas-



**Figure 4.3:** A and B Psychometric curves (coloured lines) and associated binned data (circles) for one participant. Circle size represents the amount of trials within the bin. Psychometric curves before collapsing across reference order are shown as gray lines. A World-fixed condition (purple) while fixation was either near (dark) or far (light). B Body-fixed condition (brown) while fixation was either near (dark) or far (light). C and D Individual and average points of subjective equality (PSEs). Colour scheme matches panels A and B.

ured eye excursion and the geometrically required eye displacement were the target world stationary. Figure 4.2B shows that these normalised eye displacements are about zero during body fixed fixations (brown and orange) and close to one in the world fixed fixations (purple and pink data points), for all participants. The question is whether these eye movements are inversely scaled by fixation depth in order to interpret them as linear self-motion cues.

Figure 4.3 illustrates psychophysical data on self-motion perception of a single participant for the two-fixation depths in both the world (Figure 4.3A) and body condition (Figure 4.3B). Lighter and darker colours indicate which fixation depth was the reference translation (see figure legend). The influence of fixation depth is characterised by a shift of the psychometric functions relative to the 10 cm reference translation (i.e. the PSE). For example, the rightward shift of the pink curve in Figure 4.3A,

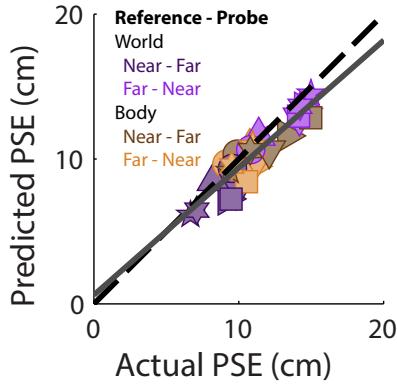
Participant	$\alpha_{50}$	$\alpha_{200}$	$\frac{d_{200}}{d_{50}}$	$\alpha$
1	0.37	0.25	1.46	0.27
2	0.51	0.41	1.25	0.27
3	0.36	0.30	1.22	0.35
4	0.14	0.29	0.49	0.06
5	0.11	0.04	2.79	0.13
6	0.49	0.15	3.34	0.33
7	0.40	0.53	0.76	0.58
8	0.42	0.35	1.25	0.21

**Table 4.2:** Best-fit parameter values for 50 and 200 cm fixation distances,  $\alpha_{50}$  and  $\alpha_{200}$  respectively (see [Equation 4.3](#)), and their ratio,  $\alpha_{200}/\alpha_{50} = d_{200}/d_{50}$ , for each participant. Best-fit parameter values,  $\alpha$ , from our previous paper ([Clemens et al., 2015a](#)) are included for reference.

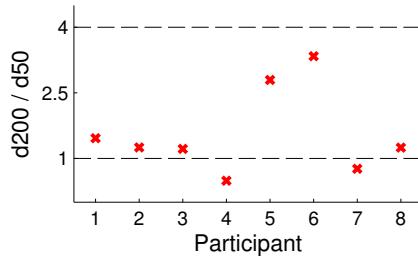
representing the world-condition, means that with a far target a longer translation ( $\approx 15$  cm) was required to be perceived equivalent to a 10 cm reference translation with nearby fixation. Likewise, the leftward shift of the purple curve indicates that a shorter translation with near fixation ( $\approx 6$  cm) is required to be perceived the same as the 10 cm reference translation with far fixation. Together, these opposite biases suggest that translations are perceived shorter for fixations further away. For the body condition, no shift of the psychometric curves is visible ([Figure 4.3B](#)), indicating that fixation depth (i.e. near versus far) has no effect in absence of eye movements.

Similar results were obtained for all participants, as shown by the individual PSEs (right column of [Figure 4.3](#)). Statistical significance of the effects of fixation depth was evaluated by comparing PSEs for the two fixation depths using a paired t-test. PSEs differed significantly between the two fixation depths in the world condition,  $t(7) = 5.42$ ,  $p < 0.01$  ([Figure 4.3C](#)), but not in the body condition,  $t(7) = -1.17$ ,  $p = 0.28$  ([Figure 4.3D](#)), confirming the single subject results ([Figure 4.3A and B](#)). Thus, increasing fixation depth does not influence self-motion perception during body-stationary fixations, but causes self-motion to be perceived as shorter during world-stationary fixations.

We used a simple linear model to quantify to what extent eye movements are scaled by fixation depth in order to serve as a linear self-motion cue (see [Materials and methods](#)). This model describes the perceived translation distance as a weighted average of a vestibular estimate, equal to the actual translation, and an oculomotor-based estimate. The latter estimate depends on the eye excursion which should be scaled by fixation depth to serve as a valid cue. Our model contained two weighting parameters,  $\alpha_{50}$  and  $\alpha_{200}$ , one per fixation depth (see [Table 4.2](#) for best-fit values). Using these parameters we predicted the PSEs, i.e.  $m_p$  in [Equation 4.3](#), and plotted them against the actually



**Figure 4.4:** Eye movement based prediction for the PSE plotted against the actual PSE. A data point (symbol) is shown for each participant (symbol shape) and condition (symbol colour) pair, following the same colour scheme as in [Figure 4.3](#). The identity line, corresponding to a perfect prediction is shown (solid line) as well as the best fit line (dashed).



**Figure 4.5:** Ratio between near and far parameter values,  $\alpha_{200}/\alpha_{50} = d_{200}/d_{50}$  for every participant. Both perfect depth compensation,  $\alpha_{200}/\alpha_{50} = 200/50 = 4$ , as well as the lack thereof,  $\alpha_{200}/\alpha_{50} = 1$ , are represented by a dashed line.

observed PSEs in [Figure 4.4](#). The positive correlation ( $\rho = 0.92$ ,  $p < 0.01$ ) between observed and predicted PSEs shows that our simple model does reasonably well in predicting perceptual performance.

By examining the ratio of these weighting parameters, we remove any depth-independent contributions. In absence of depth scaling, i.e. when  $d_{50} = d_{200}$ , the ratio should be one. For perfect compensation, that is when  $d_{50} = 50 \wedge d_{200} = 200$ , the ratio should be 4. The actual ratio between  $d_{50}$  and  $d_{200}$  is plotted for each participant in [Figure 4.5](#). While two participants show moderate compensation for depth, the majority of participants show no clear sign of scaling of eye movements by fixation distance. This is consistent with the observation that translations are perceived shorter with far compared to near fixations in the world condition.

## 4.4 Discussion

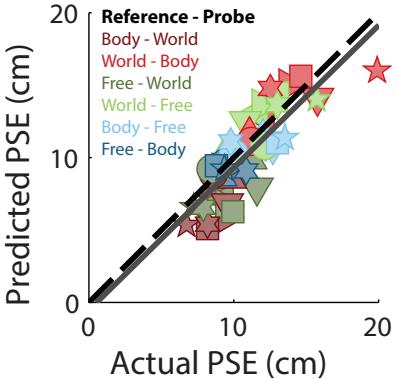
In our previous study, we demonstrated that oculomotor signals play a substantial role in the perception of translation, even in the absence of optic flow or any other visual stimulation (Clemens et al., 2015a). Although the vestibular system provided the most significant contribution, oculomotor signals were shown to account for about 20% of the overall percept. Because these experiments were performed with a single fixation depth, it was not clear whether the brain weighted the oculomotor signal in a depth-dependent manner when using it as a translation cue, or merely uses the signal as a rudimentary cue to self-motion. In the present study we tested between these two possibilities.

We assessed translation perception during both body- and world-fixed fixation at two different fixation depths. Our results show that self-motion was underestimated when comparing far and near fixation trials in the world-fixed condition, which argues against a proper scaling of the eye movement signal. Fixation depth did not influence translation perception during body-fixed fixation (where eye movements are virtually absent).

To quantify the relative depth-dependent scaling of eye movements for nearby and far away fixation targets, we fitted a straightforward linear model to the perceptual responses based on the oculomotor behaviour across four conditions. While two participants show partial scaling, the other six participants did not show any sign of scaling. Thus, we conclude that while oculomotor signals provide a robust cue to translation perception, they are not properly scaled by fixation depth.

### 4.4.1 Relation to other studies

In our previous experiment we compared translation perception with body-fixed versus world-fixed fixations at near depth (50 cm) only (Clemens et al., 2015a). Figure 4.6 shows how well our  $\alpha_{50}$  parameter explains the data in our previous paper. The positive correlation between the actual PSEs in the previous study and those predicted using the model of the present paper ( $\rho = 0.60$ ,  $p = 0.06$ ) adds confidence to the



**Figure 4.6:** Predicted versus actual PSEs for Clemens (2015a) using parameter  $\alpha_{50}$  from the present paper. A data point (symbol) is shown for each participant (symbol shape) and condition (symbol colour) pair, following the same colour scheme as in Figure 3.2: Body-world comparison; body reference, dark red; world reference, light red. World-free comparison; world reference, light green; free reference, dark green. Body-free comparison; body reference, dark blue; free reference, light blue.

parameter values presented here. The average difference between the values found here and those reported previously (see Table 4.2) is  $12 \pm 8$  percent-points, which is relatively small given the independent measurements.

The function of the LVOR is to keep the eyes stable in the world during linear translation (Paige, 1989; Busettini, Miles, Schwarz & Carl, 1994; Paige et al., 1998). Because it also needs to scale with fixation depth (Angelaki, 2004), it is possible that the LVOR and self-motion perception have the same underlying signal. Because of the visual fixation point in our paradigm, visual following mechanisms may even augment the LVOR compensation. If the oculomotor signal generated by the LVOR is used for self-motion perception, one would expect that the LVOR compensation at 50 and 200cm closely relate to the corresponding oculomotor weights in the present study (see Figure 4.5 and Table 4.2). To further explore this, we derived the LVOR gains for 50 and 200 cm from Paige et al. (1989) and computed their expected depth ratio,  $d_{200}/d_{50}$ . This ratio, 1.87, is in between the ratio of the 6 participants who did not show any sign of scaling ( $\frac{d_{200}}{d_{50}} = 1.07 \pm 0.16$ ) and the 2 participants that did show scaling ( $\frac{d_{200}}{d_{50}} = 3.06 \pm 0.27$ ), suggesting that our effect might share a pathway with the LVOR.

#### 4.4.2 Alternative explanations

It is important to point out that while the vestibular signal and thus noise is constant for a given translation distance, the noise in the associated oculomotor estimate might change with fixation distance because the magnitude of the eye movement is modulated by fixation depth and may show signal-dependent noise.

If this were the case, the nearby world-fixed fixations would cause larger eye movements with more noise compared to the far world-fixed fixations with smaller eye movements. The oculomotor based translation estimate would be weighted less for near versus far fixation, potentially explaining the partial compensation for fixation depth we have observed. In addition, the noise levels in the oculomotor estimate could also depend on fixation depth itself: the retinal displacement of a world stationary fixation point decreases with fixation depth, making it less informative about the amount of self-motion. The noise level in the oculomotor estimate would therefore be higher for far away compared to nearby fixation points. For world stationary targets, this would predict an underestimation of self-motion while fixating far away compared to nearby, which is in line with our observations. However, it also predicts a similar effect for body stationary targets. As no such effects between the near and far body stationary fixation targets have been observed, we consider it an unlikely alternative explanation.

