# Multisensory integration in self-motion perception

Mark W. Greenlee[1], Sebastian M. Frank[1,6], Mariia Kaliuzhna[2], Olaf Blanke[2], Frank Bremmer[3], Jan Churan[3], Luigi F. Cuturi[4], Paul R. MacNeilage[4] and Andrew T. Smith[5]

[1] Institute of Experimental Psychology, University of Regensburg, Regensburg, Germany
[2] Center for Neuroprosthetics, Laboratory of Cognitive Neuroscience, Ecole Polytechnique Fédérale de Lausanne, EPFL, Switzerland
[3] Department of Neurophysik, University of Marburg, Marburg, Germany
[4] German Center for Vertigo, University Hospital of Munich, LMU, Munich, Germany
[5] Department of Psychology, Royal Holloway, University of London, England
[6] Department of Psychological and Brain Sciences, Dartmouth College, Hanover NH, USA

**Table of Contents**



## Summary

Self-motion perception involves the integration of visual, vestibular, somatosensory and motor signals. This article reviews the findings from single-unit electrophysiology, functional and structural magnetic resonance imaging and psychophysics to present an update on how the human and non-human primate brain integrates multisensory information to estimate one's position and motion in space. The results indicate that there is a network of regions in the nonhuman primate and human brain that processes self-motion cues from the different sense modalities.

## Introduction

We experience our environment via a continuous exchange between our different sense modalities in order to optimally plan and execute behavioral responses to these sensory signals. Oftentimes signals from the environment are weak, noisy and ambiguous. As an example, while riding in a vehicle, retinal motion arises from self motion induced by locomotion, object motion evoked by independently moving objects (other vehicles or pedestrians), eye and head movements in order to monitor the surrounding traffic (e.g. oncoming vehicles, lead car, etc.) and accompanying vestibular and somatosensory sensations. The challenge the brain is confronted with is to analyze these incoming signals and extract an accurate representation of one's self moving in the environment. Multisensory integration of the different incoming sensory signals enhances our ability to more accurately represent the current scene, to draw conclusions about the nature of the objects that lead to these sensory signals and to execute the appropriate behavior to interact with animate (i.e. moving) and inanimate (i.e., stationary) objects in our environment, even when we ourselves are in motion (Sereno and Huang, 2014). Such complex processing of sensory information from different modalities requires efficient integration to be able to respond quickly to the demands placed on us by our environment. Indeed impairment in any one of these sense systems can lead to maladaptation and thus suboptimal behavior (Andersen et al., 1997).

Self-motion perception involves the integration of sensory signals arising from the visual, vestibular, somatosensory and motor systems (Lappe et al., 1999, Bremmer, 2011, Britten 2008, Greenlee, 2000). To differentiate between the retinal image motion evoked by one's own movement from image displacements arising from object motion the brain needs to integrate signals related to head and body motion with motion signals of external objects. Linear acceleration or rotations of the head lead to changes in endolymph flow in the otoliths and semicircular canals of the vestibular organ (Barany, 1907; Lopez et al., 2012). Optic flow, on the other hand, provides rich visual information above self motion in space (Koenderink, 1986). Since the eyes move with respect to the head and the head moves with respect to the trunk, the brain needs to parse these different signal sources to disambiguate self motion in space (i.e., heading; Gibson 1950) from object motion and retinal slip due to eye/head movements (Duffy, 2000: Lappe et al., 1999). The cortical representations of these visual, vestibular and somatosensory inputs, as well as their combinations, have been extensively studied (for reviews,

Angelaki et al., 2011; Lopez & Blanke, 2011, Hitier et al., 2014). In this review we provide an update on progress in research that should deepen our understanding of how the brain combines multisensory cues to disambiguate motion signals arising when the organism itself is in motion. In addition, we will also examine how eye movements affect perceived heading during self-motion perception and how visual-vestibular cues for self motion modulate other sense modalities like touch. We focus primarily on our own results based on the methods we have available in our own laboratories.

**Eye-movement invariant heading encoding in monkeys**

Neurophysiological research over the past thirty years has shown in the animal model, i.e., the macaque monkey, how visual, vestibular, tactile and auditory signals interact to enhance and disambiguate the perception of heading during self-motion. Two cortical areas, i.e. the medio-superior-temporal area (area MST) and the ventral intraparietal area (area VIP), proved to be of specific importance in this context. Neurons in area MST respond to visual and vestibular self-motion signals (Duffy and Wurtz, 1991a, b; Lappe et al., 1996; Duffy, 1998; Bremmer et al., 1999; Page and Duffy, 2003; Yu et al., 2010). Their causal role in heading perception has been confirmed by numerous studies (Gu et al., 2007, 2008, 2010, 2012; Morgan et al., 2008). Neurons in area VIP respond not only to visually simulated and real (vestibular) self-motion, but also to tactile and auditory stimulation (Duhamel et al., 1998; Ben Hamed et al., 2002; Schlack et al., 2002; Bremmer et al., 2002a, b; Avillac et al., 2005, 2007; Chen et al., 2011). Behavioral experiments have likewise demonstrated a causal role of area VIP for heading perception (Zhang et al., 2004; Britten, 2008; Chen et al., 2013). Importantly, functional equivalents of both areas have been identified in human visual cortex (area MST: e.g. Morrone et al., 2000; Dukelow et al 2001 and Huk et al. 2002; area VIP: Bremmer et al., 2001; Sereno and Huang (2006); Wall and Smith, 2008). Accordingly, a better understanding of the processing of self-motion information at the cellular level in the animal model will also advance our understanding of the same processes in humans.

Neurons in macaque areas MST and VIP integrate visual and vestibular self-motion signals with extraretinal or eye-movement information to dissociate self-induced motion from object motion (area MST: e.g. Bradley et al., 1996; Shenoy et al., 1999; Upadhyay et al., 2000; area VIP: Zhang

and Britten, 2011). Such multisensory convergence of self-motion signals is clearly suited to improve heading. Nevertheless, perceptual performance could be further enhanced if single neurons were capable of deducing heading information from visual signals alone. Accordingly, in two recent studies, we investigated if neurons in areas MST and VIP can encode heading solely based on visual signals (Bremmer et al., 2010; Kaminiarz et al., 2014). More specifically, we asked if such neurons would keep their heading selectivity regardless of whether or not the retinal flow resulting from a simulated forward motion was disturbed by superimposed simulated eye movements of various gains.

We performed single unit recordings in three awake, behaving monkeys (Macaca mulatta). All procedures were in accordance with published guidelines on the use of animals in research (European Communities Council Directive 86/609/ECC). Experimental methods followed standard procedures that were described in detail in (Bremmer et al., 2009, 2010; Morris et al., 2012, and Kaminiarz et al., 2014). Optic flow stimuli were back projected onto a tangent screen (90° x 90°) 48 cm in front of the monkey and simulated self-motion of a virtual observer over an extended horizontal plane located 37cm below eye–level. Stimuli simulated self–motion at 1m/s in one of three directions: 30° to the left, straight-ahead, and 30° to the right. These three self-motion directions were combined with three different gains of simulated eye movements: Gain = 0.0 (fixed gaze), Gain = 0.5 (aiming at the natural viewing behavior as indicated by Lappe et al., 1998), and Gain = 1.0 (imitating perfect tracking of a stationary target on the ground plane). These nine different stimulus conditions were presented in pseudo-randomized order across trials and were combined with blocks of trials during which the animals were allowed to perform spontaneous, unrestrained eye movements. Here, the same three different self-motion directions as in the simulated eye-movement condition were presented in pseudo-randomized order.

We recorded 84 MST neurons and 68 VIP neurons from three awake behaving monkeys. About three quarters of the cells revealed a significant stimulus driven response: 64/84 = 76% in area MST and 48/68 = 71% in area VIP (ANOVA on ranks, 9 degrees of freedom [df], $p<0.05$). An example for a response of a neuron from area VIP is shown in Figure 1. The left column shows schematically three of the nine optic flow fields presented to the monkeys. Stimuli represent self-motion to the left, in each row superimposed with one of the three simulated eye-movement behaviors. The resulting flow fields differed markedly, while heading direction, as indicated by the tip of the arrow, was always the same. The right column depicts the responses of a VIP neuron for the three self-motion directions combined with the three eye-movement behaviors.

The neuron responded strongly for self-motion to the left (indicated by the green response curves), irrespective of the underlying simulated eye-movement. Medium responses were observed for movement straight-ahead (blue response curves), while movement to the right (red response curves) induced in all three eye-movement conditions inhibition of the ongoing activity with respect to baseline (Mann-Whitney rank test, $p<0.001$). Responses for a given self-motion direction did not differ across eye movement conditions (ANOVA on ranks, 2 df, $p>0.05$ for each of the three heading directions). Accordingly, this neuron's heading tuning was invariant with respect to the simulated smooth eye-movement.

In order to quantify heading at the population level, we ranked the response strength of each individual neuron for a given self-motion direction across the three simulated eye-movement conditions. Firing rate was determined from a fixed response window covering the whole stimulus duration (2500 ms), shifted by an estimated response latency of 100ms. To give an example, responses shown in Figure 1 were quantified as follows: discharges in the 'no eye movement condition' (Gain = 0.0) were strongest for leftward heading (rank = 1), medium for straight-ahead movement (rank = 2) and lowest for rightward heading (rank = 3). The two other simulated eye-movement conditions (Gain = 0.5 and 1.0) resulted in the same rank order. Hence, the rank-orders for leftward, straight-ahead and rightward self-motion were <1-1-1> (leftward), <2-2-2> (straight-ahead), and <3-3-3> (rightward), respectively.

For a given heading direction, this ranking procedure theoretically could result in 3*3*3=27 different outcomes, which we termed *rank-order triplets*. Three of them were unique (<1-1-1>, <2-2-2>, and <3-3-3>), corresponding to an expected frequency of (1/27)*100% = 3.7% for each of them. Figures 2A and 2B show the distribution of the observed rank-order triplets for areas MST (light orange) and VIP (light green) as well as average data (Figure 2B, light blue). Given data from 64 MST neurons and 48 VIP neurons, this resulted in a set of n = 64*3 + 48*3 = 336 triplets. Peak average discharges for identical heading directions (i.e., triplet <1-1-1>) were observed in 60/336 = 17.9% of the cases. Weakest discharges for identical headings (<3-3-3>) occurred in 62/336 = 18.4% of the cases and medium discharges for identical headings (<2-2-2>) occurred in 44/336 = 13.1% of the cases. Each of these rank-order triplets indicates the invariance of a neuron's heading response with respect to simulated eye movements. Considering the occurrence of these three triplets together, eye-movement invariances were found in 168/336 = 49.4% of the cases. This proportion was significantly larger than would have been expected, if responses for a given heading direction across the different eye-movement conditions had been

independent ($\chi^2$ = 117.5, 1 df, p<0.001). Importantly, we found this over-representation of eye-movement invariance in each of the two areas individually: in 82/192 = 42.7% of the cases in area MST ($\chi^2$ = 49.4, 1 df, p<0.001), and in 84/144 = 58.3% of the cases in area VIP ($\chi^2$ = 72.6, 1 df, p<0.001).

These ranks were obtained from long response-windows (2500ms). In everyday life, such an integration time would be far too long to be functional during navigation. We hence were interested to determine, how long the neurons would take to establish these response patterns. The time-courses for eye-movement invariance <1-1-1> and a random distribution <1-2-3> are shown in Figures 2C and 2D. For increasingly longer integration windows, exponential functions with time constants tau in the order of 400ms could be significantly fitted to the data.

In the single-cell example shown in Figure 1, responses for leftward heading were strongest for all three simulated eye-movement conditions. This VIP neuron was recorded from a monkey's right hemisphere. Based on this finding, we were interested to determine whether or not such tuning for contraversive heading was representative for the populations of neurons in both areas. We determined for each simulated eye-movement condition the distribution of response maxima, i.e. left vs. straight-ahead vs. right. In area MST, we found in 83/192 = 43.2% of the cases maximal activity for contraversive heading, while ipsiversive heading resulted in the strongest response in 67/192 = 34.9% of the cases. Tuning for straight-ahead self-motion was found in only 42/192 = 21.9% of the cases. This overrepresentation for contraversive heading was statistically significant ($\chi^2$ = 3.98, 1 df, p<0.05). Recordings from area VIP revealed an even stronger contraversive bias ($\chi^2$ = 33.35, 1 df, p<0.001). Preferred contraversive heading was found in 97/144 = 67.4% of the cases. Maximum responses for ipsiversive and forward heading were found in only 20/144 = 13.9% and 27/144 = 18.7% of the cases, respectively.

- insert Figure 1 about here –

Forty-nine neurons from area MST and 37 neurons from area VIP were recorded for a sufficiently long time to investigate the self-motion tuning also during real eye movements. In blocks of trials, we removed the central fixation point. As expected from the literature (e.g. Lappe et al., 1998), optic flow stimuli elicited spontaneous eye-movements that often followed the visual motion experienced along the direction of gaze. We ranked also the responses during these

real eye-movements (not shown here). The resulting ranks were combined with the rank-order triplets from the simulated eye-movement condition, resulting in *rank-order quadruplets* (Figure 2E). In area MST, we found a coincidence of heading preferences for simulated and real eye movements (rank-order quadruplet <1-1-1-1>) in 20 out of 147 cases (13.6%). The weakest response for a given heading direction (<3-3-3-3>) was observed in 19/147 = 12.9% of the cases. A medium response for a given heading (<2-2-2-2>) was found in 13/147 = 8.8% of the cases. Each observed number of cases of the three eye-movement invariances occurred significantly more often than would have been expected if tunings had been distributed uniformly (smallest $\chi^2$ value = 8.5, 1 df, $p<0.005$). A similar result was obtained from the population of VIP neurons. Here, response peaks for a given heading direction were observed in 16/111 = 14.4% of the cases. Medium and weakest responses for a given heading direction were found in 18/111=16.2% and 10/111=9% of the cases, respectively. Again, each of these proportions differed significantly from a uniform distribution (smallest $\chi^2$ value = 5.64, 1 df, $p<0.03$). Considering both areas together, eye-movement invariances (<1-1-1-1>, <2-2-2-2>, and <3-3-3-3>) were found in 96/257 = 37% of the cases, i.e. more often than expected if responses across eye movement conditions had been independent ($\chi^2 = 87.81$, 1 df, $p<0.001$).

