# *The role of the ventral intraparietal area (VIP/pVIP) in parsing optic flow into visual motion caused by self-motion and visual motion produced by object-motion*

Field, D. T. ORCID: https://orcid.org/0000-0003-4041-8404, Biagi, N. ORCID: https://orcid.org/0000-0002-7119-0767 and Inman, L. A. (2020) The role of the ventral intraparietal area (VIP/pVIP) in parsing optic flow into visual motion caused by self-motion and visual motion produced by object-motion. NeuroImage, 213. 116679. ISSN 1053-8119 doi: 10.1016/j.neuroimage.2020.116679 Available at https://centaur.reading.ac.uk/89479/

# The role of the ventral intraparietal area (VIP/pVIP) in the perception of object-motion and self-motion

David T. Field [*], Nicolò Biagi, Laura A. Inman

*Centre for Integrative Neuroscience and Neurodynamics (CINN), Psychology and Clinical Language Sciences, University of Reading, Reading, RG6 6AL, UK*

ABSTRACT

Retinal image motion is a composite signal that contains information about two behaviourally significant factors: self-motion and the movement of environmental objects. It is thought that the brain separates the two relevant signals, and although multiple brain regions have been identified that respond selectively to the composite optic flow signal, which brain region(s) perform the parsing process remains unknown. Here, we present original evidence that the putative human ventral intraparietal area (pVIP), a region known to receive optic flow signals as well as independent self-motion signals from other sensory modalities, plays a critical role in the parsing process and acts to isolate object-motion. We localised pVIP using its multisensory response profile, and then tested its relative responses to simulated object-motion and self-motion stimuli; results indicated that responses were much stronger in pVIP to stimuli that specified object-motion. We report two further observations that will be significant for the future direction of research in this area; firstly, activation in pVIP was suppressed by distant stationary objects compared to the absence of objects or closer objects. Secondly, we describe several other brain regions that share with pVIP selectivity for visual object-motion over visual self-motion as well as a multisensory response.

## 1. Introduction

Visual motion provides the brain with behaviourally useful information about self-motion and object-motion. Self-motion produces relative motion between the eye and the world that is reflected in the global structure of the patterns of light stimulating the retina. It is possible to recover the direction of self-motion from this global optic flow pattern (Gibson, 1950), while local motion inconsistent with the global flow pattern indicates object-motion; given this computational interdependence, it is likely that the neural processes that extract these two key signals are also interlinked at some level (Rushton et al., 2018).

During self-motion through a static environment, the current heading direction and path curvature are both reflected systematically within the optic flow (Warren et al., 1991), and at the neural level it has been shown that multiple macaque brain regions contain neurons that are tuned to the necessary global patterns in optic flow as well as the location of the focus of expansion (FOE) in the flow field, which specifies the heading direction if gaze is fixed over time. These include: MSTd, Duffy and Wurtz (1991); VIP, Schaafsma and Duysens (1996); area 7a, Siegel and Read (1997); Raffi et al. (2002); and FEF, Gu et al. (2016).

While identifying the global patterns and heading direction in the flow field is behaviourally useful in itself, doing this is also a necessary part of the process of identifying object-motion because object-motion is visually specified by local retinal motion components that are inconsistent with the global flow (Warren and Rushton, 2009). In macaques, one brain region likely to be important for 'parsing' (Rushton and Warren, 2005) total visual motion into optic flow components caused by self-motion and local motion components caused by object-motion is the ventral intraparietal area (VIP). This is because many VIP neurons are tuned to extract information from optic flow generated by self-motion, but neurons are also present that respond to visual stimulation indicative of object-motion. In particular, many VIP neurons are disparity tuned; one study found that this disparity tuning was unlikely to be directly involved in multisensory integration for heading perception and suggested it could either be used to dissociate object-motion from self-motion or to provide a second self-motion signal independent from optic flow (Yang et al., 2011). Another study found that the majority of disparity-tuned VIP cells preferred crossed disparities produced by objects at near distances closer than the depth of fixation (Bremmer et al., 2013). Furthermore, a recent fMRI study of macaque brain activation in response to real objects placed in near extrapersonal space or far space found that VIP was part of the network preferentially activated for objects in near extrapersonal space (Cléry et al., 2018).