Could the lack of scaling be explained by how participants perceive the distance of the fixation points? Because the difference between body- and world-fixed fixation points is reduced at far fixation distances, the lack of scaling could - in theory - be explained by participants incorrectly perceiving both the body- and world-fixed far fixation points as being body-fixed. We consider this an unlikely explanation, because the target displacement associated with a world-fixed target was between 0.3 and 8.5 degrees in our experiment, which is easily perceived. This adds confidence to our claim that eye movements influence self-motion perception, but with moderate to no scaling for fixation depth.



---

---

***CHAPTER 5***

---

*Visual stability across  
combined eye and body  
motion*

**Abstract** In order to maintain visual stability during self-motion, the brain needs to update any ego-centric spatial representations of the environment. Here, we use a novel psychophysical approach to investigate how, and to what extent, the brain integrates visual, extraocular, and vestibular signals pertaining to this spatial update. Participants were oscillated sideways at a frequency of 0.63 Hz while keeping their gaze fixed on a stationary light. When the motion direction changed, a reference target was shown either in front or behind the fixation point. At the next reversal, half a cycle later, we tested updating of this reference location by asking participants to judge whether a briefly flashed probe was shown to the left or right of the memorised target. We show that updating is not only biased, but that the direction and magnitude of this bias depend on both gaze and object location, implying that a gaze-centred reference frame is involved. Using geometric modelling, we further show that the gaze-dependent errors can be caused by an underestimation of translation amplitude, by a bias of visually perceived objects towards the fovea (i.e., a foveal bias), or by a combination of both.

---

This chapter has been published as

**Clemens, I.A.H.**, Selen, L.P.J., Koppen, M. and Medendorp, W.P. (2012). Visual stability across combined eye and body motion. *Journal of Vision*, 12(12):8 1-11.

## 5.1 Introduction

A typical characteristic of human vision is that the position of the retina is constantly changing due to eye, head, and/or body movements. Yet, even during such self-motion, we retain a sense of whether visual objects are stable or moving with respect to an earth-centric reference frame (see e.g. Wallach, 1987). This capability is essential for a correct percept of the world and the maintenance of visual stability.

Achieving visual stability is a complex process because visual signals are coded with respect to gaze and not in an earth-fixed reference frame. When the visual scene lacks earth-centric landmarks, the brain should distinguish which changes in retinal input result from real world movement and which from eye movement. The usual view, dating back to Von Helmholtz (1867), is that this is achieved by subtracting the extraretinal signal of eye motion from the retinal image shifts (Wexler, 2005).

Visual stability experiments in which participants made head-fixed saccades suggest that efference copies of the outgoing motor commands serve this purpose. Neurons in the frontal eye fields and the lateral intraparietal area demonstrate pre-saccadic shifts of receptive fields, elicited by an efference copy (Duhamel, Colby & Goldberg, 1992; Kusunoki & Goldberg, 2003). These gaze-centred shifts could allow the brain to anticipate and cancel out the changes in retinal input due to the saccade (Sommer & Wurtz, 2006). Also, fMRI studies have reported evidence for shifting receptive fields in the human brain (Medendorp, Goltz, Vilis & Crawford, 2003).

Despite these important insights, head-fixed saccades are only one of a multitude of movements that are made in real life. Moving the body, like when driving a car, puts severe challenges on the mechanism underlying visual stability. In this case, when the body is translated passively, vestibular feedback informs the brain about the motion. This information must be combined with efference copies of orbital eye movements to interpret the changes in retinal input. Solving this problem is geometrically complicated because, during eye and body motion, the changes in retinal input depend nonlinearly on the depth and direction of objects that make up the retinal image, as in motion parallax (Medendorp, Tweed & Crawford, 2003).

Recent studies have reported fairly accurate reach or gaze responses to memorised

target locations, presented prior to whole-body translations (see for review: Klier & Angelaki, 2008; Medendorp, 2011). However, such studies do not map one-to-one to the mechanisms of visual stability. First, the requirement of a motor response may invoke different processing mechanisms, which may be subject to different constraints. Second, motor response studies probe the system after the limb or eye movement, thereby revealing the combined result of all intervening spatial computations and transformations needed to guide the action.

In the present study, we investigate visual stability across simultaneous eye and whole-body motion without involving the motor system. To this end, we used a two-alternative forced choice (2AFC) psychophysical approach in combination with a visual updating paradigm. Participants had to retain object locations during sinusoidal whole-body motion, while keeping their gaze fixed on a world stationary point either in front of or behind the object.

By systematically manipulating the parameters of retinal and extraretinal signals related to body translation, binocular fixation, and object location, we test how the brain integrates these signals for the maintenance of visual stability. Our results show consistent errors in visual stability which strongly depend on the location of the object relative to gaze. Using a modelling approach, we explore possible causes underlying these gaze-centred updating errors.

## 5.2 Methods

### 5.2.1 Participants

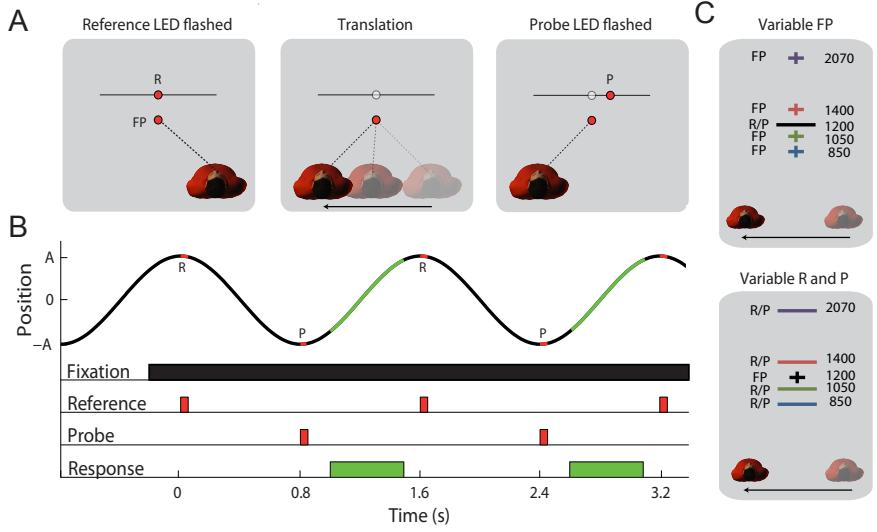
Eight participants (4 male, 4 female), aged between 22 and 41 years, provided written informed consent to participate in the experiment. All participants were free of any known vestibular or neurological disorder and had normal or corrected-to-normal visual acuity. Three participants (the authors) were knowledgeable about the purpose of the experiment, but their results did not differ from the five naïve subjects. Participants never received any feedback about their performance.

### 5.2.2 Setup

A linear sled on a 800 mm track was used to laterally translate participants. The sled, powered by a linear motor (TB15N, Technotion, Almelo, The Netherlands), was controlled by a Kollmorgen S700 (Danaher, Washington DC, USA) drive. The kinematics of the sled were controlled with an accuracy better than  $34 \mu\text{m}$ ,  $2 \text{ mm s}^{-1}$ , and  $150 \text{ mm s}^{-2}$ . The sled was configured such that participants were seated on the sled with the interaural axis aligned with the motion axis. Participants were restrained using a 5-point seat belt and a chin rest. In addition, the head was firmly held in place using an ear-fixed mold. Emergency buttons at both sides of the sled chair enabled subjects to stop the sled motion immediately if needed. Eye movements were recorded using an EyeLink II (SR Research, Kanata, Canada) eye tracking system. Its camera system, which was mounted to the sled, remained stable with respect to the head during the entire experiment. Eye positions were calibrated based on the visual fixations during the experiment, under the assumption that these fixations were accurate.

### 5.2.3 Visual stimuli

Participants had to memorise the location of an earth-centric visual target (reference, R) during half a period of sinusoidal body translation. We tested the quality of this memory by asking them to judge and report the position of a probe stimulus (P) relative to that memorised location, following a psychophysical procedure. The reference and probe stimuli were both presented using a one-dimensional 450 mm wide array, consisting of 180 red light emitting diodes (LEDs), with a spatial separation of 2.5 mm between neighbouring LEDs. The LED array was oriented in parallel with the motion direction of the sled, centred with respect to the sled's trajectory and at the same vertical level as the participant's eyes. It was positioned with an accuracy better than 5 mm, at one of five different distances (850, 1050, 1200, 1400, or 2070 mm) from the participant's eyes in front of the sled. We further positioned an LED at either 850, 1050, 1200, 1400, or 2070 mm in front of the participant, on a virtual line orthogonal to the sled's motion direction and crossing the centre of the LED array. These latter LEDs served as earth-stationary gaze fixation points (FP) during the experiment, so that gaze was directed either behind or in front of the stimulus array. The fixation



**Figure 5.1:** **A** Top-view illustrating three key events within the experiment. Left panel: At the extreme right position, the reference target (R) is flashed. Middle panel: The participant moves while keeping fixation on the fixation point (FP). Right panel: At the left most position, one of the probe locations (P) is flashed. **B** Timing of key events. Within a run of sinusoidal sled motion, participants always fixated the fixation point. At the rightmost point, the reference (R) was flashed for 50 ms. Then, at the extreme left position, a probe (P) was shown for 50 ms. Participants responded whether the probe was presented to the left or right of the reference using a joystick. **C** Locations of fixation points (plus signs) and reference/probe locations (bars) used in our experiment (see also [Table 5.1](#)).

points were displaced vertically by a few mm, such that the fixation point and the LED array did not occlude each other.

#### 5.2.4 Paradigm

The experiments employed a paradigm that studies the constancy of spatial locations during 0.63 Hz sinusoidal whole-body motion in the lateral direction (left-right motion). We tested the effects of body translation ( $T$ , 150 or 300 mm peak-to-peak amplitude), fixation depth ( $FP$ , four spatial locations), and depth of the reference target ( $R$ , four spatial locations) on the quality of perceptual stability. These quantitative data will be interpreted using the geometric framework outlined in the subsection [Model](#) below.

[Figure 5.1](#), panels A and B, illustrate the paradigm in detail; [Figure 5.1C](#) provides an overview of the experimental conditions. The experiment consisted of runs of either 30 or 15 trials. Each run started with the onset of a FP, to be fixated for the entire duration of the run. To avoid discontinuous acceleration at motion onset, sled velocity

was linearly increased over one sinusoidal cycle (see Merfeld, Park, Gianna-Poulin, Black & Wood, 2005 for a similar approach). Once the steady-state sinusoidal motion was reached, participants were tested using a visual updating task (Figure 5.1A). More specifically, at the most rightward position, when the body motion reversed direction, the reference R (here, the centre LED) was presented for 50 ms. When the sled reached the left-most position, again during motion reversal, the participant's estimate for the location of R was tested by displaying another LED, the probe P, for 50 ms. The participant then had to report the location of this probe relative to R in a two-alternative forced choice (leftward, rightward) using a joystick. While we asked participants to respond in a timely manner, we did not explicitly constrain response time. Therefore, the next trial was only presented after a response was given. In practice, most responses were given within half a cycle ( $RT \pm SD = 0.59 \text{ s} \pm 0.09$ , across participants). We used an adaptive algorithm to vary the spatial separation between reference and probe target from trial to trial (Kontsevich & Tyler, 1999), mapping out psychometrically the bias and precision of visual stability across whole-body motion.