Eye movements induce predictable distortions of the retinal image. Also from a system's point of view, visual consequences of real eye movements are predictable by means of *efference copy* or *corollary discharge* signals (von Holst and Mittelstaedt, 1950; Sperry, 1950). These signals could be used to obtain a net signal of optic flow as induced purely by self-motion. Such a compensatory mechanism, however, was not possible for simulated eye movement conditions, for which distortions were not predictable. Hence, we were interested in the question, if we could find any indication for predictive signals during real eye movements and (if so) how they would be represented at the neural level. As a first step, we determined for each of the three heading directions the median firing rates as obtained in the four different eye movement conditions. For area MST, for none of the three heading directions a difference in response strength during real and simulated eye-movements was found (repeated measures ANOVA on ranks, 3 df. Left: $p>0.4$, straight-ahead: $p>0.2$, and right: $p>0.9$, respectively). For area VIP, this was also true in almost all of the cases. Only for rightward heading, average discharges during real eye movements were significantly smaller than during simulated eye movements with Gain = 1.0 and 0.0 (each: $p<0.05$).

- insert Figure 2 about here –

In a second step of our population analysis, we computed for each eye-movement condition the response modulation (RM), i.e. the difference between the maximum (preferred heading) and the minimum response (non-preferred heading. Figure 2F). For the population of VIP neurons, most data points fell below the identity line ($\chi^2 = 13.89$, 1 df, $p<0.001$). This means that in most cases, RM during simulated eye-movements was larger than during the real eye-movement, which is indicative of an efference copy signal reducing response modulations during real eye-movements. An analogue result was obtained for area MST ($\chi^2 = 3.98$, 1 df, $p<0.05$).

Gibson (Gibson, 1950) proposed that the visual flow in the optic field surrounding a moving observer contained sufficient information to estimate heading just from invariances in the flow pattern itself. His further suggestion, however, that the focus of expansion in the flow is such an invariant is problematic, since in everyday life the optic flow as sensed by the eyes is superimposed with tracking eye movements which distort the flow structure and degrade the focus of expansion (Figure 3). Such eye movements are reflexively induced by the flow itself, and attempt to stabilize the foveal and parafoveal image (Lappe et al., 1998, 1999; Wei and Angelaki, 2006). Human observers can estimate heading from such complex spiraling flow fields in the absence of a focus of expansion (Warren and Hannon, 1988; Van den Berg, 1993).

We have used the simulated eye movement technique (Warren and Hannon, 1990) to show that neurons in macaque areas MST and VIP respond selectively to heading irrespective of the occurrence of tracking eye movements. Crucially, this invariant response was derived from purely visual mechanisms. Such invariant heading responses require complicated visual tuning of heading detectors. The receptive field (RF) structure of MST and VIP neurons is not fully understood but known to be complex (MST: Yu et al., 2010, VIP: Chen et al., 2014). The existence of neuronal mechanisms that determine heading visually from distorted flow fields was proposed in several neural models for heading detection (Lappe and Rauschecker, 1993, 1994; Perrone and Stone, 1994; Beintema and Van den Berg, 1998). Curiously, these models share the prediction of a bi-circular (RF) structure (Beintema et al., 2004). It will be interesting to see how the receptive fields of invariant heading detectors in areas MST and VIP are structured.

In addition to compensatory mechanisms based purely on visual information, we also found evidence for non-visual or *extraretinal* signals being involved in self-motion processing. In

blocks of trials, monkeys were allowed to freely move their eyes. This resulted typically in reflexive eye-movements being composed of slow tracking phases and fast resetting eye movements. Here, different from the simulated eye-movement conditions, eye movement signals (*efference copy* or *corollary discharge*) were available. Such extraretinal signals are thought to help dissociating self-induced from externally induced motion (Galletti et al., 1990; Erickson and Thier, 1991; Sommer and Wurtz, 2008). We observed an implicit neural signature of such *predictive processing*. Response modulation (RM) between strongest and weakest discharges was generally smallest for the real eye-movement condition as compared to all simulated eye-movement conditions. We suggest that this reduced response modulation is indicative of predictive processing of visual motion information: in the simulated eye-movement condition the response modulation would encode both, the eye-movement induced and the self-motion induced visual signal. On the contrary, the response modulation in the real eye-movement condition would indicate only the net visual signal resulting from the (simulated) self-motion. Further experiments, however, are needed to test this hypothesis.

In our studies, we tested MST and VIP neurons for their responses to simple radial flow fields and to distorted flow fields that simulated an eye movement during self-motion. In half of the cases cell responses compensated for such distortion and kept the same heading selectivity irrespective of the simulated eye movement. Response modulations, i.e., differences between the strongest and the weakest heading response for a given eye-movement condition, were smaller during real as compared to simulated eye movements. This latter finding is indicative of predictive mechanisms involved in the processing of visual self-motion information. Functional equivalents of macaque areas MST and VIP have been identified in humans (Bremmer et al., 2001; Greenlee, 2000; Wall and Smith, 2008). We therefore suggest that eye-movement invariant heading encoding is also at play in the human sensorimotor system during visually based navigation.

**Behavioral studies in humans**

**Visuo-Vestibular Crossmodal Aftereffects in Self-motion Perception**

Numerous psychophysical studies have investigated how visual and vestibular signals contribute to self-motion perception (Bremmer, Klam et al. 2002, MacNeilage et al., 2010; Fetsch, Turner et

al. 2009, Cardin and Smith 2010, Cuturi and MacNeilage 2013, Ni, Tatalovic et al. 2013, Frank, Baumann et al. 2014, de Winkel, Katliar et al. 2015, Kaliuzhna, Prsa et al. 2015), but less research has examined adaptation or calibration across these signals. Characterizing perceptual adaptation provides unique insight into both the architecture and dynamics of physiological processes underlying perception. In particular, extensive research on the visual motion aftereffect, caused by adaptation to a constant visual motion stimulus, has shown that visual motion processing depends on opponent motion channels that adapt with particular temporal dynamics (Sutherland 1961, Barlow and Hill 1963, Barlow 1990). This finding, in turn, has been related to changes in response properties of visual-motion-sensitive neurons thought to underlie visual motion perception (for a review: Mather, Pavan et al. 2008). Similar insights have resulted from characterization of motion aftereffects in other modalities, including auditory (Grantham and Wightman 1979, Shu, Swindale et al. 1993), tactile (Watanabe, Hayashi et al. 2007) and vestibular modalities (Crane 2012, Coniglio and Crane 2014).

Aftereffects also have the potential to shed light on multimodal interactions and the underlying neural mechanisms through characterization of so-called crossmodal aftereffects, i.e. when adaptation to a stimulus in one modality gives rise to an aftereffect in a different modality. For example, Kitagawa and Ichihara (2002) showed adaptation to visual motion led to auditory aftereffect and vice-versa, and Konkle, Wang et al. (2009) demonstrated similar interactions between visual and tactile modalities. Such findings are proposed to be indicative of multisensory neural representation that are more "process-dependent" than modality-dependent (Konkle and Moore 2009).

Following this reasoning, it seems likely that adaptation of visual-vestibular neural populations devoted to the "process" of self-motion estimation (e.g. in area MSTd) could similarly result in aftereffects that transfer between visual and vestibular modalities. Evidence in favor of such visual-vestibular crossmodal aftereffects has been described in several previous reports.

For example, Brandt and coworkers (1974) presented subjects with prolonged (5 seconds to 15 minutes) visual-only stimulation inside a rotating drum that induced the illusion of circular self-motion (i.e. vection). When the visual stimulation stopped and the light was extinguished subjects reported an "after-sensation" of self-motion in darkness in the direction opposite the previously

experienced illusory self-motion, and they also exhibited an afternystagmus (Kommerell and Thiele 1970), which agreed with the perceptual aftereffect.

More recently, Seno, Ito et al. (2010) explicitly asked subjects to rate the strength and duration of the self-motion aftereffect after exposure to sustained visual stimulation simulating linear self-motion which was long enough to elicit vection. They reported persistence of vection in the same direction as the adapted one. The problem with subjective reports is that they do not allow quantification of the aftereffect magnitude and may be vulnerable to biases induced by experimental instructions (for a recent comparsion of vection onset latency and strength for 3D and head-mounted displays see Riecke and Jordan, 2015). These findings are highly suggestive but they provide only a subjective measure of aftereffect strength.

A more objective approach was taken by Crane (2013). In this experiment, subjects were seated on a motion platform experienced small forward or backward movements and had to indicate the direction of the movement, i.e. two-alternative-forced choice, a common method in vestibular psychophysical experiments (Benson, Kass et al. 1986, MacNeilage, Banks et al. 2007, Grabherr, Nicoucar et al. 2008, Crane 2012). The magnitude and direction of the movement was varied from trial to trial in order to find the movement that was equally likely to elicit a forward or backward response. This stimulus magnitude is known as the point of subjective equality (PSE) and indicates the stimulus perceived equal to zero motion. Prior work (Crane 2012) had shown systemic shifts in the PSE when the movement was preceded on each trial by an adapting vestibular stimulus, suggestive of a within-modality vestibular motion aftereffect. However when the adapting stimulus consisted of visually-simulated self-motion only, no shift of the PSE, and therefore no crossmodal aftereffect, was observed. This result was somewhat surprising given reports of visually induced self-motion aftereffects by Brandt and Seno.

Comparisons across studies suggests that an adapter duration longer than 1.5 sec used by Crane (2013) may be required. To investigate this possibility Cuturi & MacNeilage (2014) conducted a study with methods similar to those of Crane (2013), but using different durations for the visual self-motion adapter (Fig. 3A). In the main experiment an adapter duration of 15 secs was used. Optic flow simulated either forward or backward self-motion at 3 m/s through a 3D cloud of randomly placed triangles, each with 0.5 cm base and height. Immediately after the optic flow

adapter was extinguished, the platform moved either forward or backward and subjects indicated the direction they perceived. Movement magnitude and/or direction was varied from trial to trial according to an adaptive procedure (Kontsevich and Tyler 1999). A total of 50 trials were collected per condition and psychometric functions were fit to estimate the PSE. Trials with forward and backward adaptation were collected in separate blocks. In addition, a baseline condition with no visual adaptation was run in a separate block to quantify any pre-existing biases in fore-aft self-motion perception.

- insert Figure 3 about here –

Results for the baseline condition show no significant shift in PSE whereas significant shifts are observed when a 15 sec visual adaptation stimulus is presented before each trial (Fig. 3B). These shifts indicate the amount of movement needed to cancel the self-motion aftereffect. The direction of the shifts suggests aftereffects in the direction opposite the simulated visual self-motion such that a movement in the same direction as the adaptation stimulus is required to cancel the aftereffect. In additional conditions, shorter duration adapters were used but no crossmodal aftereffects were observed (Fig. 3C) suggesting that not just vection (elicited after ~7 sec) but *sustained* vection (e.g. following a 15-sec stimulus) is required in order to elicit these aftereffects.

A trivial explanation of the crossmodal aftereffect is that a visual motion aftereffect is influencing the subsequent perception of self-motion via purely visual pathways. Indeed, recent results show that visual motion aftereffects (i.e. illusory visual motion), rather than real visual motion, are sufficient to elicit postural sway responses (Holten, van der Smagt et al. 2014). To examine this possibility, Cuturi & MacNeilage (2014) also measured the standard visual motion aftereffects in response to the same adaptation stimuli by presenting a test stimulus consisting of visual rather than physical motion. However, the magnitude of the visual and self-motion aftereffects were found to be uncorrelated across subjects (Fig. 3D) supporting the conclusion that the crossmodal aftereffects are not simply a secondary consequence of a visual motion aftereffect leading to perceived self-motion.

These behavioral findings inform our understanding of physiological substrates of self-motion perception. Aftereffects are thought reflect neural calibration to steady stimulation by way of suppressed neural responses to the adapted stimulus (Sutherland 1961, Barlow and Hill 1963, Barlow 1990). Crossmodal aftereffects prove that this calibration is transferred from the visual to the non-visual (most likely vestibular; Valko, Lewis et al. 2012, Priesol, Valko et al. 2014) domain. They provide a measure of the magnitude and temporal dynamics of visual-vestibular interaction at the level of the neurophysiological substrates of self-motion perception. Critically, because aftereffect strength is measured by finding the stimulus necessary to cancel the aftereffect, results can be more directly related to degree of neural suppression (i.e. the inverse of neural activation elicited by the cancellation stimulus). With this tool in hand it is possible to probe the system more deeply to investigate, for example, how interaction, quantified by strength of crossmodal transfer, depends on features of the visual stimulus other than duration, such as field of view, speed, contrast, movement direction, etc.

To identify the neurophysiological substrate of the crossmodal aftereffect, it would be necessary to combine behavioral measurements with neurophysiological techniques. In particular, it would be interesting to identify physiological modulations that correlate with aftereffect strength, either across subjects or across conditions. Future research taking advantage of neural recordings in animal models, brain imaging and/or stimulation techniques could reveal more details about the neural interconnections and temporal dynamics behind the crossmodal aftereffect, thus providing a more comprehensive understanding of the neural substrates of self-motion perception.