---

That self-motion signals are present in VIP is indicated by the presence of neurons tuned to the location of the FOE in a flow field (Bremmer et al., 2002a). Furthermore, VIP neurons have visual responses to flow fields that are invariant or partially invariant to eye position – thus their responses are indicative of the 'real' heading direction with respect to the head, rather than with respect to the fovea of the retina as is the case in most visual areas (Zhang et al., 2004). The observations of real heading direction signals in VIP has been generalised to the case where the eye is moving during smooth pursuit eye movements (Kaminiarz et al., 2014; Bremmer et al., 2017); although the latter study showed that invariance breaks down during saccades. Invariance of responses to eye position is also desirable from a computational and behavioural perspective for the representation of object-motion.

VIP neurons also respond to visual stimulation indicative of object-motion, and some have binocular tuning suited to signalling whether or not an object is closer to the head than the current fixation plane (Bremmer et al., 2002b). This type of response was evident in the early investigation of VIP conducted by Colby et al. (1993), in which single spots of light were displayed on a tangent screen whose distance from the monkey was varied. The authors noted that some VIP neurons "appear to be involved in the detection of the trajectory of stimulus motion and anticipation of point of contact for an approaching visual stimulus.". Later work showed that VIP is part of a visuotactile convergence network whose activation is selectively enhanced when a visual stimulus looming towards the face correctly predicts the location on the face and time of occurrence of a consequent impact on the face (Cléry et al., 2017). Corresponding to these neural effects, human work showed enhancement of tactile psychophysical thresholds in a spatial and temporal window defined by predictive information contained in visual stimuli looming towards the face (Cléry et al., 2015).

Despite observations of object-motion selectivity, most of the VIP literature has focused on the responses of VIP neurons to optic flow and the presence of heading direction signals, including the non-visual heading signals available from the vestibular system and other sensory modalities (e.g., Zhang and Britten, 2004; Chen et al., 2013a,b). Nonetheless, we believe it is worth asking whether the primary functional role of VIP (and putative pVIP in humans) is to represent heading direction and other aspects of self-motion, or object-motion, or both? Deactivation studies of MST and VIP suggest an answer. Both MST and VIP receive their visual input from MT, and electrical stimulation of columns of cells in both regions influences heading judgments (Britten and Richard J. A., 1998; Zhang and Britten, 2011). However, reversible inactivation studies produce strikingly different results in the two regions. In MSTd, inactivation severely disrupts heading judgments based on optic flow while having a weak effect on vestibular based heading judgments, whereas in a more recent study VIP deactivation had no effect on either type of heading judgment (Gu et al., 2012; Chen et al., 2016). The authors acknowledge that their VIP deactivation findings are unprecedented in that they are the first demonstration of a brain area whose responses show a strong correlation with perception, but appear to have no causal relationship to perception. They are forced to conclude that MSTd has a causal role in heading perception, while VIP does not. What is the purpose of the information about heading direction that is present in VIP if it has no influence on the perception of heading? Our suggestion is that VIP uses this signal to isolate, and presumably support perception of, the motions of objects by discounting from the flow field all visual motion that is consistent with current self-motion.

Turning to the putative human homologue of VIP (pVIP), this was first identified by exploiting the multimodal property of the region in an fMRI conjunction analysis: VIP/pVIP responds to auditory and tactile motion in addition to visual motion (Bremmer et al., 2001). However, as pointed out by Bartels et al. (2007) the type of conjunction analysis used by that study, implemented in SPM99, has subsequently been criticized for allowing voxels to survive thresholding if any one of the input contrasts is strongly activated (Nichols et al., 2005). While the authors intention was that voxels only survive conjunction thresholding if they are activated in all three sensory modalities, inspection of Fig. 1 of Bremmer et al. (2001) suggests that the visual contrast may have driven their conjunction result. The present study will resolve doubts about the multimodal conjunction method of localising pVIP by using a conceptual replication of Bremmer et al (2001) experimental stimuli combined with a valid conjunction analysis that requires all three sensory modalities to activate significantly.