Participants were tested in 16 conditions, each comprising a unique combination of translation amplitude (T), visual fixation point (FP), and reference (R) position (see Figure 5.1C). The values used in this experiment are shown in Table 5.1. For each condition we presented 135 trials, which were divided into 4 runs of 30 trials and one run of 15 trials. In every condition 70 out of 135 trials were normal trials, that is, the central LED was used as reference. The other 65 trials in each condition were catch trials, of which 25 trials had the reference location shifted 36 mm to the left of the central led; another 25 had the reference location 36 mm rightward; and in 15 trials a random LED in the stimulus array was taken as the reference location. The catch trials were to prevent participants from simply making repeated stereotypic responses. After each run, the lights were turned on. Following a 30 s break, the experiment resumed automatically. The total experiment was divided into three sessions, tested on different days. Each participant was tested on a total of 2160 trials.

<i>FP</i> (mm)	<i>R/P</i>	<i>T</i> (mm)
1200	850	75 and 150
1200	1050	75 and 150
1200	1400	75 and 150
1200	2070	75 and 150
850	1200	75 and 150
1050	1200	75 and 150
1400	1200	75 and 150
2070	1200	75 and 150

**Table 5.1:** Fixation distance (FP), distance to reference and probe targets (R/P) and translation amplitude for each of the 16 unique visual updating conditions.

### 5.2.5 Data analysis

To prevent effects caused by vergence and/or version eye-movements, we excluded trials in which participants did not maintain fixation within a 3 degree interval around FP, during the time interval starting 100 ms before presenting the reference target and ending 100 ms after cueing the probe. Overall,  $6.4\% \pm 1.7\%$  ( $\pm SD$ ) of all trials were discarded per participant based on these eye movement criteria.

For each condition, we quantified performance by calculating the probability of a rightward response as a function of the location of the probe relative to reference location. We used a maximum likelihood fit of a cumulative Gaussian function to summarise the psychometric data:

$$P(x) = \lambda + (1 - 2\lambda) \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-(y-\mu)^2/2\sigma^2} dy, \quad (5.1)$$

in which  $x$  represents the size of probe displacement. The mean of the Gaussian,  $\mu$ , represents the bias in visual stability (positive  $\mu$  corresponding to a rightward bias). The width of the curve, corresponding to the standard deviation  $\sigma$  of the Gaussian, is inversely related to precision, and serves as a measure of the participant's variability in the visual updating task. Parameter  $\lambda$ , representing the lapse rate, accounts for stimulus-independent errors caused by subject lapses or mistakes and was restricted to small values ( $\lambda < 0.06$ ). Fits were performed using the 'psignifit' Matlab toolbox (Wichmann & Hill, [2001a](#), [2001b](#)).

### 5.2.6 Model

We investigated whether the observed bias could be explained by allowing a gain factor in the processing of the lateral translation by the vestibular system. That is, we assume that  $\tilde{T} = \alpha T$ , where  $\tilde{T}$  is the perceived and  $T$  the actual translation (Medendorp, van Asselt & Gielen, 1999). If the spatial update is performed entirely in a head-centred system, the effect of this gain would be straightforward. The reference flash  $R$  is presented when the sled is in the rightmost position and the following translation of the sled by  $T$  mm to the left in world coordinates amounts to a translation of the world, including the reference point, by  $T$  mm to the right in head-coordinates. Due to the gain of the vestibular system the perceived translation equals  $\alpha T$  mm to the right, leading to a predicted bias of

$$\mu = \tilde{T} - T = (\alpha - 1)T \quad (5.2)$$

in mm on the LED array. Thus, when processed in a head-centred system, the bias would be negative for  $\alpha < 1$ , positive for  $\alpha > 1$ ; it would be proportional to the translation amplitude, but would not depend at all on the reference and fixation point positions. Previous experiments (van Pelt & Medendorp, 2007) have shown that reach targets are updated not in head-centred coordinates, but rather within a gaze-dependent frame of reference. Following up on this, we also model the effect of the translation gain in a gaze-centred system. Let  $OF$  be the vector from the cyclopean eye to the fixation point and, similarly,  $OR$  the vector to the reference point. The translation by  $T$  mm to the left in world coordinates is in head-coordinates well approximated by a rotation of  $OF$  by  $T/|OF|$  radians to the right and a rotation of  $OR$  by  $T/|OR|$  radians to the right. (The approximation is good, since both  $T \ll |OF|$  and  $T \ll |OR|$ . To express the gist of the prediction of the gaze-dependent model, this first-order approximation is very useful; in the actual calculations the precise geometry was used, without noticeable differences.) Consequently, in gaze-centred coordinates (i.e.,  $OF$  fixed straight ahead) the vector  $OR$  rotates by an angle of

$$\phi \approx T\left(\frac{1}{|OR|} - \frac{1}{|OF|}\right) \quad (5.3)$$

radians to the right. In modelling the perceived rotation angle  $\tilde{\phi}$  we again replace  $T$  by  $\tilde{T} = \alpha T$ , but we also have to consider possible biases in the perception of  $|OR|$  and  $|OF|$ . Following previous literature (Gogel, 1977; Medendorp, Tweed & Crawford, 2003), we assume that the depth of the constantly visible fixation point is perceived accurately, i.e.,  $|\tilde{OF}| = |OF|$ , but we allow that the perceived depth of the 50 ms flashed reference stimulus,  $|\tilde{OR}|$ , is biased towards this fixation point depth. Since the depth signals available in this experiment (vergence angle and disparity) express more directly in terms of inverse depth than depth itself. The simplest way to implement such a bias is to model the perceived reference depth as a weighted harmonic mean of the actual reference and fixation depths:

$$\frac{1}{|\tilde{OR}|} = \beta \frac{1}{|OR|} + (1 - \beta) \frac{1}{|OF|} \quad (5.4)$$

where  $\beta = 1$  represents the limiting case of accurate depth perception of the reference stimulus (no bias) and  $\beta = 0$  the limiting case of full "assimilation" to fixation point depth. In total this leads to a perceived rotation angle of

$$\tilde{\phi} = \alpha T \cdot \beta \left( \frac{1}{|OR|} - \frac{1}{|OF|} \right) \quad (5.5)$$

radians to the right. Comparing Equation 5.5 with Equation 5.3 shows that our assumptions amount to a total gain of  $\alpha\beta$  on the rotation angle, with freely interchangeable contributions of the parameters  $\alpha$  and  $\beta$ . We substitute  $\gamma = \alpha\beta$  and note that the resulting bias in angle,  $\tilde{\phi} - \phi$ , is observed as a bias in mm on the LED array at a distance  $|OR|$ :

$$\mu = (\tilde{\phi} - \phi)|OR| = (\gamma - 1)T \left( 1 - \frac{|OR|}{|OF|} \right) \quad (5.6)$$

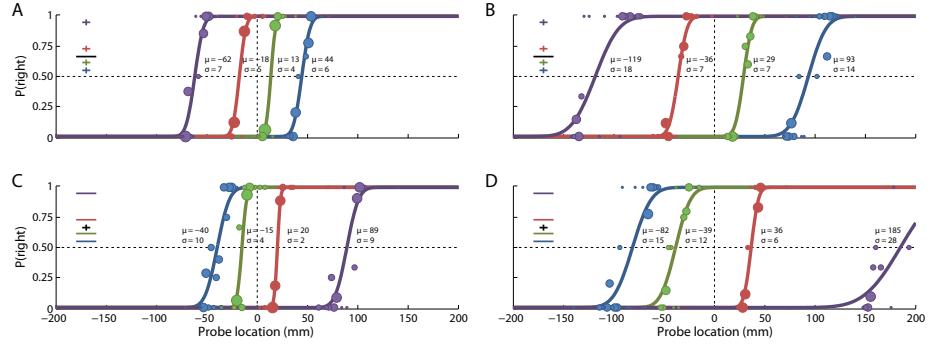
Thus, in the gaze-centred model, the bias is again proportional to translation amplitude, but now it also depends critically on the fixation and reference point positions. In particular, the bias flips sign according to presenting the reference point in front of or behind the fixation point. On top of this, there is an overall (across all conditions) sign dependence on the combined values of the translation gain and fixation depth bias factors.

## 5.3 Results

Participants were tested in an experimental paradigm that studies the stability of spatial locations across combined eye and body motion. The task, illustrated in [Figure 5.1](#), requires that subjects fixate an earth-stationary central fixation point, FP, which is visible throughout the run. At two successive reversals of motion direction, at the right and left excursion point of the sinusoidal motion, a reference (R) and a probe (P) target are briefly flashed. In a two-alternative forced choice task, the participant has to indicate whether the probe location was to the left or to the right of the reference location. The resulting psychometric data provide a quantitative assessment of the bias ( $\mu$ ) and precision ( $\sigma^{-2}$ ) of visual stability across self-motion (see Methods for details). Depending on the stimulus conditions (FP, R, and T), participants may erroneously judge the location of R, and hence provide biased responses.

[Figure 5.2](#) shows the results of a typical participant, plotting the fraction of rightward responses (indicated by the circles) as a function of horizontal probe location relative to the reference. The 16 conditions are split into 4 panels according to the manipulated variable: FP distance (top panels), reference distance (bottom panels) and translation amplitude (left vs. right panels). Data for all individual probes are presented (circles). In an ideal observer, all psychometric functions would constitute a step response centred at zero, indicating no bias and no uncertainty. However, the actual data shows consistent biases and non-zero variance.