**Self-motion perception and visuo-vestibular-somatosensory interactions**

The vestibular system is tuned to detect self-motion, and vestibular signals are integrated with additional information in the form of visual optic flow, as well as tactile and auditory cues (Cohen et al., 1981; Probst et al., 1985; Lackner and DiZio, 2005; Prsa et al., 2012). Such integration is necessary for navigating in the environment by detecting changes in the supporting surface, shifts of body weight, perceiving head orientation on the trunk as well as distinguishing object motion from self-motion (Mergner et al., 1983; Mergner and Rosemeier, 1998; Wexler et al., 2001). The reviewed research explores multisensory integration of vestibular cues in the form of passive whole body rotations, visual stimuli in the form of optic flow, and tactile stimulation.

Extending previous work that has shown visual and vestibular cues for heading to be optimally integrated (Fetsch et al., 2009; Butler et al., 2010), we recently demonstrated similar integrative mechanisms for visual and vestibular cues in the case of angular self-motion (Prsa et al., 2012). In several experiments participants were seated in a human motion platform, which delivered rotation stimuli, while observing a 3-D display delivering a visual stimulus (optic flow). Three conditions were tested: unimodal vestibular (yaw rotations); unimodal visual (yaw rotations simulated by visual optic flow; stationary motion platform); bimodal condition (visual and vestibular rotations occurred simultaneously in opposing directions). On each trial two consecutive rotations of one type (visual, vestibular, or bimodal) were presented and participants had to judge whether the first rotation was bigger or smaller than the second. The results show that participants' discrimination thresholds were always better in the bimodal as compared to the unimodal conditions. In addition, participants' performance in the bimodal empirical condition was predicted by a Bayesian optimal observer model. Thus, participants optimally integrated visuo-vestibular cues signaling self-rotations. In an additional experiment we also demonstrate that visual and vestibular cues are fused, whereby the access to the individual cues is lost and only the integrated percept is retained (Prsa et al., 2012). To further explore the limits of visuo-vestibular integration we tested the impact of additional conflicts between visual and vestibular self-rotation cues. Research in other sensory modalities (such as vision, audition and touch) has shown that multisensory integration is altered when conflict is introduced between the two cues, and can even break down when the conflict is substantial (Wallace et al., 2004; Roach et al., 2006). We investigated whether under comparable visuo-vestibular conflicts for idiothetic self-motion cues (which are never in conflict under natural circumstances) integration still occurs. Performing the task described above (using vestibular yaw rotation) we now tested a visual stimulus simulating rotation either around the roll axis (experiment 1) or the pitch axis (experiment 2). The results in Figure 4 show that participants optimally integrated the conflicting cues in both experiments and thus despite the axial visual-vestibular incongruency (Kaliuzhna et al., 2015). Multisensory integration is thought to occur when a common cause is inferred for the cues (Körding et al., 2007). It is possible that despite the directional conflict a common cause was attributed to the visual and vestibular cues through the formation of a subjective percept of some untested intermediate (e.g. diagonal) direction of self-motion, thus resolving the conflict (in fact, one of our subjects spontaneously reported this). Alternatively, multisensory integration could also occur because, despite directional conflict along the axis of rotation, other characteristics of

the two cues were not in conflict, such as their identical motion profile, their angle of rotation and their speed. Thus, visuo-vestibular combinations remain optimally integrated, possibly due to the functional necessity of merging multiple information sources for an adequate estimation of self-motion.

- Insert Figure 4 about here -

Another important source for computing self-motion, as well as balance maintenance, is provided by the tactile system (Mergner and Rosemeier, 1998). Moreover, do vestibular or visuo-vestibular stimuli modulate touch? Vestibular effects on touch were first documented in patient studies, where caloric vestibular stimulation (CVS) was shown to transiently improve hemianaesthesia (Vallar et al., 1990; Bottini et al., 2005). Subsequent studies confirmed this result in healthy volunteers, demonstrating that CVS and galvanic vestibular stimulation (GVS) improved tactile sensitivity (Ferrè et al., 2011; Ferrè et al., 2013a). CVS and GVS represent, however, a highly artificial way to stimulate the vestibular system and simultaneously activate other processing pathways, such as touch and pain. Our work explored whether natural vestibular stimulation in the form of passive whole-body rotations achieved the same effect of improving tactile sensitivity. Yaw-rotation using the above-described motion platform also allowed us to explore whether a spatial attentional component could play a role in the vestibular-tactile effects previously observed (Bottini et al., 2013; Ferrè et al., 2013b). While participants received yaw-rotation, they also received near-threshold tactile stimuli at their index fingers (Figure 5, A). On every trial a tactile stimulus would either be delivered to the left or the right finger, or no stimulation would occur (catch trials). We tested three conditions: a no-rotation baseline; a congruent condition (where the rotation was in the direction of the finger stimulated); and an incongruent condition (where the rotation was in the direction opposite to the finger stimulated). If attention mediated vestibulo-tactile interactions, higher tactile sensitivity would be expected in the congruent condition with respect to the incongruent condition. Our results show improved tactile sensitivity during rotation, independently of congruency, as compared to the no-rotation baseline (Figure 5, B) (Ferrè et al., 2014). Thus, natural vestibular stimulation improves tactile detection and this effect is independent of spatial attention. Note that previous work employing CVS and GVS could not precisely control for spatial attention: the lateralised thermal and tactile stimulation produced by CVS and GVS could orient spatial attention thus indirectly improving

tactile detection. The natural vestibular stimuli used here provide an input signal, which is balanced between the two hemispheres and affects somatosensation in a spatially unspecific way.

- Insert Figure 5 about here -

Multisensory integration and mandatory fusion of visual and vestibular cues as well as vestibular-tactile interactions are largely driven by the functional necessity for accurate self-motion perception, gaze stabilisation and balance maintenance (Mergner and Rosemeier, 1998; Wexler et al., 2001). Next to self-motion perception, these processes are also important for the perception of one's body in space allowing self-identification with one's body, as well as the construction of a first-person perspective (Ionta et al., 2011; Pfeiffer et al., 2013; Pfeiffer et al., 2014), and are subtended by an anatomical pathway combining visual, vestibular and tactile information already at the level of the brainstem and the thalamus up to the cortex (Lopez and Blanke, 2011). Our work opens an avenue for a more controlled and systematic study of vestibular effects on perception, cognition, and self-consciousness. Exploring the limits of trimodal integration the use of a motion platform provides precise control over the onset, magnitude, and duration of vestibular stimulation and its congruency with visual and tactile signals and is additionally selective for a given semicircular canal. Thus, one direction of future research should explore the characteristics of the vestibular stimulus producing tactile facilitation in terms of strength (different self-motion speeds), timing (different self-motion duration), direction (stimulation around e.g. different rotation axes), and whether the tactile stimulus is occurring at a functionally relevant site (e.g., foot soles vs. fingertips). Trimodal visual-vestibular-tactile interaction also remains underexplored despite the tight link between the three modalities. Future behavioural and imaging work (e.g. recording of somatosensory evoked potentials, Pfeiffer et al., in revision) may also allow us to establish the order as well as the timing and the anatomical locus of these effects in the human brain.

**Functional MRI studies of Vestibular and Visuo-vestibular processing in Humans**

Brain imaging studies of vestibular and visuo-vestibular functions have been hindered by the simple fact that the participant's head should not move during imaging. Artificial stimulation of the vestibular nerve can be achieved by galvanic vestibular stimulation (GVS; Smith et al., 2012)

or by caloric vestibular stimulation (CVS; Frank & Greenlee, 2014) techniques. Below we summarize recent findings using these two techniques to investigate how vestibular stimulation modifies responses in visual areas to visual motion. Using these techniques we can also explore areas that integrate visual and vestibular cues for self-motion perception.

**Visual-vestibular interactions in the human cerebral cortex studied with fMRI and galvanic vestibular stimulation**

A major limitation in the study of how and where visual and vestibular signals interact in the human brain is that natural vestibular stimulation is not possible in an MRI scanner, necessitating the use of artificial methods. In galvanic vestibular stimulation (GVS), a controlled electric current is passed between two electrodes attached to the skin, just behind the ears. This stimulates the 8$^{th}$ cranial nerve, which connects the vestibular organs of the inner ear to the brainstem, resulting in an illusory vestibular sensation. Typically, the stimulation waveform is sinusoidal with a frequency in the region of 1 Hz and a current of about ±2mA. Unlike caloric stimulation, which typically causes a sensation of leftward and rightward translation (see next section), GVS causes a predominant sensation of clockwise or anticlockwise roll.

Several authors have previously used GVS in conjunction with fMRI (e.g. Lobel et al, 1998; Stephan et al, 2005). GVS results in activation of a fairly consistent and well-defined set of cortical regions. The work reported here focussed on area hMST (human MST) because MSTd is a key area for visual-vestibular interaction in macaques (Gu et al, 2006; Takahashi et al, 2007; Fetsch et al, 2013). Area hMST can be defined by dividing the MT complex into two sub-regions on the basis that one portion (hMT) responds mainly to contralateral stimuli but the other (hMST) responds well also to ipsilateral stimuli. The relationship between hMST and macaque MSTd is unclear and hMST defined in this way may in reality contain more than one functional region, but this definition of hMST has been widely used and provides a starting point for vestibular studies.

In a study in which GVS was delivered in complete darkness, Smith et al (2012) explored whether hMT and hMST respond to vestibular stimulation. Three other visual areas were also examined, all areas that are implicated in processing visual cues to self-motion, namely VIP, CSv and V6. These can all be identified based on the fact that they respond well to a standard optic flow stimulus but respond less well or not at all to an array of flow stimuli (Wall & Smith, 2008; Cardin & Smith, 2010). As in the case of hMST, there are uncertainties about the relationships between

some of these areas and their macaque counterparts. VIP corresponds to human VIP of Bremmer et al (2001) but it is not certain that this is functionally homologous with macaque VIP and it is unclear how it maps onto alternative classifications of visual areas in the human intraparietal sulcus. Human V6 probably corresponds well to macaque V6, but V6 as defined with a flow localizer may also include V6A (Pitzalis et al, 2013, 2015). CSv was first defined in the human brain and a homologue has not so far been identified in macaques.

Figure 6A shows the locations of all these visual regions in the brain of a typical participant. Figure 6B shows, again for one participant, that vestibular activity is present in hMST. In a group of participants, vestibular activity was usually present in hMST but, without exception, it was absent in hMT. Vestibular activity was also seen quite strongly and consistently in CSv and sometimes also in VIP. It was not seen in area V6, suggesting that human V6 is not involved in processing vestibular signals, or at least not the analysis of roll. When vestibular activity was present in hMST, it did not fill the hMST region of interest as defined visually but was only evident in the anterior part of it (as in Figure 6B). This is consistent with the idea that hMST encompasses more than one functional region and suggests that not all sub-regions have vestibular input. It is possible, though a matter of speculation, that the part of hMST that has vestibular responses might correspond to macaque MSTd, the part that has vestibular responses in macaques.

The work of Smith et al. (2012) identifies three cortical regions that respond to both visual and vestibular stimuli: hMST, CSv and VIP. A fourth such is area PIC, which was not studied by Smith et al (2012) but was shown by Frank et al (2014) to respond to both types of stimulus. The presence of both types of response does not necessarily imply that integration occurs. Nonetheless, these areas are candidates for visual-vestibular interactions. There is good evidence from Angelaki's group (reviewed by Fetsch et al, 2013) that integration occurs in macaque MSTd. In a series of studies, they moved the animal to create vestibular activity and also presented visual cues to self-motion. They found that many MSTd cells are tuned for direction of self-motion. Some have the same preferred direction for both visual and vestibular stimulation, suggesting that these neurons may be integrating the two signals, and detailed analysis showed that direction tuning is often better during combined stimulation than for either modality alone, a signature of integration. They found other MSTd cells that were tuned for direction in both modalities but with opposite direction preferences, indicating antagonistic comparison of visual and vestibular signals, perhaps

used for discounting self-motion when assessing object motion. Similar cells exist in macaque VIP and also in VPS, which may be a homologue of human PIC.

- Insert figure 6 about here –

To study visual-vestibular interactions with fMRI, Billington & Smith (2015) developed a nulling procedure. They presented participants with a circular patch of white dots on a black background, with all visual cues from the surroundings occluded. In the real world, when one's head moves, the image of a stationary object will move across the retina. If the head is moving yet the retinal image is static, the object must be moving (in synchrony with the head). Consequently, the illusion of self-motion from GVS caused the static patch of dots in this study to appear to move. The dominant sensation from GVS is roll so the patch of dots, which was centrally fixated, appeared to rotate about its centre. Billington & Smith now physically rotated the dots in the opposite direction, so as to cancel the illusory motion (Figure 6C). They then conducted an fMRI experiment in which this nulled motion was presented. They also employed a second condition in which the direction of the physical motion was reversed, so that instead of cancelling the illusory motion, it summed with it. Allowance was made for the effect of VOR when equating retinal motion. Thus, the two stimulus conditions had exactly the same retinal motion and exactly the same vestibular motion, the only difference between them being the relative phase of the two sinusoidal motions. One condition was perceived as stationary, because the two signals cancelled, and the other appeared to be moving quite strongly. This raises the interesting question of whether cortical activity during the nulled condition follows what is perceived or what is happening on the retina. It was found that in all four areas (hMST, hVIP, CSv, PIC), the BOLD response was similar in the two conditions. Retinal motion always caused cortical activity, whether or not the motion was perceived, and no cortical region was found that was active only when motion was perceived.