It has subsequently been confirmed that visual responses in pVIP are head-centred like those of macaque VIP, and the same study showed that its somatosensory receptive fields map the face and are co-aligned with the visual map (Sereno and Huang, 2006). The first fMRI study claiming to identify a specific role for pVIP in processing optic flow to detect heading direction did so by confirming a preference in pVIP for coherent flow patterns – i.e. retinal motion inconsistent with self-motion (Wall and Smith, 2008). A follow-up study demonstrated that BOLD signals recorded from pVIP can decode the direction of heading changes in a flow field (Furlan et al., 2014). While these studies establish that pVIP is responsive to self-motion signals in optic flow, they did not compare these responses to the response to object-motion. Thus, the question of whether pVIP is specialised for heading detection or instead uses information about self-motion as part of the process of identifying the motion of objects was not answered by these studies. Furthermore, these studies did not localise pVIP using the multisensory criteria that Sereno and Huang (2006) argued should be used in order to guarantee localising the human area homologous to macaque VIP, and thus it is possible that their findings apply to a nearby functional region.

Turning to studies of pVIP that have included objects as stimuli, one study localised pVIP using a ball that approached and receded from the face compared to stationary presentation of the ball (Quinlan and Culham, 2007). In a third condition the ball approached and touched the face, and activation in this condition was not distinguishable from the approach/recede condition. This method of localising pVIP produced somewhat different peak coordinates from the multisensory conjunction method used by Bremmer et al. (2001). Quinlan & Culham's results indicate that pVIP, or possibly a nearby region, is responsive to the visual expansion and looming and/or other depth cues produced when an object comes close to the face.

Calabro and Vaina (2012) conducted an exploratory whole brain fMRI investigation of optic flow and object-motion processing. Participants experienced either simulated forwards or backwards motion in the presence of an array of nine stationary objects, or forwards motion in a similar scene in which one of the objects also moved. These two experimental conditions were not contrasted statistically with each other; instead they were separately contrasted with a no visual motion baseline. This analysis served to identify sets of regions of interest that were then used to examine the network connectivity underlying object-motion processing. Four sub-networks were identified, and pVIP was part of one of those together with DIPSM and the right precuneus. Connectivity within this particular sub-network was reduced when all visual motion was consistent with self-motion, which is consistent with our proposal of a specific role for pVIP in perceiving object-motion. However, this study does not answer the question of whether pVIP specifically is specialised for object-motion processing because connectivity in the self-motion only condition was also weakened within two other sub-networks. Furthermore, the authors acknowledge that behavioural task difficulty was lower in the self-motion only condition, and so the reported differences in network connectivity are also open to explanation in terms of task difficulty, effort, or arousal.

Overall, studies have shown that both VIP and pVIP respond to simulated self-motion and also object-motion, but no properly controlled direct comparison of the two classes of event has been carried out in order to establish which produces the stronger response. In the present study, to ensure we focused on the area homologous to monkey VIP we localised pVIP using the multisensory method of Bremmer et al. (2001), as recommended by Huang et al. (2017). We then tested its relative responses to simulated object-motion and self-motion stimuli, matching
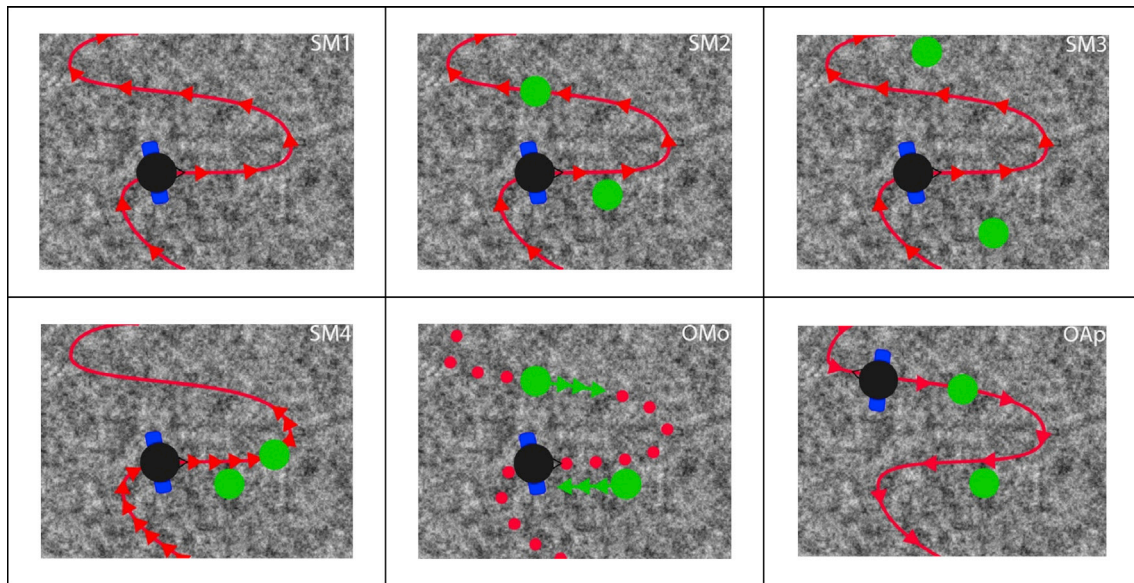
**Fig. 1.** Experimental conditions. Birdseye views of simulated self-motion and object-motion unfolding over the course of 16 s fMRI blocks. Note that these illustrative diagrams are not drawn to scale and depict only one third of a block. In SM1 the curved line with arrows depicts the simulated course travelled over a textured ground plane, which generated optic flow at the moving point of observation. The line was not visible to the participant, whose task was to continuously adjust the angle of a joystick fixed in position near their right hip to reflect their perception of the current rate of change of heading. In SM2 static pole objects were positioned on or close to the course travelled; two are shown here, but at any moment all but the closest pole was hidden from the point of observation. In SM3 the poles were positioned at a greater lateral separation from the course travelled. In SM4 the viewpoint traversed the course at a reduced speed; this is indicated by the reduced distance between the arrows. In SM4 poles were positioned closer together so that the number of poles presented per 16 s block remained the same as in SM2 and SM3. OMo was created by modifying SM2 such that the ground plane texture object was attached to the moving viewpoint, removing all optic flow and creating the impression that the viewpoint was stationary with objects moved towards it. During OMo, joystick movements tracked the lateral motion of the object. OAp was a modification of SM2 in which the viewpoint was repositioned to the end of the course and then moved backwards along the course to its start, equivalent to a passenger in a car looking out of the rear window; objects became visible once the viewpoint had 'passed through' them.