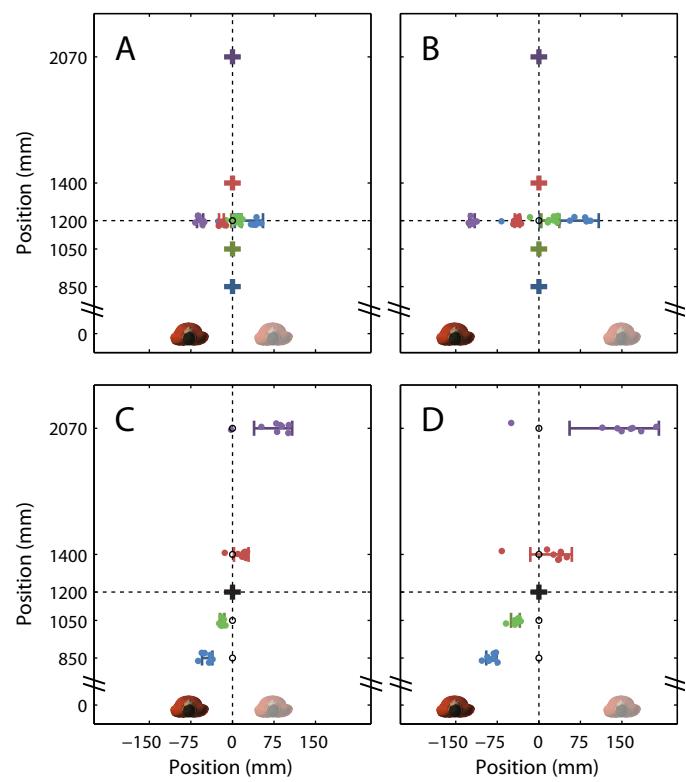
When FP was behind R, we observed a leftward bias (top panels; red and purple curves), that increased when fixation was further away from the reference location (red vs. purple dots). When FP was in front of R (green and blue dots), the opposite pattern was seen. Furthermore, as T increases, psychometric curves move away from zero (t-test;  $t(63) = -4.55, p < 0.05$ ) and become less steep (t-test;  $t(63) = -4.64, p < 0.05$ ), a sign that there is decay in both accuracy and precision (compare left and right panels). Similar biases are observed when keeping FP constant, and varying the location of R, as demonstrated by the bottom panels. We derived estimates of the bias ( $\mu$ ) and corresponding standard deviation ( $\sigma$ ) values in each of the 16 conditions, for all subjects.



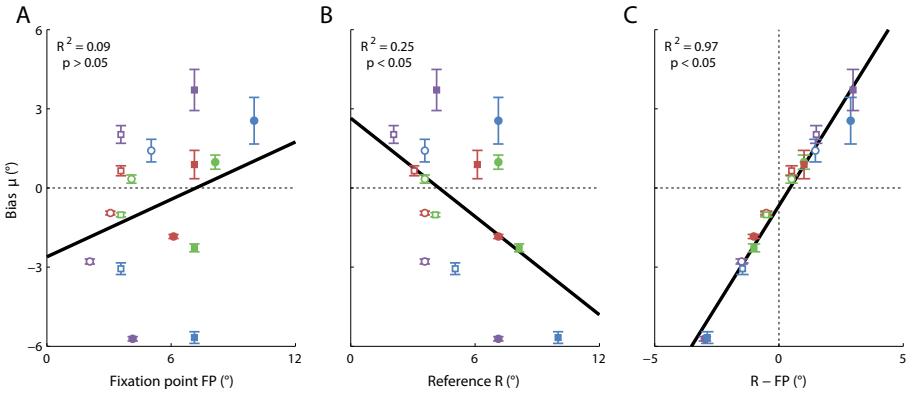
**Figure 5.2:** Performance in one subject (S1). The proportion of rightward responses is plotted against probe location relative to the reference. Size of a data point represents the number of trials tested. Solid lines, best-fit cumulative Gaussians, characterised by bias ( $\mu$ ) and standard deviation ( $\sigma$ ). **A** Constant reference depth, variable fixation depth, 150 mm translation. **B** Constant reference depth, variable fixation depth, 300 mm translation. **C** Variable reference depth, constant fixation depth, 150 mm translation. **D** Variable reference depth, constant fixation depth, 300 mm translation.

Figure 5.3 depicts the bias ( $\mu$ ) for each subject (dots), together with the mean bias  $\pm SD$  across subjects (error bars), in top-view panels. This shows that the pattern in Figure 5.2 holds across all participants, with biases ranging between -126 and 212 mm. Clearly, the bias in updating of the central target increases with  $T$  and depends on FP, reversing for gaze fixation behind versus in front of the R (two top panels). Likewise, when FP was kept constant, the updating bias is not only larger for the larger  $T$ , but also depends on the location of R, with the bias in opposite directions for targets presented in front versus behind fixation (two bottom panels). Taken together, these observations suggest that the location of R relative to gaze, rather than the head-centric locations of FP or R, is a crucial factor in determining the updating bias.

To further analyze these observations, Figure 5.4 plots the bias values ( $\pm SE$  across participants) as a function of gaze fixation FP (panel A), target location R (panel B) and reference location relative to gaze fixation FP - R (panel C). Both the location of FP and R, as well as the bias are expressed in units of degrees instead of millimetres because the former is more closely associated with native visual coordinates. (In practice, however, because of the large distance, visual angles are about proportional to the associated horizontal distances). While in panel A no clear relationship is observed ( $R^2 = 0.09$ ,  $F(1, 14) = 1.32$ ;  $p > 0.05$ ), panel B reveals only a weak linear relationship ( $R^2 = 0.25$ ,  $F(1, 14) = 4.71$ ;  $p < 0.05$ ). However, in panel C the data for all conditions are rearranged such that they fall into a single response curve. A



**Figure 5.3:** Top view of the updating biases ( $\mu$ ). Dots, individual bias values; error bars, averages ( $\pm$ SD) across participants; +, fixation point. **A** to **D**: Conditions as in [Figure 5.2](#).



**Figure 5.4:** Head- versus gaze-centred effects in updating bias ( $l$ ). **A** Average bias (6 SE) across participants plotted against head- centric version angle of fixation point, for each of the 16 conditions. **B** Same data plotted as a function of head-centric angle to the reference location. **C** Same data plotted against the gaze-centric location of the reference target. Open symbols, 150 mm translation. Closed symbols, 300 mm translation. Circles, constant fixation depth. Squares, constant reference depth. Colour scheme as in [Figure 5.1C](#)

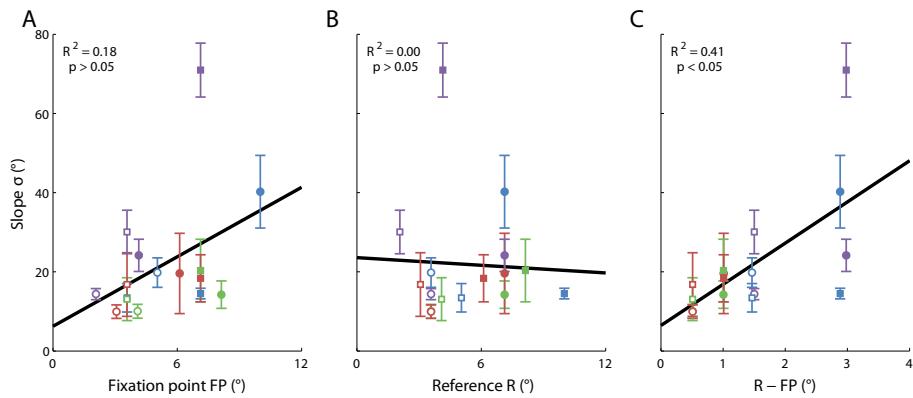
linear fit shows a very strong correlation in this case ( $R^2 = 0.97$ ,  $F(1, 14) = 483$ ,  $p < 0.05$ ). This suggests that the observed errors almost solely depend on the location of  $R$  relative to gaze.

To validate this notion, we fit two different models to explain the updating biases: a head- and gaze-centred model (see Equations 5.2 and 5.6 respectively, in [Methods](#)). Because the updating bias systematically depends on gaze, we expect the gaze-centred model to outperform the head-centred model. Indeed, the RMSE of the gaze-centred model was significantly lower (t-test;  $t(7) = -3.68$ ,  $p < 0.05$ ) than that of the head-centred model. [Table 5.2](#) presents the RMSE values for both models and the fit-results of the gaze-centred model for each participant. According to this latter model, the best-fit value of the gain  $\gamma$  (mean  $0.25 \pm 0.08$  SE) is considerably lower than the ideal value of one. In the Discussion we will address the possible implications of this small value.

Finally, in addition to accuracy, we also quantified the precision of the updated  $R$ . [Figure 5.5](#) shows the standard deviation ( $\sigma \pm \text{SE}$  across participants) of the psychometric functions as a function of either FP (panel A), the head-centred location of  $R$  (panel B) or the gaze-centred location of  $R$  (panel C), in the same format [Figure 5.4](#). No significant effects can be observed in panels A and B ( $R^2 = 0.18$ ,  $F(1, 14) = 3.14$ ,  $p > 0.05$  and  $R^2 = 0.00$ ,  $F(1, 14) = 0.03$ ,  $p > 0.05$  respectively). Panel C shows

Participant	Head-centred RMSE ( $^{\circ}$ )	FIX $\gamma$	ME RMSE ( $^{\circ}$ )
1	0.32	0.13	0.06
2	0.29	0.17	0.05
3	0.13	0.81	0.23
4	0.64	0.29	0.59
5	0.31	0.14	0.07
6	0.31	0.16	0.11
7	0.28	0.23	0.07
8	0.33	0.07	0.10

**Table 5.2:** RMSE values for both models and best-fit values for the gaze-centred model parameter ( $\gamma$ , [Equation 5.6](#)).



**Figure 5.5:** Head- versus gaze-centred effects on standard deviations ( $\sigma$ ). Format as in [Figure 5.4](#)

a significant linear relationship ( $R^2 = 0.41$ ,  $F(1, 14) = 9.68$ ,  $p < 0.05$ ). From this, we conclude that precision decreases for targets that are further or nearer in depth relative to fixation, and therefore, also more peripheral in gaze-coordinates.

## 5.4 Discussion

We investigated how the brain integrates retinal and extraretinal signals in order to maintain visual stability across combined eye and body motion. Participants had to remember the location of a world-fixed reference target, flashed in the periphery, while their body was passively translated and their binocular gaze actively changed in order to fixate a world-stationary target LED. When body motion reversed direction, a probe target was presented and the participant indicated whether it was shown to the left or right of the memorised reference. The resulting psychometric curves revealed substantial biases in the updating of the reference target, which increased with depth from fixation and reversed in sign for reference targets presented at opposite

depths from fixation. In addition, precision of visual stability decreased when the distance between this target and the fixation point increased, likely due to the lower spatial resolution in the retinal periphery (Westheimer, 1982). Geometric modelling suggests that these observations are consistent with spatial updating in a gaze-centred reference frame. In the following, we compare our results to previous work, and explore possible explanations of our observations in context of the gaze-dependent updating model.

#### 5.4.1 Relation to previous studies

To our knowledge, there have been no other studies that have psychophysically investigated perceptual stability during combined eye and body motion. So far, related studies have tested spatial stability using paradigms in which participants make saccades or reaches to previously flashed targets after intervening self-motion (see Klier & Angelaki, 2008; Medendorp, 2011, for review).

For actively generated self-motion, Medendorp, Tweed and Crawford (2003) had participants make saccade-vergence movements to remembered targets that were presented before they made a sidestep. Although their participants initially misperceived the targets, i.e. they underestimated the depths of distant targets and overestimated depths of near targets (Gogel, 1977; Komoda & Ono, 1974; Philbeck & Loomis, 1997), they accurately compensated for the intervening motion in the updating of the perceived target location, following the required non-linear updating patterns. Similar observations were made in relation to the updating of spatial locations across active self-motion for reaching (Admiraal, Keijsers & Gielen, 2004; Flanders, Daghstani & Berthoz, 1999; Medendorp et al., 1999; van Pelt & Medendorp, 2007). Compared to the present study, compensation for active intervening whole body motion was substantially better in all these studies.