Given similar response amplitudes for the two conditions, Billington & Smith were able to look at whether the responses to the two stimuli could be distinguished with multi-voxel pattern analysis (MVPA). If they could be distinguished (classified) based on differences in the pattern of activity across voxels in a given cortical area, this would suggest the existence of neurons that are not only responsive to both visual and vestibular stimuli but are also sensitive to the relative phase in which the two stimuli are presented. Conceptually, the two conditions correspond to the congruent and opposite conditions used by Angelaki's group, so applying MVPA should probe

whether a given visual/vestibular area has two neural populations, one with congruent and one with opposite direction preferences.

Figure 6D summarizes the results. In hMST, classification performance reached about 75% correct, where chance performance is 50%. This suggests that human MST does indeed contain neurons that are responsive to both modalities and are sensitive to whether the stimuli are in the same or opposite phase, like those in macaque MSTd. There were two other areas that showed this property, namely VIP and PIC. As a control, the authors looked at primary visual cortex (V1), which is not thought to receive vestibular signals, and found as expected that performance was at chance. They also found that it is at chance in CSv, suggesting that this area perhaps does not integrate the two signals, or at least not in the same way as hMST, VIP and PIC, even though both senses are represented.

In summary, human cortical visual areas hMST, hVIP, CSv and PIC, but not MT or hV6, show vestibular responses. Areas hMST, hVIP and PIC, but not CSv, appear to integrate visual and vestibular cues to direction of self-motion and may contain neurons with direction preferences that are congruent in some cases and opposite in others.

### Organization of human vestibular cortex in lateral sulcus studied with fMRI and caloric vestibular stimulation

The posterior lateral sulcus (also called posterior sylvian fissure) and its surrounding regions (perisylvian cortex) are the major sites of the cortical vestibular network in humans (Lopez and Blanke, 2011; Dieterich & Brandt, 2015). Although the exact location remains disputed, several imaging studies in humans (for meta-analyses see: Lopez et al., 2012; zu Eulenburg et al., 2012) indicate that the central hub of vestibular processing in this network is in an area referred to as parieto-insular vestibular cortex (PIVC). However, it remains an open question, which other areas belong to the vestibular network in posterior lateral sulcus, and which role they play for vestibular processing.

Previous imaging studies have reported distributed activations in lateral sulcus during vestibular stimulation (e.g., Fasold et al., 2002; Dieterich et al., 2003), extending into the posterior end of the sulcus. Interestingly, imaging studies employing visual motion stimuli found evidence for visual motion processing at this posterior site as well, in a region referred to as the posterior

insular cortex area (PIC) (Beer et al. 2009; Biagi et al. 2015; Claeys et al. 2003; Orban et al. 2003; Sunaert et al. 1999). The location of PIC suggests that it might also be part of the vestibular network but this hypothesis has not been tested in the past.

Therefore, as a first step to understanding the organization of the vestibular network in the sylvian fissure, we explored the sensitivity of area PIC to caloric stimuli (Frank et al., 2014; Frank & Greenlee, 2014; Frank et al., in prep.). If PIC is part of the vestibular network it should respond to vestibular information.

Participants were in supine position with eyes open in the MRI-scanner (3-Tesla Siemens Allegra) and fixated on a static point in the screen center, while caloric vestibular stimulation (CVS) was performed. For CVS we used a custom-built MRI-compatible, micro-pump system, where hot (48°C), cold (5°C), or neutral (30°C) water flowed through left and right ear pods leading to differential caloric vestibular stimulation conditions (see Frank & Greenlee, 2014). Periods of bithermal stimulation were always followed by periods of neutral stimulation (warm on both sides). During each trial, participants indicated the presence or absence of self-motion sensations and, if present, the main direction of self motion. The fMRI BOLD response was contrasted between conditions of caloric and neutral stimulation.

Area PIC was localized using its known responsiveness to visual stimuli. Therefore, in a separate experiment, periods with purely (100%) coherent motion vs. static dots were presented (for more information see Frank et al., 2014).

PIC could be localized in all participants using the visual motion localizer. An example of one participant is presented in Figure 7a. After localizing PIC, we could determine its responses to vestibular stimulation. The results showed that PIC was significantly activated during CVS (Frank et al. 2014). Example activations in one participant are presented in Figure 7b.

Based on these results we conclude that area PIC, in addition to the already known PIVC, appears to be part of the cortical vestibular network in lateral sulcus and plays a role in the integration of visual and vestibular motion cues for the perception of self motion. The responsiveness to visual motion sets it apart from area PIVC that does not respond to visual motion (primates: Chen et al. 2010) and might even be suppressed by visual motion cues (humans: Brandt et al. 1998;

Deutschländer et al. 2002; Dieterich et al. 1998; Kleinschmidt et al. 2002). Therefore, previously reported activations in posterior lateral sulcus during self motion induced by visual motion (e.g. Cardin & Smith, 2010; Huang et al., 2015; Uesaki & Ashida, 2015) might be identified with PIC, or at least partially overlap with PIC, rather than PIVC.

In conclusion, we find evidence for a more complex organization of vestibular cortex in posterior lateral sulcus. In addition to area PIVC, there is at least one other area, named PIC, that responds to visual and vestibular motion, presumably supporting the integration of motion-information from the visual and vestibular senses. Future experiments might investigate if other vestibular areas exist in the vicinity of PIVC and PIC and further clarify the organization of the vestibular network in posterior lateral sulcus.

- insert figure 7 about here –

**Conclusions**

The findings reviewed above suggest that sensory signals about self motion are integrated over visual, vestibular, somatosensory and motor systems. To differentiate between self and object motion the brain integrates signals arising from the vestibular system and compares them with those stemming from retinal displacements and from other sensory systems. Since the eyes within the head are also in relative motion, corollary discharge signals arising from the motor commands to move the eye in a certain direction in space with a specific velocity also need to be taken into account when calculating self-motion in space. Using a simulated eye movement technique, we showed above that neurons in areas MST and VIP appear to be able to discount the retinal displacements during eye movements when calculating the focus of expansion of optic flow fields (see Figure 1). Moreover, prolonged stimulation with directed motion in the one sensory system can lead to aftereffects of in the same or other sensory modality, leading to a prominent (but illusory) sensation of self motion in a direction opposite that of the adapting direction (Figure 3). Natural vestibular stimulation leads to direction-invariant enhancements in tactile perception (Figure 5). The results of functional MRI studies using either galvanic (Figure 6) or caloric vestibular stimulation (Figure 7) coupled with visual motion stimulation indicate that the human brain also contains a set of cortical regions that are inter-connected and exchanging

information regarding visual and vestibular cues regarding self motion in space. Exactly how the receptive fields of these multisensory neurons are constructed and what sorts of sensory cues they respond to remains an unresolved research domain requiring further investigation.

**Acknowledgements**

Author ATS thank the Leverhulme Trust for financial support. Authors FB & JC acknowledge that their work was supported by DFG-CRC/TRR-135-A2. Author MWG was supported by German Research Council (DFG, FOR 1075).

**References**


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. http://doi.org/10.1038/nrn1804

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2011). Visual and vestibular cue integration for heading perception in extrastriate visual cortex. The Journal of Physiology, 589(Pt 4), 825–833. http://doi.org/10.1113/jphysiol.2010.194720

Avillac M, Ben Hamed S, Duhamel JR. (2007). Multisensory integration in the ventral intraparietal area of the macaque monkey. J Neurosci 27: 1922-1932.

Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR. (2005). Reference frames for representing visual and tactile locations in parietal cortex. Nat Neurosci 8: 941-949,.

Barany, R. (1907). Physiologie und Pathologie des Bogengangapparates beim Menschen. Leipzig: Franz Deutsche-Verlag.

Barlow, H. B. (1990). A theory about the functional role and synaptic mechanism of visual after-effects. Vision: Coding and Efficiency. e. C.B. Blakemore, Cambridge: Cambridge University Press**:** 363–375.

Barlow, H. B. and R. M. Hill (1963). "Selective sensitivity to direction of movement in ganglion cells of the rabbit retina." Science **139**(3553): 412-414.

Beer, A. L., Watanabe, T., Ni, R., Sasaki, Y., & Andersen, G. J. (2009). 3D surface perception from motion involves a temporal-parietal network. European Journal of Neuroscience, 30(4), 703–713. http://doi.org/10.1111/j.1460-9568.2009.06857.x

Beintema JA, Van den Berg AV, Lappe M. (2004). Circular receptive field structures for flow analysis and heading detection. In: The structure of receptive fields for flow analysis and heading detection (Vaina LM, Beardsley SA, Rushton S, eds), pp 223-248. Norwell, MA, USA: Kluwer Academic Publishers.

Beintema JA, Van den Berg AV. (1998). Heading detection using motion templates and eye velocity gain fields. Vision Res 38: 2155-2179.

Ben Hamed S, Duhamel JR, Bremmer F, Graf W. (2002). Visual receptive field modulation in the lateral intraparietal area during attentive fixation and free gaze. Cereb Cortex 12: 234-245.

Benson, A. J., J. R. Kass and H. Vogel (1986). "European vestibular experiments on the Spacelab-1 mission: 4. Thresholds of perception of whole-body linear oscillation." Exp Brain Res **64**(2): 264-271.

Biagi, L., Crespi, S. A., Tosetti, M., & Morrone, M. C. (2015). BOLD Response Selective to Flow-Motion in Very Young Infants. *PLoS Biology*, *13*(9), e1002260.

Billington J, Smith AT. (2015). Neural mechanisms for discounting head-roll-induced retinal motion. Journal of Neuroscience 35:4851-4856.

Bottini G, Gandola M, Sedda A, Ferrè ER (2013). Caloric vestibular stimulation: interaction between somatosensory system and vestibular apparatus. Frontiers in integrative neuroscience 7.

Bottini G, Paulesu E, Gandola M, Loffredo S, Scarpa P, Sterzi R, Santilli I, Defanti C, Scialfa G, Fazio F (2005). Left caloric vestibular stimulation ameliorates right hemianesthesia. Neurology 65:1278-1283.

Bradley DC, Maxwell M, Andersen RA, Banks MS, Shenoy KV. (1996). Mechanisms of heading perception in primate visual cortex. Science 273: 1544-1547.

Brandt, T., & Dieterich, M. (1999). The vestibular cortex. Its locations, functions, and disorders. Annals of the New York Academy of Sciences, 871, 293–312.

Brandt, T., Bartenstein, P., Janek, A., & Dieterich, M. (1998). Reciprocal inhibitory visual-vestibular interaction. Visual motion stimulation deactivates the parieto-insular vestibular cortex. Brain : a Journal of Neurology, 121 ( Pt 9), 1749–1758.

Brandt, T., J. Dichgans and W. Buchle (1974). "Motion habituation: inverted self-motion perception and optokinetic after-nystagmus." Exp Brain Res **21**(4): 337-352.

Bremmer F, Duhamel J-R, Ben Hamed S, Graf W. (2002a ). Heading encoding in the macaque ventral intraparietal area (VIP). Eur J Neurosci 16: 1554-1568.

Bremmer F, Klam F, Duhamel J-R, Ben Hamed S, Graf W. (2002b ). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). Eur J Neurosci. 16: 1569-1586.

Bremmer F, Kubischik M, Hoffmann KP, Krekelberg B. (2009). Neural dynamics of saccadic suppression. J Neurosci 29: 12374-12383,.

Bremmer F, Kubischik M, Pekel M, Hoffmann KP, Lappe M. (2010). Visual selectivity for heading in monkey area MST. Exp Brain Res 200: 51-60.

Bremmer F, Kubischik M, Pekel M, Lappe M, Hoffmann KP. (1999). Linear vestibular self-motion signals in monkey medial superior temporal area. Ann N Y Acad Sci 871: 272-281.

Bremmer F, Schlack A, Shah NJ, Zafiris O, Kubischik M, Hoffmann K, Zilles K, Fink GR. (2001). Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. Neuron, 29(1):287-96.

Bremmer, F. (2011). Multisensory space: from eye-movements to self-motion. The Journal of Physiology, 589(Pt 4), 815–823.

Bremmer, F., F. Klam, J. R. Duhamel, S. Ben Hamed and W. Graf (2002). "Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP)." Eur J Neurosci **16**(8): 1569-1586.

Bremmer, F., Klam, F., Duhamel, J.-R., Ben Hamed, S., & Graf, W. (2002). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). European Journal of Neuroscience, 16(8), 1569–1586.

Britten KH. (2008). Mechanisms of self-motion perception. Annual Rev Neurosci.; 31:389–410,

Butler JS, Smith ST, Campos JL, Bülthoff HH (2010). Bayesian integration of visual and vestibular signals for heading. Journal of vision 10:23.

Cardin, V. and A. T. Smith (2010). "Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation." Cereb Cortex **20**(8): 1964-1973.

Chen A, Deangelis GC, Angelaki DE. (2013). Functional specializations of the ventral intraparietal area for multisensory heading discrimination. J Neurosci. 33(8):3567-81.

Chen A, DeAngelis GC, Angelaki DE. (2011). Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. J Neurosci 31: 12036-12052.

Chen X, DeAngelis GC, Angelaki DE. (2014). Eye-centered visual receptive fields in the ventral intraparietal area. J Neurophysiol., 112(2):353-61.

Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011). Representation of Vestibular and Visual Cues to Self-Motion in Ventral Intraparietal Cortex. The Journal of Neuroscience, 31(33), 12036–12052.