low-level stimulus properties and behavioural task difficulty as closely as possible. Our results show a much stronger response in pVIP for object-motion and support the hypothesis that optic flow and other self-motion signals are present in pVIP not to provide perception of those signals, but to enable pVIP to isolate the motion of objects.

## 2. Materials and methods

### 2.1. Participants

Nineteen participants were recruited (8 male, age range 20–50, mean age 28.3 years). All had normal or corrected-to-normal vision. Written informed consent was collected prior to taking part. The study was approved by the University of Reading Research Ethics Committee. Due to a technical problem with presentation of auditory stimuli in the MRI scanner, the sample size was reduced to 15 for analysis of the auditory motion part of the multi-sensory pVIP localiser.

### 2.2. Functional localisers

To replicate Bremmer et al. (2001) we attempted to localise pVIP on the basis of its sensitivity to motion in multiple sensory modalities. Our stimuli were similar in conception, but not identical in terms of low-level details to those used by Bremmer et al. (2001). Visual motion, auditory motion, and tactile stimulation by air flowing across the face were presented in separate scans; pVIP was expected to respond to all three. Each scan used an AB block design with 16 s blocks and 8 repetitions. In practice, we found that responses to the auditory stimulus were not reliable enough to use for localization at the individual participant level, and so individual participant regions of interest (ROI) were made up of voxels that responded to both the visual and tactile stimulus. However, it was possible to include the auditory scan in a 3-way conjunction analysis that successfully located pVIP at the group level.

The visual motion localiser was presented using an MRI-compatible BOLDscreen (Cambridge Research Systems), displaying 1920 * 1200 pixels with a field of view of 23 * 14 degrees of visual angle. The screen was positioned at the bore of the magnet and was viewed by participants via a mirror fixed to the head coil and positioned above the eyes. Visual motion consisted of a cloud of approximately 100 dots forming an optic flow pattern that was a combination of translation and radial motion. The optic flow pattern was generated by simulating self-motion on a winding course produced by summing sine waves of differing frequencies and amplitude. The viewpoint was continuously rotated to face the instantaneous heading direction. Participants fixated a central cross while passively viewing the stimulus. During the baseline blocks static dots replaced optic flow. With the exception of the use of the BOLDscreen visual display, the stimulus was identical to that used in the investigation of CIngulate Sulcus Visual area by Field et al., 2015, where full technical details are provided.