Regarding passively induced self-motion, previous work by Israël and Berthoz (1989) and more recent observations by Klier and Angelaki (2008) showed that human participants can also update the locations of saccade targets for passive whole body motion. Similar experiments in non-human primates have also demonstrated compensation for translational motion in the updating of saccadic space (Li et al., 2005). Although

the amount of compensation depended on the depth of fixation, it was typically less than geometrically required (see their Figure 4B), as in the present results. The same experiments were also conducted in labyrinthectomised monkeys, showing that their updating is even more compromised (Li & Angelaki, 2005; Wei, Li, Newlands, Dickman & Angelaki, 2006). This suggests that otolith information interacts with visual information to update saccade goals.

Thus, in view of previous studies, our results are consistent with the notion that spatial ability is better maintained across active compared to passive body motion, perhaps due to the presence of efference copies of motor commands during active motion. Furthermore, based on the present findings it seems that perceptual updating is worse when compared to the action-oriented updating in previous studies. Should this be interpreted in favour of the proposal that visuospatial updating is organised in distinct processing pathways, one for conscious perception and one for the control of action (Goodale & Milner, 1992)? We do not want to suggest this. There may be other factors that contribute to the relatively low updating performance in the present study. Using geometric models (i.e., Equations 5.2 and 5.6) we will now explore such factors in more detail.

#### 5.4.2 Modelling implications

In order to systematically explore possible explanations for the updating performance found in the present study, we now return to the head- and gaze-centred models of the updating mechanism presented in Equations 5.2 and 5.6 respectively. These models were inspired by the models proposed by van Pelt and Medendorp (2007), with the addition of the possibility of a foveal bias. In the head-centred model (Equation 5.2), the updating bias is proportional to the translation amplitude, but independent of reference and/or fixation point positions. However, since our data show a clear and systematic dependence on these positions (see Figures 5.2-5.4), this model is not viable.

This leaves us with the gaze-centred model of Equation 5.6, which incorporates these dependencies. Estimating the overall gain parameter  $\gamma$  in this gaze-centred model yielded a mean value of  $\gamma = 0.25$  across participants. Since this  $\gamma$  is the product of

parameters  $\alpha$  and  $\beta$  (see [Equation 5.5](#)), this entails that at least one of these parameters must be considerably smaller than one, the veridical value. That is, in the updating process the translation is perceived with a small gain ( $\alpha \ll 1$ ) and/or there is a distinct bias towards fixation depth ( $\beta \ll 1$ ). We now explore the plausibility of these explanations in turn.

For the perception of body translation at least two signals may be important: the vestibular signal from the otoliths and the changes in eye position while tracking the visual FP. Both linear acceleration (peak: 231 cm/s<sup>2</sup>) and frequency (0.63 Hz) were well above the detection thresholds of the otoliths (Benson et al., [1986](#); Yu, Dickman & Angelaki, [2012](#)). Furthermore, the firing rate of otolith afferents increases monotonically with acceleration in our frequency range (Fernández & Goldberg, [1976b](#); Yu et al., [2012](#)), and can therefore be used to correctly decode acceleration. However, this does not mean that further processing of acceleration into a velocity or displacement signal is veridical (Merfeld et al., [2005](#)). In fact, it has been shown that the translational vestibulo-ocular reflex is not perfectly compensatory at the frequency that we have tested. However, when the vestibular signal is complemented by visual following mechanisms, participants are able to maintain fixation (Medendorp et al., [2002](#); Paige et al., [1998](#)). This indicates that a near veridical percept of translation is possible by combining vestibular and eye position information. Yet, higher level processing of the translation signal might still be biased. For instance, the conversion of translated distance into an updating angle might be faulty, and/or the actual updating process itself could misinterpret an otherwise veridical updating angle. It has been shown that near-veridical updating takes place for e.g. reach targets (Henriques, Klier, Smith, Lowy & Crawford, [1998](#); van Pelt & Medendorp, [2007](#)) where errors are attributed to the reference frame transformation instead. This suggests that the gaze-centred remapping process itself, which is thought to drive spatial updating, is not biased.

Thus, when considering previous work, it is most likely the higher level processing of the translation signal that governs the observed biases. One such processing step concerns the problem of attributing visual motion to either self-motion or object-motion (Von Helmholtz, [1867](#)). If this attribution is flawed, it can have a profound influence on updating and might be the cause of our low updating gain. Support for

this idea is found in work by Dyde and Harris (2008) who showed that participants make such attribution errors, in particular in conditions of passive translation and darkness, both of which apply to our study. In the active translation studies mentioned earlier, this effect is likely diminished by the presence of an efference copy that helps in disambiguating self-motion from object-motion.

A further explanation for our low overall gain is that depth perception of the reference point is biased ( $\beta << 1$ ). Because the reference and probe lights were flashed for only 50ms at the zero velocity points of the sinusoidal motion and the head is unable to move relative to the body, depth perception of these lights is likely to be compromised. Actually, the spatial updating process that takes place in our experiment can alternatively be described in terms of a Bayesian model. To represent the brain's assumption that, lacking any precision information, the depth of peripheral stimuli is at or close to fixation point depth, such a model will involve a prior distribution centred at this fixation depth. The full specification of such a Bayesian model is beyond the scope of this paper. Here, we have opted for a more straightforward geometrical modelling approach (Equations 5.2 to 5.6), in which such a foveal depth bias appears in Equation 5.4 with the weight  $1 - \beta$ . While such foveal influences have been reported previously (Brenner, Mamassian & Smeets, 2008; Mateeff & Gourevich, 1983), for this to be the sole explanation for our low gain would require the foveal bias to be 80%, which is quite extreme.

In conclusion, we have shown systematic biases in visual stability across combined eye and body movements. These biases are consistent with a gaze-centred updating model, with simple gain factors on both translation and depth perception.



---

---

***CHAPTER 6***

---

***Summary and discussion***

Navigating through the environment evokes complex changes of visual, auditory, vestibular, tactile and motor inputs to the brain. Yet, despite these motion induced changes of input, we perceive the world as a stable reality, maintain an integrated sense of where we are, how we are oriented, and are able to track and act rather effortlessly upon surrounding objects. How this ability comes about is also the topic of this thesis. The objective of the research project described in this thesis was to build computational models and perform thorough psychometric testing in order to unravel the physical and biological constraints on the interaction between the vestibular and other sensory systems for spatial orientation and self-motion perception. This resulted in the following contributions to the field:

**Chapter 2** Statistical optimality can account for the way body somatosensory, neck proprioceptive and vestibular signals are integrated in spatial orientation perception. While these sensors cannot be examined in isolation, by using optimality as a starting point, their noise properties can be determined and linked to clinical deficits that are seen in particular patient groups.

**Chapter 3** Oculomotor signals influence self-motion perception, even in the absence of optic flow or other visual stimulation, and even when in conflict with vestibular information.

**Chapter 4** This oculomotor signals should be regarded as rudimentary cue to self-motion perception; it is not veridically scaled by fixation distance in the perception of body translation.

**Chapter 5** Self-motion signals interact with the dynamic perception of external world. Errors that arise in this process suggest the use of a gaze-centred reference frames in the underlying computations.

In the following sections, we will provide a detailed summary of each result.

## 6.1 Multisensory processing in spatial orientation

Many studies have shown that the brain combines noisy sensory signals in a statistically optimal way. This is normally done by showing that the noise levels of each separate modality can predict performance in combined conditions. In spatial orientation it is difficult to measure sensory signals in isolation as one cannot switch off the vestibular sense when measuring somatosensory contributions. In [Chapter 2](#) we therefore used a statistically optimal integration model as a starting point, and attempted to account for the perceptual differences found when probing head-in-space versus body-in-space orientation. Using a psychometric approach, we tested both the perception of body tilt (subjective body tilt; SBT) as well as the perception of visual vertical (subjective visual vertical; SVV) in seven participants. Because both the SVV and SBT make use of the same sensory signals, we were able to fit a 7-parameter probabilistic model to the response data. One of the estimated parameters represented the noise of the neck proprioceptors. This allowed us to independently confirm that the derived values for neck noise matched those that were measured in isolation. We further validated our model by showing that predictions made by our model are consistent with previously published deficits in vestibular and somatosensory patients. We conclude that Bayesian computations can account for the typical differences in spatial orientation judgements associated with different task requirements. In a follow-up to this work, this approach was recently applied to a patient population with complete vestibular loss ([Alberts, Selen, Verhagen & Medendorp, 2015](#)). Performance in those patients was similar to that of controls, suggesting that the sensory weights had shifted from the vestibular to the somatic sensors.

## 6.2 Eye movements influence self-motion perception

Eye movement typically accompany self-motion in order to minimise retinal slip and maximise dynamic visual acuity. In [Chapter 3](#) we investigated whether these eye movements also have a reversed role, by serving as a cue for self-motion perception. To address this question, we asked participants to compare perceived translation distances from two successive, passive, lateral whole-body translations. Eye movements during these translations were either world-stationary or body-stationary. Results show that translations were perceived shorter with body-fixed gaze compared to world-fixed gaze, indicating that eye movements indeed influence self-motion perception. Using a linear model, we estimated the relative contribution of the vestibular versus the eye movement based displacement signal: the eye movement based displacement signal contributes approximately 25 percent to the perceived motion. We independently validated the model by successfully predicting the effects of eye movements on self-motion perception during trials in which the eye movements were unconstrained. This further shows that eye movement signals influence self-motion perception, even in the absence of visual stimulation, and even when oculomotor and vestibular estimates are in conflict, e.g. during body-fixed gaze. We hypothesise that adverse consequences of this seemingly inflexible arrangement are minimal under natural conditions because eye movements and self-motion are highly correlated, and because eye movements are most often accompanied by veridical optic flow cues to self-motion.

## 6.3 Partial compensation for fixation depth in self-motion perception

If eye movements are used in the estimation of self-motion magnitude, the brain should also take the accompanying fixation depth into account for a veridical translation estimate. The reason is that the amplitude of these eye movements, for the same physical translation, depends on the depth of fixation: when fixating faraway they are small-

ler than when fixating nearby in the world. In [Chapter 4](#) we investigated whether the brain indeed takes fixation depth into account when using eye movements to augment self-motion perception. Participants had to judge self-motion during different eye movement conditions in the absence of full-field optic flow. In a 2-AFC task, similar to the one used in [Chapter 3](#), participants indicated whether the second of two successive passive lateral whole-body translations was longer or shorter than the first. During each translation, participants fixated either a nearby or far away target, which was either body- or world-stationary. Results show that the perceived translations were shorter for nearby world-fixed gaze compared to faraway world-fixed gaze, indicating that eye movements are not properly scaled in self-motion perception. Together with the observation that self-motion perception is not affected by the depth of a body stationary fixation target, we conclude that eye movements are merely a rudimentary cue to self-motion, with a compensation for fixation depth that is partial at best.