Claeys, K. G., Lindsey, D. T., de Schutter, E., & Orban, G. A. (2003). A higher order motion region in human inferior parietal lobule: evidence from fMRI. Neuron, 40(3), 631–642.

Cohen B, Henn V, Raphan T, Dennett D (1981). Velocity storage, nystagmus, and visual-vestibular interactions in humans. Annals of the New York Academy of Sciences 374:421-433.

Coniglio, A. J. and B. T. Crane (2014). "Human yaw rotation aftereffects with brief duration rotations are inconsistent with velocity storage." J Assoc Res Otolaryngol **15**(2): 305-317.

Crane, B. T. (2012). "Fore-aft translation aftereffects." Exp Brain Res **219**(4): 477-487.

Crane, B. T. (2013). "Limited interaction between translation and visual motion aftereffects in humans." Exp Brain Res **224**(2): 165-178.

Cuturi, L. F. and P. R. MacNeilage (2013). "Systematic biases in human heading estimation." PLoS One **8**(2): e56862.

Cuturi, L. F. and P. R. MacNeilage (2014). "Optic flow induces nonvisual self-motion aftereffects." Curr Biol **24**(23): 2817-2821.

de Winkel, K. N., M. Katliar and H. H. Bulthoff (2015). "Forced fusion in multisensory heading estimation." PLoS One **10**(5): e0127104.

Dieterich, M., & Brandt, T. (2008). Functional brain imaging of peripheral and central vestibular disorders. Brain : a Journal of Neurology, 131(Pt 10), 2538–2552.

Duffy CJ, Wurtz RH. (1991a ). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J Neurophysiol 65: 1329-1345.

Duffy CJ, Wurtz RH. (1991b ). Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. J Neurophysiol 65: 1346-1359.

Duffy, C.J. (2000) Optic flow analysis for self-movement perception. International Rev Neurobiology. 44, 199 – 218.

Duffy CJ. (1998). MST neurons respond to optic flow and translational movement. J Neurophysiol 80: 1816-1827.

Duhamel JR, Colby CL, Goldberg ME. (1998). Ventral intraparietal area of the macaque: congruent visual and somatic response properties. J Neurophysiol 79: 126-136.

Eickhoff, S. B., Weiss, P. H., Amunts, K., Fink, G. R., & Zilles, K. (2006). Identifying human parieto-insular vestibular cortex using fMRI and cytoarchitectonic mapping. Human Brain Mapping, 27(7), 611–621. http://doi.org/10.1002/hbm.20205

Erickson RG, Thier P. (1991). A neuronal correlate of spatial stability during periods of self-induced visual motion. Exp Brain Res 86: 608-616.

Ferrè ER, Bottini G, Iannetti GD, Haggard P (2013b). The balance of feelings: vestibular modulation of bodily sensations. Cortex 49:748-758.

Ferrè ER, Day BL, Bottini G, Haggard P (2013a). How the vestibular system interacts with somatosensory perception: a sham-controlled study with galvanic vestibular stimulation. Neuroscience letters 550:35-40.

Ferrè ER, Kaliuzhna M, Herbelin B, Haggard P, Blanke O (2014). Vestibular-Somatosensory Interactions: Effects of Passive Whole-Body Rotation on Somatosensory Detection. PloS one 9:e86379.

Ferrè ER, Sedda A, Gandola M, Bottini G (2011). How the vestibular system modulates tactile perception in normal subjects: a behavioural and physiological study. Experimental brain research 208:29-38.

Fetsch CR, DeAngelis GC, Angelaki DE. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. Nature Reviews Neuroscience 14:429-442.

Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. The Journal of Neuroscience 29:15601-15612.

Fetsch, C. R., A. H. Turner, G. C. DeAngelis and D. E. Angelaki (2009). "Dynamic reweighting of visual and vestibular cues during self-motion perception." J Neurosci **29**(49): 15601-15612.

Frank SM, Baumann O, Mattingley JB, Greenlee MW. (2014). Vestibular and visual responses in human posterior insular cortex. Journal of Neurophysiology 112:2481-2491.

Frank, S. M., & Greenlee, M. W. (2014). An MRI-compatible caloric stimulation device for the investigation of human vestibular cortex. Journal of Neuroscience Methods, 235, 208–218.

Galletti C., Battaglini P.P., Fattori P. (1990). 'Real-motion' cells in area V3A of macaque visual cortex. Exp Brain Res 82: 67-76.

Gibson J.J. (1950). The perception of the visual world. Boston: Houghton Mifflin.

Grabherr, L., K. Nicoucar, F. W. Mast and D. M. Merfeld (2008). "Vestibular thresholds for yaw rotation about an earth-vertical axis as a function of frequency." Exp Brain Res **186**(4): 677-681.

Grantham, D. W. and F. L. Wightman (1979). "Auditory motion aftereffects." Percept Psychophys **26**(5): 403-408.

Greenlee, M.W. (2000).Human cortical areas underlying the perception of optic flow: brain imaging studies. Int Rev Neurobiol. 44: 269-92.

Grüsser, O. J., Pause, M., & Schreiter, U. (1990). Localization and responses of neurones in the parieto-insular vestibular cortex of awake monkeys (Macaca fascicularis). The Journal of Physiology.

Gu Y, Angelaki DE, DeAngelis GC. (2008). Neural correlates of multisensory cue integration in macaque MSTd. Nat Neurosci 11: 1201-1210.

Gu Y, DeAngelis GC, Angelaki DE. (2007). A functional link between area MSTd and heading perception based on vestibular signals. Nat Neurosci 10: 1038-1047.

Gu Y, DeAngelis GC, Angelaki DE. (2012). Causal Links between Dorsal Medial Superior Temporal Area Neurons and Multisensory Heading Perception. J Neurosci 32: 2299-2313.

Gu Y, Fetsch CR, Adeyemo B, DeAngelis GC, Angelaki DE. (2010). Decoding of MSTd population activity accounts for variations in the precision of heading perception. Neuron 66: 596-609.

Gu Y, Watkins PV, Angelaki DE, DeAngelis GC. (2006). Visual and Nonvisual Contributions to Three-Dimensional Heading Selectivity in the Medial Superior Temporal Area. Journal of Neuroscience 26:73-85.

Guldin, W. O., & Grüsser, O. J. (1998). Is there a vestibular cortex? Trends in Neurosciences, 21(6), 254–259.

Holten, V., M. J. van der Smagt, S. F. Donker and F. A. Verstraten (2014). "Illusory motion of the motion aftereffect induces postural sway." Psychol Sci **25**(9): 1831-1834.

Ionta S, Heydrich L, Lenggenhager B, Mouthon M, Fornari E, Chapuis D, Gassert R, Blanke O (2011) Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. Neuron 70:363-374.

Kaliuzhna M, Prsa M, Gale S, Lee SJ, Blanke O (2015) Learning to integrate contradictory multisensory self-motion cue pairings. J Vis 15.

Kaliuzhna, M., M. Prsa, S. Gale, S. J. Lee and O. Blanke (2015). "Learning to integrate contradictory multisensory self-motion cue pairings." J Vis 15(1): 15 11 10.

Kaminiarz A, Schlack A, Hoffmann KP, Lappe M, Bremmer F. (2014). Visual selectivity for heading in the macaque ventral intraparietal area. J Neurophysiol., 112(10):2470-80.

Kitagawa, N. and S. Ichihara (2002). "Hearing visual motion in depth." Nature **416**(6877): 172-174.

Kleinschmidt, A., Thilo, K. V., Büchel, C., Gresty, M. A., Bronstein, A. M., & Frackowiak, R. S. J. (2002). Neural correlates of visual-motion perception as object- or self-motion. NeuroImage, 16(4), 873–882.

Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*(1), 161–179.

Kommerell, G. and H. Thiele (1970). "[Optokinetic short-stimulation nystagmus]." Albrecht Von Graefes Arch Klin Exp Ophthalmol **179**(3): 220-234.

Konkle, T. and C. I. Moore (2009). "What can crossmodal aftereffects reveal about neural representation and dynamics?" Commun Integr Biol **2**(6): 479-481.

Konkle, T., Q. Wang, V. Hayward and C. I. Moore (2009). "Motion aftereffects transfer between touch and vision." Curr Biol **19**(9): 745-750.

Kontsevich, L. L. and C. W. Tyler (1999). "Bayesian adaptive estimation of psychometric slope and threshold." Vision Res **39**(16): 2729-2737.

Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L (2007). Causal inference in multisensory perception. PLoS one 2:e943.

Lackner JR, DiZio P (2005). Vestibular, proprioceptive, and haptic contributions to spatial orientation. Annu Rev Psychol 56:115-147.

Lappe M, Bremmer F, Pekel M, Thiele A, Hoffmann KP. (1996). Optic flow processing in monkey STS: a theoretical and experimental approach. J Neurosci 16: 6265-6285,

Lappe M, Bremmer F, Van den Berg AV. (1999). Perception of self-motion from visual flow. Trends in Cognitive Sciences 3: 329-336.

Lappe M, Pekel M, Hoffmann K-P. (1998). Optokinetic eye movements elicited by radial optic flow in the macaque monkey. J Neurophysiol 79: 1461-1480.

Lappe M, Rauschecker JP. (1993). A neural network for the processing of optic flow from ego-motion in man and higher mammals. Neural Comp 5: 374-391.

Lappe M, Rauschecker JP. Heading detection from optic flow. (1994). Nature 369: 712-713.

Lappe, M., Bremmer, F., & van den Berg AV. (1999). Perception of self-motion from visual flow. Trends in Cognitive Sciences, 3(9), 329–336.

Lobel E, Kleine J, Le Bihan D, Leroy-Willig A, Berthoz A. (1998). Functional MRI of galvanic vestibular stimulation. Journal of Neurophysiology, 80: 2699-2709.

Lobel, E., Kleine, J. F., Leroy-Willig, A., van de Moortele, P. F., Le Bihan, D., Grüsser, O. J., & Berthoz, A. (1999). Cortical areas activated by bilateral galvanic vestibular stimulation. Annals of the New York Academy of Sciences, 871, 313–323.

Lopez, C., & Blanke, O. (2011). The thalamocortical vestibular system in animals and humans. Brain Research Reviews, 67(1-2), 119–146. http://doi.org/10.1016/j.brainresrev.2010.12.002

Lopez, C., Blanke, O., & Mast, F. W. (2012). The human vestibular cortex revealed by coordinate-based activation likelihood estimation meta-analysis. Neuroscience, 1–21.

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. *The Journal of Neuroscience*, *30*(27), 9084-9094.

Mather, G., A. Pavan, G. Campana and C. Casco (2008). "The motion aftereffect reloaded." Trends Cogn Sci **12**(12): 481-487.

Mergner T, Nardi G, Becker W, Deecke L (1983). The role of canal-neck interaction for the perception of horizontal trunk and head rotation. Experimental brain research 49:198-208.

Mergner T, Rosemeier T (1998). Interaction of vestibular, somatosensory and visual signals for postural control and motion perception under terrestrial and microgravity conditions—a conceptual model. Brain Research Reviews 28:118-135.

Morgan ML, DeAngelis GC, Angelaki DE. (2008). Multisensory integration in macaque visual cortex depends on cue reliability. Neuron 59: 662-673.

Morris AP, Kubischik M, Hoffmann K-P, Krekelberg B, Bremmer F. (2012). Dynamics of eye-position signals in the dorsal visual system. Current Biology 22: 173-179.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience*, *3*(12), 1322–1328.

Ni, J., M. Tatalovic, D. Straumann and I. Olasagasti (2013). "Gaze direction affects linear self-motion heading discrimination in humans." Eur J Neurosci **38**(8): 3248-3260.

Orban, G. A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., et al. (2003). Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. Neuropsychologia, 41(13), 1757–1768.

Page WK, Duffy CJ. (2003). Heading representation in MST: sensory interactions and population encoding. J Neurophysiol 89: 1994-2013.

Perrone JA, Stone LS. (1994). A model of self-motion estimation within primate extrastriate visual cortex. Vision Res 34: 2917-2938.

Pfeiffer C, Lopez C, Schmutz V, Duenas JA, Martuzzi R, Blanke O (2013). Multisensory origin of the subjective first-person perspective: visual, tactile, and vestibular mechanisms. PLoS One 8:e61751.

Pfeiffer C, Schmutz V, Blanke O (2014). Visuospatial viewpoint manipulation during full-body illusion modulates subjective first-person perspective. Exp Brain Res 232:4021-4033.

Pitzalis, S., Sereno, M. I., Committeri, G., Fattori, P., Galati, G., Tosoni, A., & Galletti, C. (2013). The human homologue of macaque area V6A. *NeuroImage*, *82*, 517–530.

Pitzalis, S., Fattori, P., & Galletti, C. (2015). The human cortical areas V6 and V6A. *Visual Neuroscience*, *32*, E007.

Priesol, A. J., Y. Valko, D. M. Merfeld and R. F. Lewis (2014). "Motion Perception in Patients with Idiopathic Bilateral Vestibular Hypofunction." Otolaryngol Head Neck Surg **150**(6): 1040-1042.

Probst T, Straube A, Bles W (1985). Differential effects of ambivalent visual-vestibular-somatosensory stimulation on the perception of self-motion. Behavioural brain research 16:71-79.

Prsa M, Gale S, Blanke O (2012). Self-motion leads to mandatory cue fusion across sensory modalities. Journal of neurophysiology 108:2282-2291.

Riecke, B. E., & Jordan, J. D. (2015). Comparing the effectiveness of different displays in enhancing illusions of self-movement (vection). *Frontiers in Psychology*, *6*, 713.

Roach NW, Heron J, McGraw PV (2006). Resolving multisensory conflict: a strategy for balancing the costs and benefits of audio-visual integration. Proceedings of the Royal Society B: biological sciences 273:2159-2168.