## 6.4 Gaze-dependent effects in spatial updating

The brain also needs self-motion signals to update ego-centric spatial representations of the environment. In [Chapter 5](#) we investigate how, and to which extent, the brain integrates the various self-motion signals for the spatial update. Participants were oscillated sideways while keeping gaze fixed on a stationary target. When the motion direction changed, a reference target was shown either in front or behind the fixation point. Half a cycle later, at the next reversal, we tested updating of this reference location by asking participants to judge whether a briefly flashed probe was shown to the left or right of the memorised target. Results show that both the direction and magnitude of the bias in spatial updating depends on the location of the object being updated with respect to gaze, implying that a gaze-centred reference frame is involved. We further show that these biases can be caused by an underestimated of translation amplitude, a bias of visually perceived objects towards the fovea, or by a combination of both.

## 6.5 Concluding remarks

All experimental chapters in this thesis report on subjects' spatial perception. In [Chapter 2](#) we use an optimal statistical integration model of all the contributing sensory systems to explain both the bias and precision of these percepts. In these models the contributions of the senses are weighted according to their uncertainty. In the chapters on lateral displacement perception ([Chapters 3 and 4](#)) we also use weightings of the vestibular and eye-movement based estimates of displacement. However, these weights are only based on the observed biases in self-motion perception. Ideally one would use the same approach as in [Chapter 2](#) to model and analyse these self-motion perception data. We have made initial efforts to arrive at such a model, but there appear at least two major challenges. First, the geometry of the self-motion perception experiments makes that the straightforward Gaussian distributions from the spatial orientation model become less well defined, skewed distributions and we would have to use particle filter models to run our simulations. Second, those models contain many free parameters for the sensory modalities and priors. With the current set of experiments we do not have enough data to fit these parameters in a consistent manner. Future research should be done to derive a complete optimal integration account of self-motion perception.

In [Chapter 5](#) we show that eye movements elicited by fixating a world-stationary target are taken into account when updating the remembered location of previously seen targets. If the underlying signals used for self-motion perception as found in [Chapters 3 and 4](#) are also used for spatial updating, the type of fixation (world- versus body-stationary) should have an effect on the observed biases in spatial updating as well. Indeed, results from a pilot experiment in the early stage of this thesis suggest that fixation type influences spatial updating performance under translation (Clemens, Selen & Medendorp, [2010](#)). Participants had to reach towards remembered targets after an intervening translation. Preliminary results suggest that reaching errors are large when fixating a body-stationary target during translation, while no such errors exist while fixating world-stationary targets (Clemens et al., [2010](#)), suggesting that indeed the eye-movement based self-motion cues are also used in spatial updating.

In conclusion, I hope to have made new advances in the understanding of the mechanisms for spatial orientation and self-motion perception. Of course, there are many remaining questions for further study. These studies should not only address the computational and theoretical mechanisms but also on the neural implementation and pathways that can be found in the brain.



---

## Bibliography

- Admiraal, M. A., Keijser, N. L. W. & Gielen, C. C. A. M. (2004). Gaze affects pointing toward remembered visual targets after a self-initiated step. *Journal of Neurophysiology*, 92(4), 2380–2393.
- Alais, D. & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, 14, 257–262.
- Alberts, B. B., Selen, L. P. J., Verhagen, W. I. M. & Medendorp, W. P. (2015). Sensory substitution in bilateral vestibular a-reflexic patients. *Physiological Reports*.
- Andersen, G. J. & Braunstein, M. L. (1985). Induced self-motion in central vision. *Journal of Experimental Psychology: Human Perception and Performance*, 11(2), 122.
- Angelaki, D. E. (2004). Eyes on target: what neurons must do for the vestibuloocular reflex during linear motion. *Journal of Neurophysiology*, 92, 20–35.
- Angelaki, D. E. & Cullen, K. E. (2008). Vestibular system: the many facets of a multimodal sense. *Annual Reviews Neuroscience*, 31, 125–150.
- Angelaki, D. E. & Hess, B. J. M. (2005). Self-motion induced eye movements: effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6(12), 966–976.
- Aubert, H. (1861). Eine scheinbare bedeutende drehung von objecten bei neigung des kopfes nach rechts oder links. *Archiv für pathologische Anatomie*

- und Physiologie und für klinische Medicin*, 20(3–4), 381–393.
- Battaglia, P. W., Jacobs, R. A. & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Journal of the Optical Society of America A*, 20(7), 1391–1397.
- Bays, P. M. & Wolpert, D. M. (2007). Computational principles of sensorimotor control that minimize uncertainty and variability. *Journal of Physiology*, 578, 387–396.
- Benson, A. J., Hutt, E. C. & Brown, S. F. (1989). Thresholds for the perception of whole body angular movement about a vertical axis. *Aviation, Space and Environmental Medicine*, 60, 205–213.
- Benson, A. J., Kass, J. R. & Vogel, H. (1986). European vestibular experiments on the spacelab-1 mission: 4. thresholds of perception of whole-body linear oscillation. *Experimental Brain Research*, 64(2), 264–271.
- Bisdorff, A. R., Wolsley, C. J., Anastasopoulos, D., Bronstein, A. M. & Gresty, M. A. (1996). The perception of body verticality (subjective postural vertical) in peripheral and central vestibular disorders. *Brain*, 119, 1523–1534.
- Bortolami, S. B., Pierobon, A., DiZio, P. & Lackner, J. R. (2006). Localization of the subjective vertical during roll, pitch, and recumbent yaw body tilt. *Experimental Brain Research*, 173, 364–373.
- Bremmer, F., Klam, F., Duhamel, J. R., Ben Hamad, S. & Graf, W. (2002). Visual-vestibular interactive responses in macaque ventral intraparietal area (vip). *European Journal of Neuroscience*, 16(8), 1569–1586.
- Brenner, E., Mamassian, P. & Smeets, J. B. J. (2008). If i saw it, it probably wasn't far from where i was looking. *Journal of Vision*, 8(2), 1–10.
- Bronstein, A. M., Yardley, L., Moore, A. P. & Cleeves, L. (1996). Visually and posturally mediated tilt illusion in parkinson's disease and in labyrinthine defective subjects. *Neurology*, 47, 651–656.
- 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, 484–494.
- Carriot, J., Brooks, J. X. & Cullen, K. E. (2013). Multimodal integration of self-motion cues in the vestibular system: Active versus passive translation. *Journal of Neuroscience*, 33(50), 19555–19566.
- Carriot, J., Bryan, A., DiZio, P. & Lackner, J. R. (2011). The oculogyral illusion: retinal and oculomotor factors. *Experimental Brain Research*, 209(3), 415–423.
- Chen, A., DeAngelis, G. C. & Angelaki, D. E. (2010). Macaque parieto-insular vestibular cortex: Responses to self-motion and optic flow. *Journal of Neuroscience*, 30(8), 3022–3042.
- Chen, A., DeAngelis, G. C. & Angelaki, D. E. (2011). Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *Journal of Neuroscience*, 31(33), 12036–12052.
- Clark, B. & Graybiel, A. (1963). Perception of the postural vertical in normals and subjects with labyrinthine defects. *Journal of Experimental Psychology*, 65, 490–494.
- Clark, B. & Graybiel, A. (1964). Perception of the postural vertical following prolonged bodily tilt in normals and subjects with labyrinthine defects. *Acta Oto-laryngologica*, 58, 143–148.
- Clemens, I. A. H., de Vrijer, M., van Gisbergen, J. A. M. & Medendorp, W. P. (2011). Multisensory processing in spatial orientation: An inverse probabilistic approach. *Journal of Neuroscience*, 31(14), 5365–5377.
- Clemens, I. A. H., Selen, L. P. J., Koppen, M. & Medendorp, W. P. (2012). Visual stability across combined eye and body motion. *Journal of Vision*, 12(12).
- Clemens, I. A. H., Selen, L. P. J., MacNeilage, P. R. & Medendorp, W. P. (2015a). Integration of ocular and vestibular signals for self-motion perception in darkness. *In preparation*.
- Clemens, I. A. H., Selen, L. P. J., MacNeilage, P. R. & Medendorp, W. P. (2015b). Translation perception is modulated by eye movements that are

*partially scaled by fixation depth.*

- Clemens, I. A. H., Selen, L. P. J. & Medendorp, W. P. (2010). Oculo-motor contributions to visuospatial updating during passive whole-body motion. *2010 Neuroscience Meeting Planner, Society for Neuroscience*, 76.1/QQ5.
- Daunton, N. & Thomsen, D. (1979). Visual modulation of otolith-dependent units in cat vestibular nuclei. *Experimental Brain Research*, 37, 173–176.
- de Kleyn, A. & Versteegh, C. (1933). über verschiedene formen von ménieres syndrom. *Deutsche Zeitschrift für Nervenheilkunde*, 132, 157–189.
- de Vrijer, M., Medendorp, W. P. & van Gisbergen, J. A. M. (2008). Shared computational mechanism for tilt compensation accounts for biased verticality percepts in motion and pattern vision. *Journal of Neurophysiology*, 99, 915–930.
- de Vrijer, M., Medendorp, W. P. & van Gisbergen, J. A. M. (2009). Accuracy-precision trade-off in visual orientation constancy. *Journal of Vision*, 9(2).
- Dichgans, J. & Brandt, T. (1978). Visual-vestibular interaction: Effects on self-motion perception and postural control. In *Perception* (pp. 755–804). Springer.
- Dichgans, J., Diener, H. C. & Brandt, T. (1974). Optokinetic-graviceptive interaction in different head positions. *Acta Oto-laryngologica*, 78, 391–398.
- Dobson, A. J. (2001). *An introduction to generalized linear models*. (Second ed.). Chapman & Hall/CRC.
- Dokka, K., MacNeilage, P. R., DeAngelis, G. C. & Angelaki, D. E. (2013). Multisensory self-motion compensation during object trajectory judgments. *Cerebral cortex*.
- Duhamel, J. R., Colby, C. L. & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, 255(5040), 90–92.
- Dyde, R. T. & Harris, L. R. (2008). The influence of retinal and extra-retinal

- motion cues on perceived object motion during self-motion. *Journal of Vision*, 8, 1–10.
- Eggert, T. (1998). *Der einfluss orientierter texturen auf die subjektive visuelle vertikale und seine systemtheoretische analyse*. Unpublished doctoral dissertation, Munich Technical University.
- Einhäuser, W., Schumann, F., Bardins, S., Barti, K., Böning, G., Schneider, E. & König, P. (2007). Human eye-head co-ordination in natural exploration. *Network*, 18(3), 267–297.
- Ernst, M. O. & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415, 429–433.
- Ernst, M. O. & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, 8, 162–169.
- Faisal, A. A., Selen, L. P. J. & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4), 292–303.
- Fechner, G. (1860). *Elemente der psychophysik*. Breitkopf und Härtel.
- Fernández, C. & Goldberg, J. M. (1976a). Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. ii. directional selectivity and force-response relations. *Journal of Neurophysiology*, 39(5), 985–995.
- Fernández, C. & Goldberg, J. M. (1976b). Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. i. response to static tilts and to long-duration centrifugal force. *Journal of Neurophysiology*, 39, 970–984.
- Fetsch, C. R., DeAngelis, G. C. & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature reviews neuroscience*, 14(6), 429–442.
- Flanagan, J. R. & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, 424(6950), 769–771.
- Flanders, M., Daghestani, L. & Berthoz, A. (1999). Reaching beyond reach. *Experimental Brain Research*, 126(1), 19–30.
- Flash, T. & Hogan, N. (1985). The coordination of arm movements: an exper-

- imentally confirmed mathematical model. *Journal of Neuroscience*, 5(7), 1688–703.
- Foulsham, T., Walker, E. & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, 51(17), 1920–1931.
- Freeman, T. C. A. & Banks, M. S. (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research*, 38(7), 941–945.
- Freeman, T. C. A., Banks, M. S. & Crowell, J. A. (2000). Extraretinal and retinal amplitude and phase errors during filehne illusion and path perception. *Perception & Psychophysics*, 62(5), 900–909.
- Gibson, J. J., Olum, P. & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *The American journal of psychology*, 68(3), 372–385.
- Glasauer, S. (1995). Linear acceleration perception: frequency dependence of the hilltop illusion. *Acta Oto-Laryngologica*, 115(S520), 37–40.
- Glasauer, S., Schneider, E., Jahn, K., Strupp, M. & Brandt, T. (2005). How the eyes move the body. *Neurology*, 65(8), 1291–1293.
- Gogel, W. C. (1977). An indirect measure of perceived distance from oculomotor cues. *Perception & Psychophysics*, 21(1), 3–11.
- Goldberg, J. (2012). *The vestibular system: A sixth sense*. Oxford University Press.
- Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20–25.
- Grabherr, L., Nicoucar, K., W., M. F. & Merfeld, D. M. (2008). Vestibular thresholds for yaw rotation about an earth-vertical axis as a function of frequency. *Experimental Brain Research*, 186, 677–681.
- Grace Gaerlan, M., Alpert, P. T., Cross, C., Louis, M. & Kowalski, S. (2012). Postural balance in young adults: The role of visual, vestibular and somatosensory systems. *Journal of the American Academy of Nurse Practitioners*, 24(6), 375–381.

- Gray, H. & Lewis, W. H. (1918). *Anatomy of the human body*. Lea and Febiger.
- Graybiel, A. (1952). **Oculogravie illusion.** *Archives of Ophthalmology*, 48, 605–615.
- Green, A. M., Meng, H. & Angelaki, D. E. (2007). **A reevaluation of the inverse dynamic model for eye movements.** *Journal of Neuroscience*, 27, 1346–1355.
- Gu, Y., Angelaki, D. E. & DeAngelis, G. C. (2008). **Neural correlates of multi-sensory cue integration in macaque mstd.** *Nature Neuroscience*, 11(10), 1201–1210.
- Guedry, F. E. & Harris, C. S. (1963). Labyrinthine function related to experiments on the parallel swing. *Research report. Naval School of Aviation Medicine, NSAM-86*(Project MR005.13-6001; subtask 1), 1–32.
- Harris, L. R., Jenkin, M. & Zikovitz, D. C. (2000). **Visual and non-visual cues in the perception of linear self motion.** *Experimental Brain Research*, 135(1), 12–21.
- Hayhoe, M. & Ballard, D. (2014). **Modeling task control of eye movements.** *Current Biology*, 24(13), R622–R628.
- Henn, V., Young, L. R. & Finley, C. (1974). Vestibular nucleus units in alert monkeys are also influenced by moving visual fields. *Brain Research*, 71(1), 144-149.
- Henriques, D. Y., Klier, E. M., Smith, M. A., Lowy, D. & Crawford, J. D. (1998). Gaze-centered remapping of remembered visual space in an open-loop pointing task. *Journal of Neuroscience*, 18(4), 1583–1594.
- Hopkins, W. G. (2011, feb). *A new view of statistics*. Retrieved from <http://www.sportsci.org/resource/stats/>
- Israël, I. & Berthoz, A. (1989). Contribution of the otoliths to the calculation of linear displacement. *Journal of Neurophysiology*, 62(1), 247–263.
- Jacobs, R. A. (1999). **Optimal integration of texture and motion cues to depth.** *Vision Research*, 39, 3621–3629.
- Jaeger, R., Kondrachuk, A. V. & Haslwanter, T. (2008). **The distribution**

- of otolith polarization vectors in mammals: comparison between model predictions and single cell recordings. *Hearing Research*, 239, 12–19.
- Jarchow, T. & Mast, F. W. (1999). The effect of water immersion on postural and visual orientation. *Aviation, Space, and Environmental Medicine*, 70, 879–886.
- Kaptein, R. G. & van Gisbergen, J. A. M. (2004). Interpretation of a discontinuity in the sense of verticality at large body tilt. *Journal of Neurophysiology*, 91, 2205–2214.
- Kelly, D. (1962). Information capacity of a single retinal channel. *Information Theory, IRE Transactions on*, 8(3), 221–226.
- Kitazaki, M. & Sato, T. (2003). Attentional modulation of self-motion perception. *Perception*, 32(4), 475–484.
- Klier, E. M. & Angelaki, D. E. (2008). Spatial updating and the maintenance of visual constancy. *Neuroscience*, 156(4), 801–818.
- Knill, D. C. & Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27, 712–719.
- Koch, K., McLean, J., Segev, R., Freed, M. A., Berry, M. J., Balasubramanian, V. & Sterling, P. (2006). How much the eye tells the brain. *Current Biology*, 16(14), 1428–1434.
- Komoda, M. K. & Ono, H. (1974). Oculomotor adjustments and size-distance perception. *Attention, Perception & Psychophysics*, 15(2), 353–360.
- Kontsevich, L. L. & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39(16), 2729–2737.
- Kording, K. P. & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427, 244–247.
- Kusunoki, M. & Goldberg, M. E. (2003). The time course of perisaccadic receptive field shifts in the lateral intraparietal area of the monkey. *Journal of Neurophysiology*, 89(3), 1519–1527.
- Land, M. F. & Lee, D. N. (1994). Where we look when we steer. *Nature*, 369(6483), 742–744.

- Landy, M. S., Maloney, L. T., Johnston, E. B. & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, 35, 389–412.
- 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.
- Lepecq, J.-C., Giannopulu, I. & Baudonnier, P.-M. (1995). Cognitive effects on visually induced body motion in children. *Perception*, 24, 435–435.
- Li, N. & Angelaki, D. E. (2005). Updating visual space during motion in depth. *Neuron*, 48(1), 149-158.
- Li, N., Wei, M. & Angelaki, D. E. (2005). Primate memory saccade amplitude after intervened motion depends on target distance. *Journal of Neurophysiology*, 94(1), 722–733.
- Lorincz, E. N. & Hess, B. J. M. (2008). Dynamic effects on the subjective visual vertical after roll rotation. *Journal of Neurophysiology*, 100, 657–669.
- MacNeilage, P. R., Banks, M. S., Berger, D. R. & Bültlhoff, H. H. (2007). A bayesian model of the disambiguation of gravitoinertial force by visual cues. *Experimental Brain Research*, 179, 263–290.
- MacNeilage, P. R., Banks, M. S., DeAngelis, G. C. & Angelaki, D. E. (2010). Vestibular heading discrimination and sensitivity to linear acceleration in head and world coordinates. *Journal of Neuroscience*, 30(27), 9084–9094.
- MacNeilage, P. R., Turner, A. H. & Angelaki, D. E. (2010). Canal-otolith interactions and detection thresholds of linear and angular components during curved-path self-motion. *Journal of Neurophysiology*, 104(2), 765–773.
- MacNeilage, P. R., Zhang, Z., DeAngelis, G. C. & Angelaki, D. E. (2012). Vestibular facilitation of optic flow parsing. *PLoS One*, 7(7), e40264.
- Mast, F. & Jarchow, T. (1996). Perceived body position and the visual horizontal. *Brain Research Bulletin*, 40, 393–397.
- Mateeff, S. & Gourevich, A. (1983). Peripheral vision and perceived visual direction. *Biological Cybernetics*, 49(2), 111–118.

- Mayne, R. (1974). A systems concept of the vestibular organs. In *Vestibular system part 2: Psychophysics, applied aspects and general interpretations* (pp. 493–580). Springer.
- McGuire, L. M. & Sabes, P. N. (2009). Sensory transformations and the use of multiple reference frames for reach planning. *Nature Neuroscience*, 12, 1056–1061.
- McHenry, M. Q. & Angelaki, D. E. (2000). Primate translational vestibuloocular reflexes. ii. version and vergence responses to fore-aft motion. *Journal of Neurophysiology*, 83(3), 1648–1661.
- McKenna, G. J., Peng, G. C. & Zee, D. S. (2004). Neck muscle vibration alters visually perceived roll in normals. *Journal of the Association for Research in Otolaryngology*, 5, 25–31.
- Medendorp, W. P. (2011). Spatial constancy mechanisms in motor control. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1564), 476–491.
- Medendorp, W. P., Goltz, H. C., Vilis, T. & Crawford, J. D. (2003). Gaze-centered updating of visual space in human parietal cortex. *Journal of Neuroscience*, 23(15), 6209–6214.
- Medendorp, W. P., Tweed, D. B. & Crawford, J. D. (2003). Motion parallax is computed in the updating of human spatial memory. *Journal of Neuroscience*, 23(22), 8135–8142.
- Medendorp, W. P., van Asselt, S. & Gielen, C. C. (1999). Pointing to remembered visual targets after active one-step self-displacements within reaching space. *Experimental Brain Research*, 125(1), 50–60.
- Medendorp, W. P., van Gisbergen, J. A. M. & Gielen, C. C. A. M. (2002). Human gaze stabilization during active head translations. *Journal of Neurophysiology*, 87(1), 295–304.
- Merfeld, D. M., Park, S., Gianna-Poulin, C., Black, F. O. & Wood, S. (2005). Vestibular perception and action employ qualitatively different mechanisms. i. frequency response of vor and perceptual responses during trans-

- lation and tilt. *Journal of Neurophysiology*, 94(1), 186–198.
- Merfeld, D. M. & Young, L. R. (1995). The vestibulo-ocular reflex of the squirrel monkey during eccentric rotation and roll tilt. *Experimental brain research*, 106(1), 111–122.
- Merfeld, D. M. & Zupan, L. H. (2002). Neural processing of gravitoinertial cues in humans. iii. modeling tilt and translation responses. *Journal of neurophysiology*, 87(2), 819–833.
- Mittelstaedt, H. (1983). A new solution to the problem of the subjective vertical. *Naturwissenschaften*, 70, 272–281.
- Mittelstaedt, H. (1992). Somatic versus vestibular gravity reception in man. *Annals of the New York Academy of Sciences*, 656(1), 124–139.
- Mittelstaedt, H. (1997). Interaction of eye-, head-, and trunk-bound information in spatial perception and control. *Journal of Vestibular Research*, 7, 283–302.
- Mittelstaedt, H. (1998). Origin and processing of postural information. *Neuroscience & Biobehavioral Reviews*, 22, 473–478.
- Nelson, J. G. (1968). Effect of water immersion and body position upon perception of the gravitational vertical. *Aerospace Med*, 39, 806–811.
- Nesti, A., Barnett-Cowan, M., MacNeilage, P. R. & Bültlhoff, H. H. (2014). Human sensitivity to vertical self-motion. *Experimental Brain Research*, 232(1), 303–314.
- Olshausen, B. A. et al. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583), 607–609.
- Orgel, L. E. (1968). Evolution of the genetic apparatus. *Journal of Molecular Biology*, 38(3), 381–393.
- Ormsby, C. C. & Young, L. R. (1977). Integration of semicircular canal and otolith information for multisensory orientation stimuli. *Mathematical Biosciences*, 34(1), 1–21.
- Paige, G. D. (1989). The influence of target distance on eye movement responses during vertical linear motion. *Experimental Brain Research*, 77(3), 585–

593.