Schlack A, Hoffmann KP, Bremmer F. (2002). Interaction of linear vestibular and visual stimulation in the macaque ventral intraparietal area (VIP). Eur J Neurosci 16: 1877-1886.

Seemungal, B. M. (2014). The cognitive neurology of the vestibular system. Current Opinion in Neurology, 27(1), 125–132. http://doi.org/10.1097/WCO.0000000000000060

Seno, T., H. Ito and S. Sunaga (2010). "Vection aftereffects from expanding/contracting stimuli." Seeing Perceiving **23**(4): 273-294.

Sereno, M. I., & Huang, R.-S. (2014). Multisensory maps in parietal cortex. *Current Opinion in Neurobiology*, *24*(1), 39–46.

Shenoy KV, Bradley DC, Andersen RA. (1999). Influence of gaze rotation on the visual response of primate MSTd neurons. J Neurophysiol 81: 2764-2786.

Shu, Z. J., N. V. Swindale and M. S. Cynader (1993). "Spectral motion produces an auditory after-effect." Nature **364**(6439): 721-723.

Smith AT, Wall MB, Thilo KV. (2012). Vestibular inputs to human motion-sensitive visual cortex. Cerebral Cortex 22:1068-1077.

Sommer MA, Wurtz RH. (2008). Brain circuits for the internal monitoring of movements. Annu Rev Neurosci 31: 317-338,.

Sperry RW. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. J Comp Physiol Psychol 43:482-489.

Stephan T, Deutschländer A, Nolte A, Schneider E, Wiesmann M, Brandt T, Dieterich M. (2005). Functional MRI of galvanic vestibular stimulation with alternating currents at different frequencies. NeuroImage 26:721-732.

Sunaert, S., van Hecke, P., Marchal, G., & Orban, G. A. (1999). Motion-responsive regions of the human brain. Experimental Brain Research, 127(4), 355–370.

Sutherland, N. S. (1961). "Figural aftereffects and apparent size." Q. J. Exp. Psychol. **13**: 222-228.

Takahashi K, Gu Y, May PJ, Newlands SD, DeAngelis GC, Angelaki DE. (2007). Multimodal Coding of Three-Dimensional Rotation and Translation in Area MSTd: Comparison of Visual and Vestibular Selectivity. Journal of Neuroscience, 27: 9742-9768.

Uesaki, M., & Ashida, H. (2015). Optic-flow selective cortical sensory regions associated with self-reported states of vection. *Frontiers in Psychology*, *6*, 775.

Upadhyay UD, Page WK, Duffy CJ. (2000). MST responses to pursuit across optic flow with motion parallax. J Neurophysiol, 84: 818-826.

Valko, Y., R. F. Lewis, A. J. Priesol and D. M. Merfeld (2012). "Vestibular labyrinth contributions to human whole-body motion discrimination." J Neurosci **32**(39): 13537-13542.

Vallar G, Sterzi R, Bottini G, Cappa S, Rusconi ML (1990). Temporary remission of left hemianesthesia after vestibular stimulation. A sensory neglect phenomenon. Cortex 26:123-131.

Van den Berg AV. (1993). Perception of heading. Nature 365: 497-498,.

von Holst E, Mittelstaedt H. (1950). Das Reafferenzprinzip. Naturwissenschaften 37: 464-476.

Wall MB, Smith AT. (2008). The representation of egomotion in the human brain. Current Biology 18:191-194.

Wallace MT, Roberson G, Hairston WD, Stein BE, Vaughan JW, Schirillo JA (2004). Unifying multisensory signals across time and space. Experimental Brain Research, 158: 252-258.

Warren WH, Hannon DJ. (1988). Direction of self-motion is perceived from optical flow. Nature 336: 162-163,.

Warren WHJ, Hannon DJ. (1990). Eye movements and optical flow. J Opt Soc Am [A] 7: 160-169.

Watanabe, J., S. Hayashi, H. Kajimoto, S. Tachi and S. Nishida (2007). "Tactile motion aftereffects produced by appropriate presentation for mechanoreceptors." Exp Brain Res **180**(3): 577-582.

Wei M, Angelaki DE. (2006). Foveal visual strategy during self-motion is independent of spatial attention. J Neurosci 26: 564-572.

Wexler M, Panerai F, Lamouret I, Droulez J (2001). Self-motion and the perception of stationary objects. Nature 409:85-88.

Yu CP, Page WK, Gaborski R, Duffy CJ. (2010). Receptive field dynamics underlying MST neuronal optic flow selectivity. J Neurophysiol 103: 2794-2807.

Zhang T, Britten KH. (2011). Parietal area VIP causally influences heading perception during pursuit eye movements. J Neurosci 31: 2569-2575.

Zhang T, Heuer HW, Britten KH. (2004). Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. Neuron 42: 993-1001.

Zu Eulenburg, P., Baumgärtner, U., Treede, R.-D., & Dieterich, M. (2013). Interoceptive and multimodal functions of the operculo-insular cortex: tactile, nociceptive and vestibular representations. NeuroImage, 83, 75–86.

## Figure legends:

Figure 1: Retinal flow fields and neuronal responses. The panels in the left column depict retinal flow fields seen by an observer moving over a ground plane. Heading as indicated by the arrow was always to the left, but the simulated eye movements differed. Monkeys had to fixate a central target (o) in all cases. The right columns shows the time resolved responses of a neuron from area VIP for the three different headings (as indicated by the colored PSTHs) and the three different eye movement conditions.

Figure 2. Population responses. Panels A and B depict the distribution of rank-order triplets for areas MST (light orange) and VIP (light green) as raw numbers (A) and in percent (B). Panels C and D depict the time-courses of the establishment of eye-movement invariance (<1-1-1>) and of a random response scheme (<1-2-3>). Panel E depicts the distribution of the rank-order-quadruplets. Panel F shows the distribution of the response-modulation for simulated (abscissa) and real (ordinate) eye movements. For details see main text.

Figure 3. Crossmodal self-motion aftereffects. (A) Illustration of the experimental protocol. During the adaptation phase (left), subjects were presented with an optic flow simulating either forward or backward self-motion at constant velocity (3 m/s). In the test phase (right), subjects experienced a 2-sec passive linear fore-aft translation with a Gaussian velocity profile (vestibular-only test), and subsequently indicated the perceived movement direction. In a control condition, a visual-only test stimulus (far right) composed of expanding or contracting optic flow was used instead, and subjects indicated the direction of optic flow. Different adaption conditions were run in separate blocks of 50 trials each: baseline (no adaptation), forward (15 sec), backward (15 sec), as well as 3 forward blocks with shorter adapter durations of 7.5, 3.75, or 1.5 secs. (B) Mean aftereffect (PSE) across subjects (n=20) following forward and backward adaptation with 15 sec duration compared with the no-adapter baseline. (C) Mean aftereffect (expressed as forward minus baseline PSE) as a function of adapter duration (n = 17). Subjects additionally rated their vection on a scale of 1 to 7, with 1 representing perception of object motion only and 7 representing perception of self-motion only. Gray dots indicate mean subjective ratings. (D) Crossmodal and visual-only aftereffects (expressed as forward or

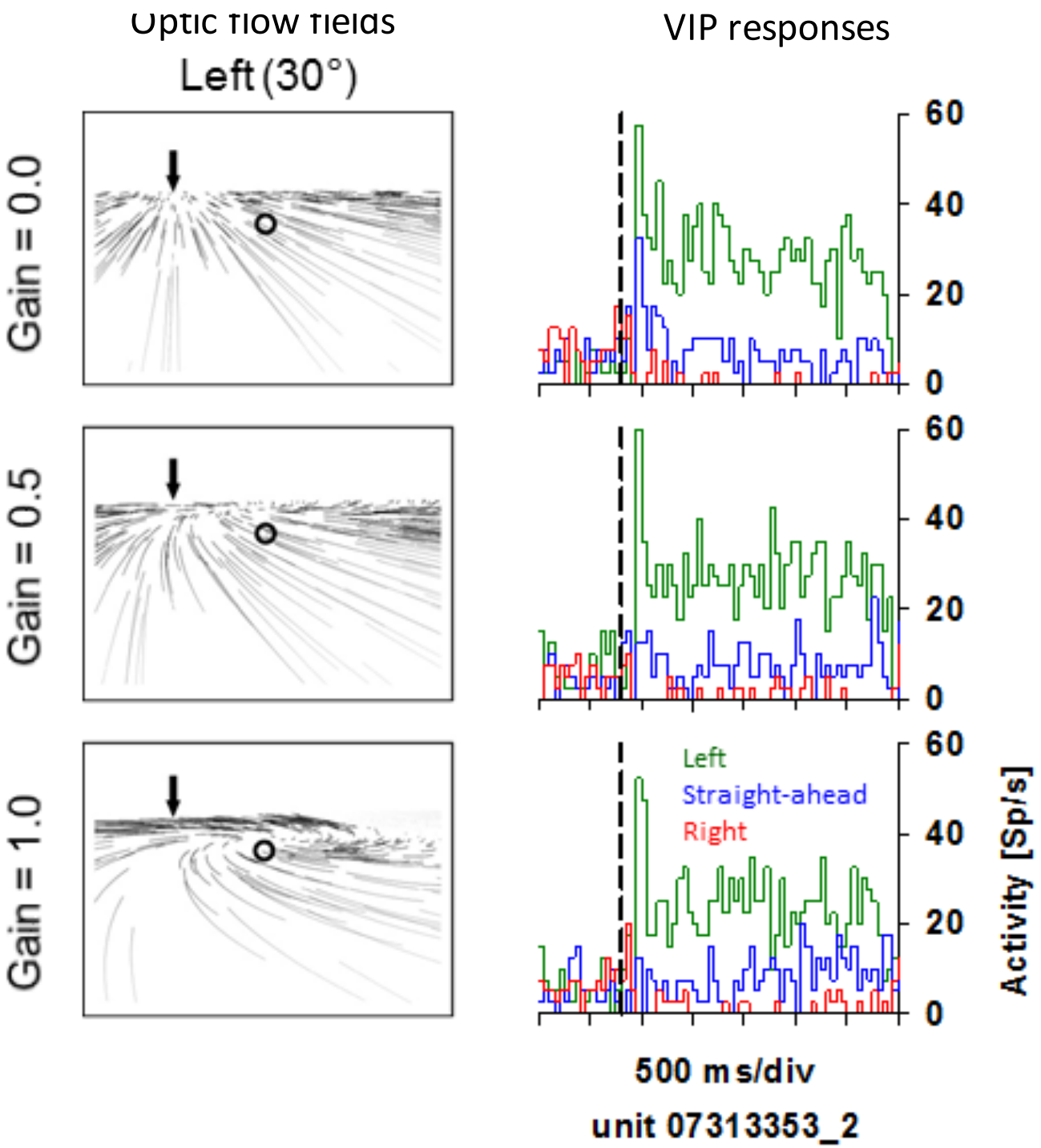
backward minus baseline PSE) are uncorrelated ($r = 0.003$; $p = 0.98$) ($n=15$). To allow for comparison, backward aftereffects are multiplied by -1. All error bars show SE.

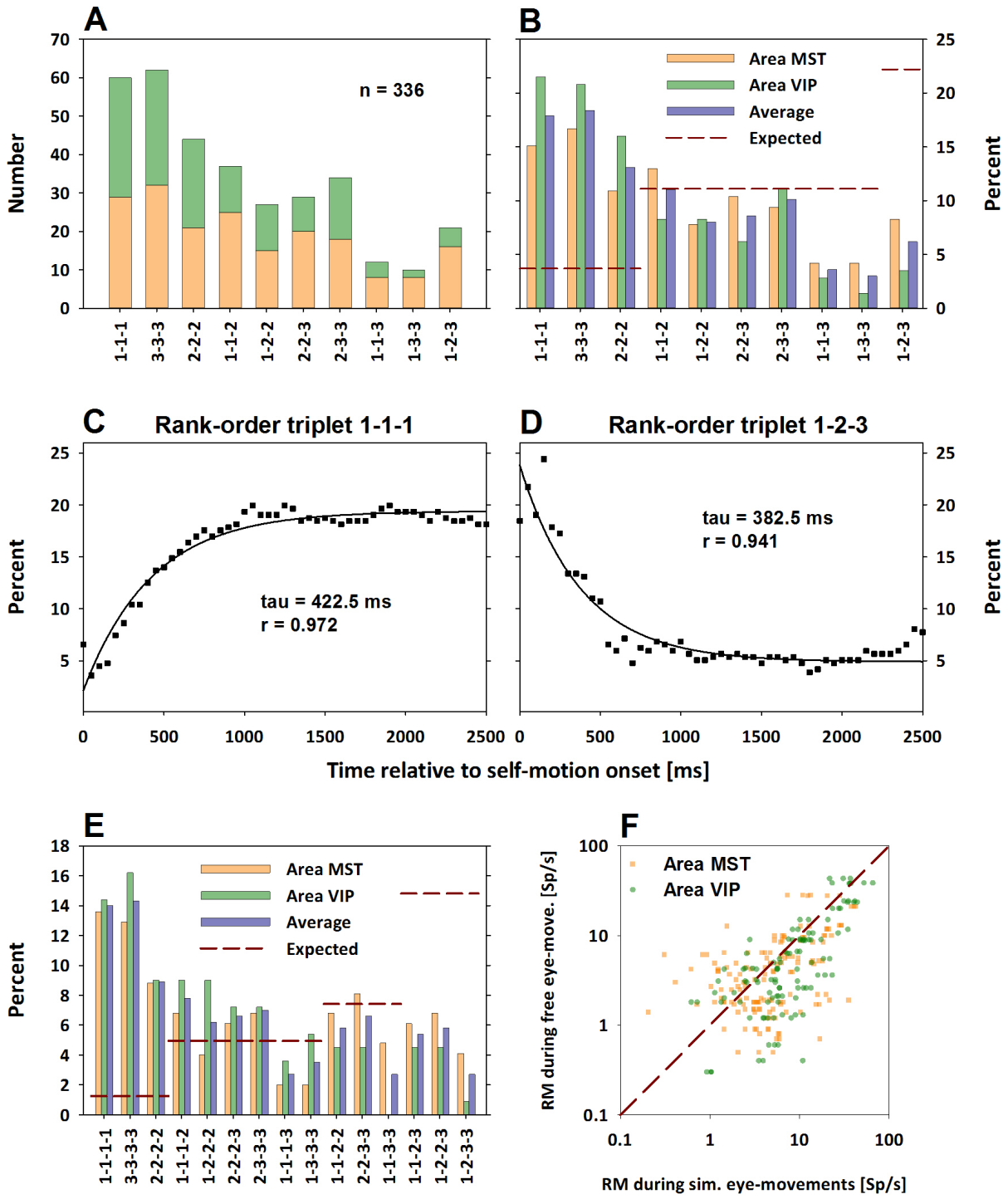
Figure 4. Experimental setup and results. Participants are seated in the motion platform with a screen in front of them. Upper panel: Experiment 1. Lower panel: Experiment 2. Participants show optimal visuo-vestibular integration, their bimodal threshold (red square) being lower than their unimodal visual (blue) and vestibular (black) thresholds, and not different from the predictions of a Bayesian optimal integration model (red circle).

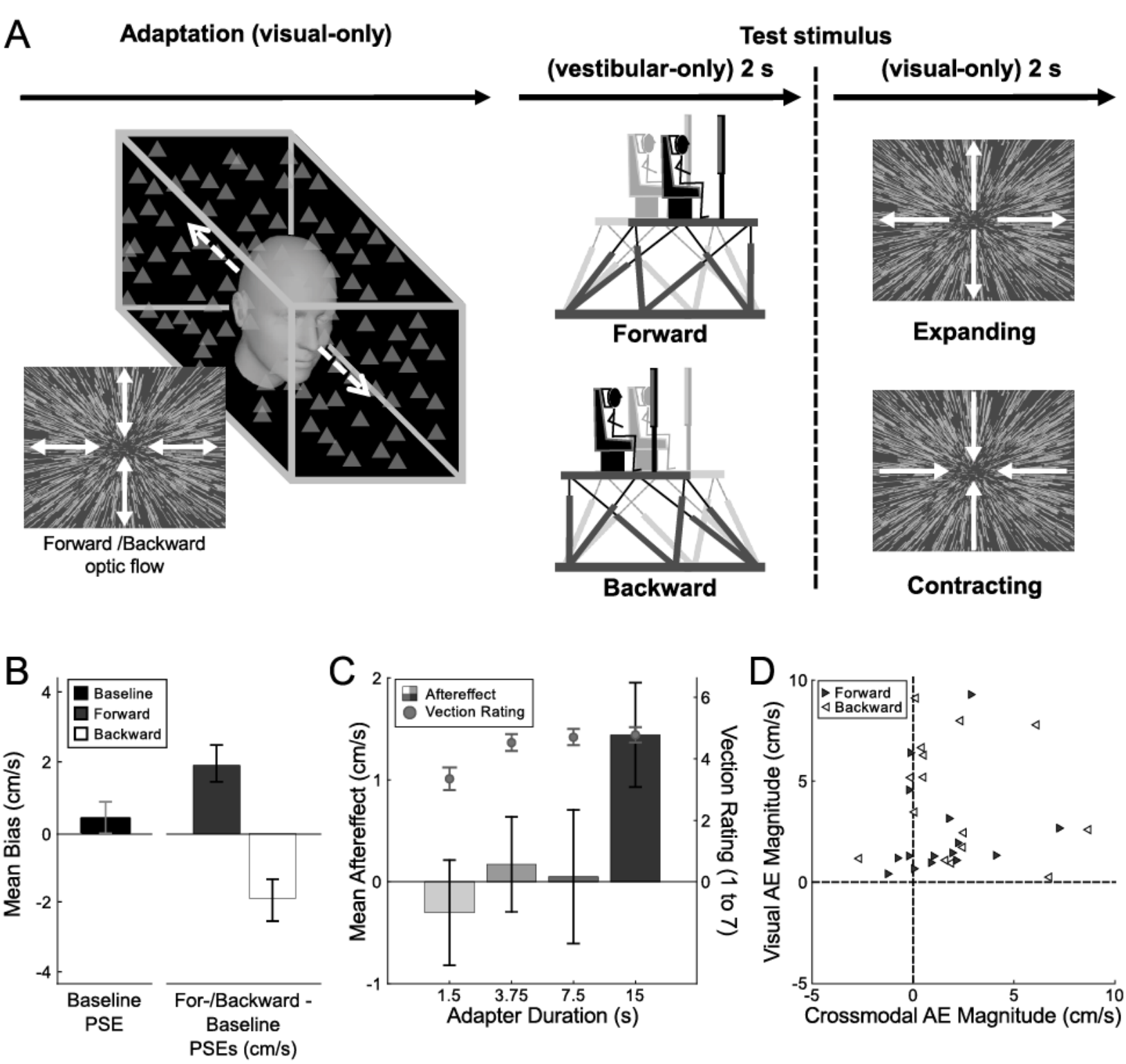
Figure 5. A: Experimental setup testing for the effects of natural vestibular stimulation on tactile detection. A congruent trial is depicted: rotation direction corresponds to the side of tactile stimulation. B: Results. Independently of congruency tactile sensitivity was improved during rotation in comparison to a no-rotation baseline.

Figure 6. (A) Cortical regions of interest studied. The figure shows an inflated representation of the right cerebral hemisphere of one individual with the locations and extents of hMT, hMST, hVIP, V6 and CSv indicated as coloured overlays. In each case, BOLD activity elicited by the visual localizer that was used to define the region is shown in a slice through the brain of the same participant. A corresponding set of visual areas is present in the left hemisphere (not shown). Modified from Smith et al (2012). (B) A 'cut-out' section of the flattened grey matter representation from each hemisphere of one participant, centred on the MT complex (dashed white line is the superior temporal sulcus). Vestibular activity (orange/yellow) is superimposed, together with the outlines of hMT (green) and hMST (magenta) as defined with a visual localizer. Vestibular activity is apparent in hMST but not hMT and is confined to the anterior portion of hMST. Modified from Smith et al (2012). (C) Visual stimulus used by Billington & Smith (2015). A circular patch of white dots appears to rotate during GVS. Physical rotation of the patch on the screen was used to null this illusory motion. (D) MVPA results for classifying the temporal phase of sinusoidal visual and vestibular rotations in the roll plane for five cortical visual areas. Chance performance is shown, along with the 95th percentile obtained from permutation testing as an indicator of statistical significance. Modified from Billington & Smith et al (2015).

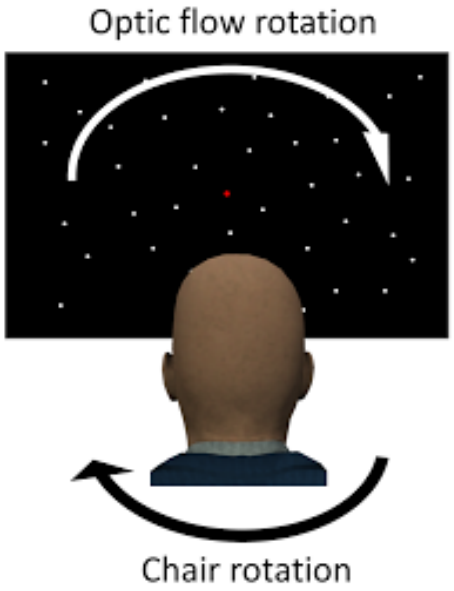
Figure 7. Organization of human vestibular cortex in lateral sulcus (also called sylvian fissure). Shown are activations in a sample participant (left hemisphere) during stimulation with visual motion and caloric vestibular cues ($p < 0.001$, uncorrected). (A) Visual motion stimulation shows significant activations in the posterior insular cortex area (PIC) in the posterior end of the lateral sulcus (LS). In addition to PIC, other motion-sensitive regions in visual and parietal cortex respond well to visual motion stimuli. (B) Caloric stimulation elicits activations in the vestibular network in lateral sulcus, including the putative center of cortical vestibular processing, the parieto-insular vestibular cortex area (PIVC). Activations during caloric stimulation are also evident in area PIC, suggesting that PIC is part of both, the vestibular and the visual motion processing networks After Frank et al., 2014 (with permission of the publisher) and Frank et al., in preparation.

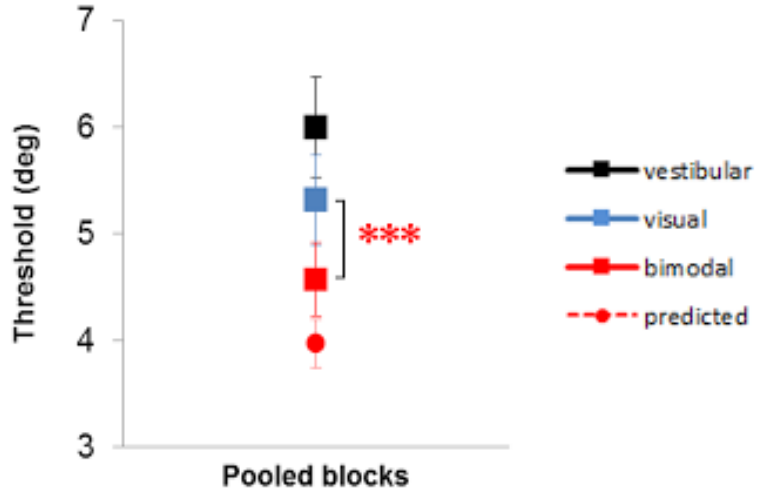

Optic flow fields
VIP responses
Left (30°)
Gain = 0.0
Gain = 0.5
Gain = 1.0
60
40
20
0
Left
Straight-ahead
Right
Activity [Sp/s]
500 ms/div
unit 07313353_2


A
B
C
D
E
F
Number
Percent
n = 336
Area MST
Area VIP
Average
Expected
1-1-1
3-3-3
2-2-2
1-1-2
1-2-2
2-2-3
2-3-3
1-1-3
1-3-3
1-2-3
Rank-order triplet 1-1-1
tau = 422.5 ms
r = 0.972
Rank-order triplet 1-2-3
tau = 382.5 ms
r = 0.941
Time relative to self-motion onset [ms]
1-1-1-1
3-3-3-3
2-2-2-2
1-1-1-2
1-2-2-2
2-2-2-3
2-3-3-3
1-1-1-3
1-3-3-3
1-1-2-2
2-2-3-3
1-1-3-3
1-1-2-3
1-2-2-3
1-2-3-3
RM during free eye-move. [Sp/s]
RM during sim. eye-movements [Sp/s]

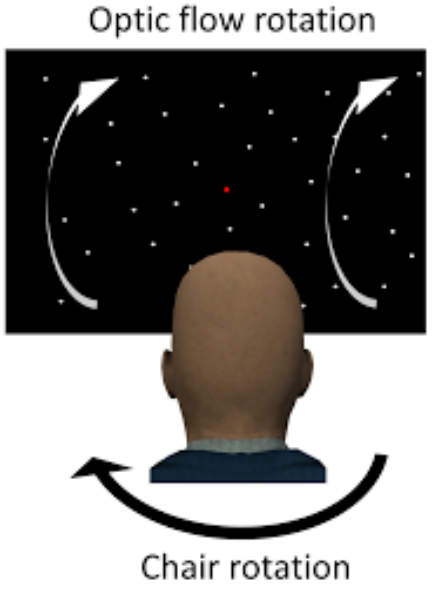

A
Adaptation (visual-only)
Test stimulus
(vestibular-only) 2 s
(visual-only) 2 s
Forward /Backward
optic flow
Forward
Backward
Expanding
Contracting
B
Baseline
Forward
Backward
Mean Bias (cm/s)
Baseline
PSE
For-/Backward -
Baseline
PSEs (cm/s)
C
Aftereffect
Vection Rating
Mean Aftereffect (cm/s)
Vection Rating (1 to 7)
Adapter Duration (s)
D
Forward
Backward
Visual AE Magnitude (cm/s)
Crossmodal AE Magnitude (cm/s)

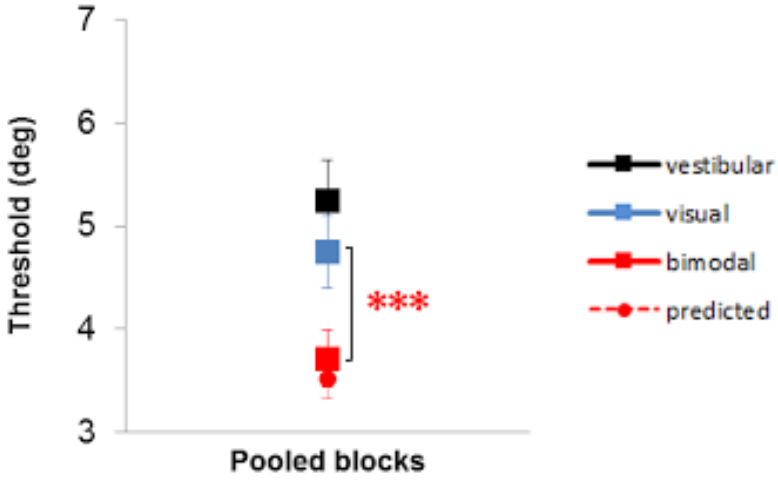
Experiment 1. Vestibular Yaw + Visual Roll