- Paige, G. D., Telford, L., Seidman, S. H. & Barnes, G. R. (1998). Human vestibuloocular reflex and its interactions with vision and fixation distance during linear and angular head movement. *Journal of Neurophysiology*, 80(5), 2391–2404.
- Paige, G. D. & Tomko, D. L. (1991). Eye movement responses to linear head motion in the squirrel monkey. i. basic characteristics. *Journal of Neurophysiology*, 65(5), 1170–1182.
- Philbeck, J. W. & Loomis, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1), 72–85.
- Pickles, J., Comis, S. & Osborne, M. (1984). Cross-links between stereocilia in the guinea pig organ of corti, and their possible relation to sensory transduction. *Hearing research*, 15(2), 103–112.
- Purves, D. (2012). *Neuroscience*. Sinauer Associates.
- Ramat, S., Straumann, D. & Zee, D. S. (2005). Interaural translational vor: Suppression, enhancement, and cognitive control. *Journal of Neurophysiology*, 94(4), 2391–2402.
- Ramat, S. & Zee, D. S. (2003). Ocular motor responses to abrupt interaural head translation in normal humans. *Journal of Neurophysiology*, 90(2), 887–902.
- Roditi, R. E. & Crane, B. T. (2012). Suprathreshold asymmetries in human motion perception. *Experimental Brain Research*, 219, 369–379.
- Rodrigues, S. T., Polastri, P. F., Carvalho, J. C., Berela, J. A., Moraes, R. & Barbieri, F. A. (2015). Saccadic and smooth pursuit eye movements attenuate postural sway similarly. *Neuroscience Letters*, 548, 292–295.
- Rosenhall, U. (1972). Vestibular macular mapping in man. *Annals of Otology, Rhinology & Laryngology*, 81(3), 339–351.
- Rosenhall, U. (1974). The vestibular sensory regions in man: a morphological

- study. *Acta Universitatis Upsala*, 191, 1–37.
- Royden, C., Banks, M. & Crowell, J. (1992). The perception of heading during eye movements. *Nature*, 360, 583–585.
- Rushton, S. K. & Warren, P. A. (2005). Perception of object movement during self-movement. In *Electronic imaging 2005* (pp. 473–480).
- Schöne, H. (1964). On the role of gravity in human spatial orientation. *Aerospace Medicine*, 35, 764–772.
- Schöne, H. & Udo de Haes, H. (1968). Perception of gravity–vertical as a function of head and trunk position. *Zeitschrift für vergleichende Psychologie*, 60, 440–444.
- Schwarz, U., Busettini, C. & Miles, F. A. (1989). Ocular responses to linear motion are inversely proportional to viewing distance. *Science*, 245(4924), 1394–1396.
- Sommer, M. A. & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, 444(7117), 374–377.
- Soodak, R. E. & Simpson, J. I. (1988). The accessory optic system of rabbit. i. basic visual response properties. *Journal of neurophysiology*, 60(6), 2037–2054.
- Tarnutzer, A. A., Bockisch, C., Straumann, D. & Olasagasti, I. (2009). Gravity dependence of subjective visual vertical variability. *Journal of Neurophysiology*, 102(3), 1657–1671.
- Tarnutzer, A. A., Bockisch, C. J. & Straumann, D. (2010). Roll-dependent modulation of the subjective visual vertical: contributions of head- and trunk-based signals. *Journal of Neurophysiology*, 103(2), 934–941.
- Telford, L., Seidman, S. H. & Paige, G. D. (1997). Dynamics of squirrel monkey linear vestibuloocular reflex and interactions with fixation distance. *Journal of Neurophysiology*, 78(4), 1775–1790.
- Trousselard, M., Barraud, P.-a., Nougier, V., Raphel, C. & Cian, C. (2004). Contribution of tactile and interoceptive cues to the perception of the direction of gravity. *Cognitive Brain Research*, 20(3), 355–362.

- Udo de Haes, H. A. (1970). Stability of apparent vertical and ocular counter-torsion as a function of lateral tilt. *Perception, & Psychophysics*, 8(3), 137–142.
- Vaitl, D., Mittelstaedt, H. & Baisch, F. (1997). Shifts in blood volume alter the perception of posture. *Int J Psychophysiol*, 27(2), 99–105.
- Vaitl, D., Mittelstaedt, H., Saborowski, R., Stark, R. & Baisch, F. (2002). Shifts in blood volume alter the perception of posture: further evidence for somatic graviception. *Int J Psychophysiol*, 44(1), 1–11.
- Valko, Y., Lewis, R. F., Prieson, A. J. & Merfeld, D. M. (2012). Vestibular labyrinth contributions to human whole-body motion discrimination. *Journal of Neuroscience*, 32, 13537–13542.
- van Beers, R. J., Sittig, A. C. & Denier van der Gon, J. J. (1999). Integration of proprioceptive and visual position-information: an experimentally supported model. *Journal of Neurophysiology*, 81(3), 1355–1364.
- van Beuzekom, A. D., Medendorp, W. P. & van Gisbergen, J. A. (2001). The subjective vertical and the sense of self orientation during active body tilt. *Vision Research*, 41, 3229–3242.
- van Beuzekom, A. D. & van Gisbergen, J. A. M. (2000). Properties of the internal representation of gravity inferred from spatial-direction and body-tilt estimates. *Journal of Neurophysiology*, 84(1), 11–27.
- van den Berg, A. V. & Beintema, J. A. (2000). The mechanism of interaction between visual flow and eye velocity signals for heading perception. *Neuron*, 26, 747–752.
- Vandenbussche, E., Vogels, R. & Orban, G. A. (1986). Human orientation discrimination: changes with eccentricity in normal and amblyopic vision. *Invest Ophthalmol Vis Sci*, 27, 237–245.
- van Pelt, S. & Medendorp, W. P. (2007). Gaze-centered updating of remembered visual space during active whole-body translations. *Journal of Neurophysiology*, 97(2), 1209–1220.
- Vaziri, S., Diedrichsen, J. & Shadmehr, R. (2006). Why does the brain predict

- sensory consequences of oculomotor commands? optimal integration of the predicted and the actual sensory feedback. *The Journal of neuroscience*, 26(16), 4188–4197.
- Vingerhoets, R. A., Medendorp, W. P. & van Gisbergen, J. A. (2008). **Body-tilt and visual verticality perception during multiple cycles of roll rotation.** *Journal of Neurophysiology*, 99(5), 2264–2280.
- Von Helmholtz, H. (1867). *Handbuch der physiologischen optik*. Leopold Voss.
- Wade, N. J. (1968). **Visual orientation during and after lateral head, body, and trunk tilt.** *Attention, Perception & Psychophysics*, 3(3), 215–219.
- Waespe, W., Büttner, U. & Henn, V. (1981). **Visual-vestibular interaction in the flocculus of the alert monkey.** *Experimental Brain Research*, 43, 337 – 348.
- Walker, M. F., Tian, J., Tamargo, R. J., Ying, H. & Zee, D. S. (2010). **The cerebellar nodulus/uvula integrates otolith signals for the translational vestibulo-ocular reflex.** *PLoS ONE*, 5, e13981.
- Wallach, H. (1987). **Perceiving a stable environment when one moves.** *Annual Review of Psychology*, 38, 1–27.
- Warren, P. A. & Rushton, S. K. (2007). **Perception of object trajectory: Parsing retinal motion into self and object movement components.** *Journal of Vision*, 7(11), 2.
- Warren, P. A. & Rushton, S. K. (2009). **Optic flow processing for the assessment of object movement during ego movement.** *Current Biology*, 19(18), 1555–1560.
- Warren, W. H. & Hannon, D. J. (1988). **Direction of self-motion is perceived from optical flow.** *Nature*, 336, 162–163.
- Wei, M., Li, N., Newlands, S. D., Dickman, J. D. & Angelaki, D. E. (2006). **Deficits and recovery in visuospatial memory during head motion after bilateral labyrinthine lesion.** *Journal of Neurophysiology*, 96(3), 1676–1682.
- Wertheim, A. H. (1994). **Motion perception during selfmotion: The direct**

- versus inferential controversy revisited. *Behavioral and Brain Sciences*, 17(02), 293–311.
- Wertheim, A. H., Mesland, B. S. & Bles, W. (2001). Cognitive suppression of tilt sensations during linear horizontal self-motion in the dark. *Perception*, 30(6), 733–742.
- Westheimer, G. (1982). The spatial grain of the perifoveal visual field. *Vision Research*, 22(1), 157–162.
- Wexler, M. (2005). Anticipating the three-dimensional consequences of eye movements. *PNAS*, 102(4), 1246–1251.
- Wexler, M., Panerai, F., Lamouret, I. & Droulez, J. (2001). Self-motion and the perception of stationary objects. *Nature*, 409(6816), 85–88.
- Wichmann, F. A. & Hill, N. J. (2001a). The psychometric function: I. fitting, sampling, and goodness of fit. *Perception & Psychophysics*, 63(8), 1293–1313.
- Wichmann, F. A. & Hill, N. J. (2001b). The psychometric function: II. bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, 63(8), 1314–1329.
- Witkin, H. A. & Asch, S. E. (1948). Studies in space orientation. iv. further experiments on perception of the upright with displaced visual fields. *Journal of Experimental Psychology*, 38(6), 762.
- Wright, W., DiZio, P. & Lackner, J. (2005). Vertical linear self-motion perception during visual and inertial motion: more than weighted summation of sensory inputs. *Journal of Vestibular Research*, 15(4), 185–195.
- Young, L. R., Oman, C. M. & Dichgans, J. M. (1975). Influence of head orientation on visually induced pitch and roll sensation. *Aviation, Space, and Environmental Medicine*, 46(3), 264–268.
- Yu, X.-j., Dickman, J. D. & Angelaki, D. E. (2012). Detection thresholds of macaque otolith afferents. *Journal of Neuroscience*, 32(24), 8306–8316.
- Zhaoping, L. (2006). Theoretical understanding of the early visual processes by data compression and data selection. *Network: Computation in neural*

*systems*, 17(4), 301–334.