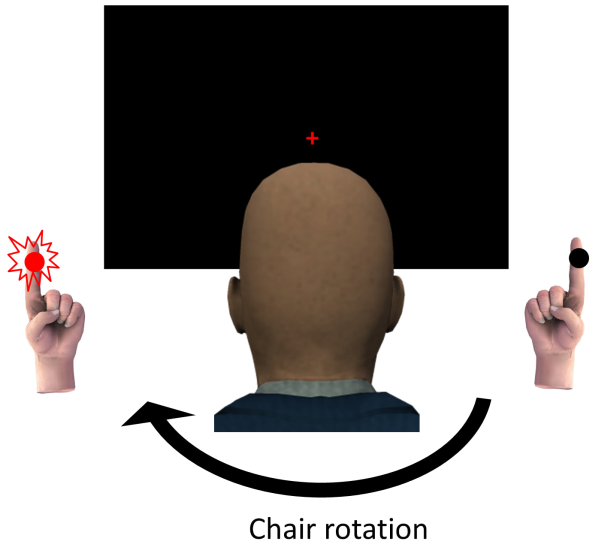




Experiment 2. Vestibular Yaw + Visual Pitch

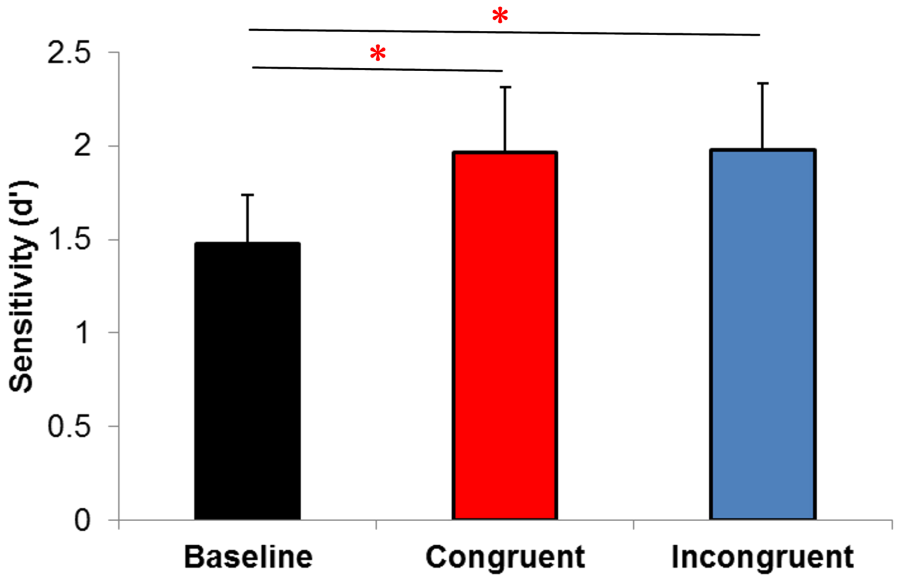


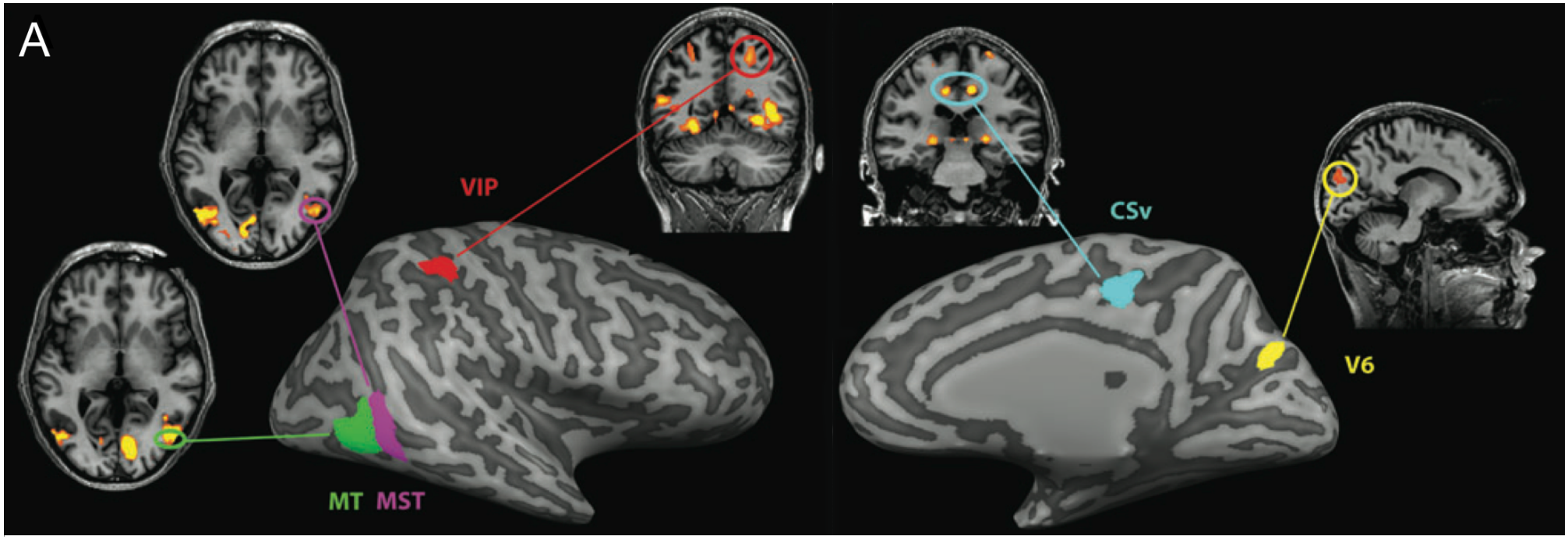



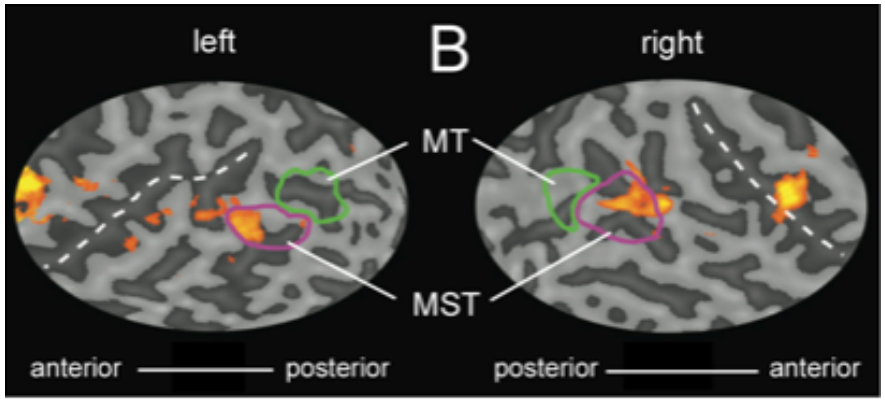
A. Experimental setup

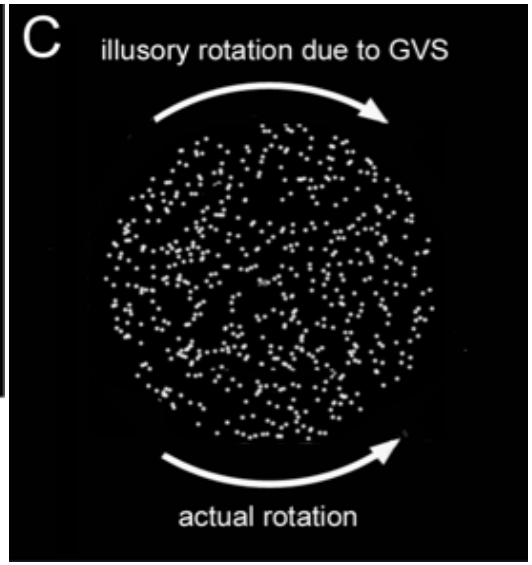


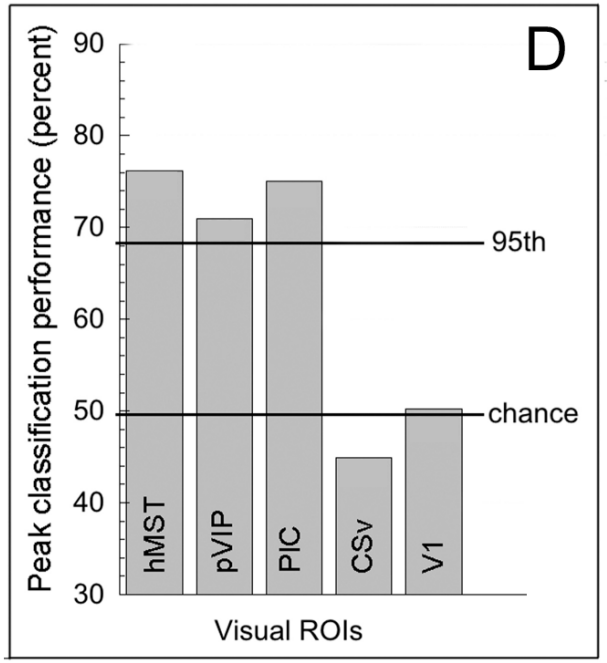

B. Results

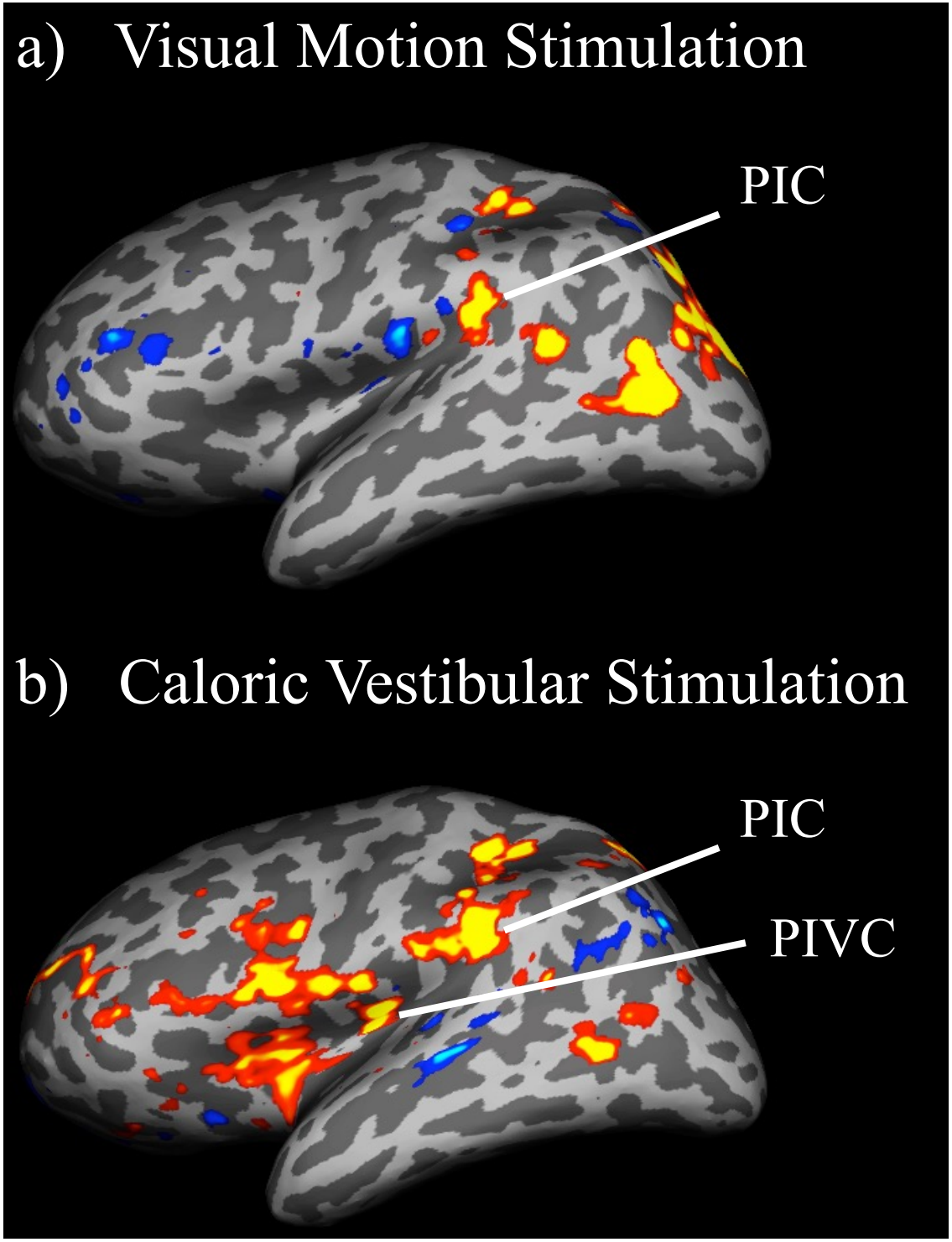



A
VIP
CSv
V6
MT MST


left
B
right
MT
MST
anterior
posterior
posterior
anterior


C
illusory rotation due to GVS
actual rotation


D
Peak classification performance (percent)
90
80
70
60
50
40
30
95th
chance
hMST
pVIP
PIC
CSv
V1
Visual ROIs


a) Visual Motion Stimulation
PIC
b) Caloric Vestibular Stimulation
PIC
PIVC