The sensation of tactile motion was produced by propelling room air across the participants face from a tube fixed to the head coil. The tube passed through a wave guide into the MRI control room where it was attached to a 'heat gun' from which the heating element had been removed. This apparatus allowed for two different rates of air flow to be produced; we changed the rate every 4 s during stimulation blocks to minimise sensory adaptation. The baseline blocks consisted of rest with no air flow.

The sensation of auditory motion was produced by the stereo presentation of a 'whooshing' sound, which was perceived as moving from one side of the head to the other. Individual sounds lasted approximately 2 s, and after two sounds were presented moving from left to right the motion direction was reversed. This cycle continued until the end of the 16 s block. The baseline was rest without auditory stimulation. The sounds were presented to participants via NordicNeuroLab MRI compatible stereo headphones.

## 2.3. Experimental stimuli and behavioural task

The main experiment consisted of an ABCDEFG block design with 8 repetitions of 6 experimental conditions plus rest. For schematic diagrams of the experimental conditions see Fig. 1. Individual blocks were 16 s in duration and were separated by 1 s information screens that indicated to the participant which of the pre-trained experimental conditions was about to occur; this allowed participants to know which of two different versions of the joystick task they were about to perform. Perspective correct 2D visual stimuli were generated using a virtual reality environment programmed in Vizard 3.0, and were presented using the same visual display device as the visual motion pVIP localiser scan.

The first experimental condition, self-motion 1 (SM1) simulated moving on a winding trajectory across a textured but otherwise empty ground plane, with the viewpoint rotating so that it always faced the instantaneous heading direction. Trajectories were generated by summing two sine waves of different frequencies and amplitudes. The participant continuously adjusted the angle of an MRI compatible joystick fixed in position near their right hip to reflect their perception of the current rate of change of heading. Participants were trained on this task before scanning; as well as the experimental stimuli, training included 'active steering' trials in which joystick lateral position controlled the rate of change of heading – these trials allowed participants to establish a mapping between joystick position and the visual turning rate in the display. By design, the ground plane was a rich source of optic flow; it was made of a texture that contained luminance contrast over a wide range of spatial frequencies. The simulated eye height was 1.1 m above the ground and the simulated travel speed was 16.9 m/s. The course travelled was mirror imaged in half of the blocks to prevent participants initial joystick responses becoming habitual. We have previously used this task with similar stimuli to investigate other aspects of the perception of self-motion and steering (Field et al., 2007; Billington et al., 2010).

To test whether the presence of objects during simulated self-motion influences pVIP responses, self-motion 2 (SM2) was identical to SM1 but with the addition of pole objects that were fixed on the ground plane at locations on or near to the course travelled. 250 msec after the start of the block a pole faded in over a period of 1.0 s at a simulated distance from the viewpoint of 56 m. While the viewpoint 'slalomed' towards the pole it remained fully visible for 2 s and then faded out over 0.5 s. After a 250 msec interval the next pole appeared on the horizon; during one block six cycles of pole appearance, approach, and disappearance occurred. Half the poles were positioned such that the viewpoint would pass through them if this had not been prevented by their fading out, and half were positioned such that the viewpoint would pass 2 m to the left or right of them. The retinal image sizes of the objects on appearance was 1.21 deg high * 0.54 deg wide. Immediately prior to fadeout the retinal size of poles had grown to 3.67 deg high and 1.63 deg wide. The height and width of a pole averaged across the time it remained on the screen was approximately 2.0 deg * 0.9 deg.

To test whether the relative proximity of objects fixed to the ground plane during simulated self-motion modulated pVIP activation levels, self-motion 3 (SM3) was included for comparison to SM2. SM3 was identical to SM2 except that the lateral separation between the moving viewpoint and the poles fixed to the ground plane was increased. For each SM3 pole, the viewpoint passed somewhere between 3.45 m and 5.80 m to the left or right of the poles' ground plane location, while in SM2 it was either on a collision course or 2 m to the left/right. Although the poles were visually more lateral than in SM2 as they neared their fade out point, their optical size was very similar across their lifetimes. On average the minimum size of a pole, just after it faded in, was 1.21 deg high * 0.54 deg wide. Immediately prior to fadeout the retinal size of poles had grown to 3.69 deg high and 1.65 deg wide. The height and width of a pole averaged across the time it remained on the screen was 1.21 deg * 0.54 deg.

To test whether the relative speed of simulated self-motion influenced pVIP activation levels, self-motion 4 (SM4) was included for comparison to

SM2. The main difference between SM4 and SM2 was that the simulated speed of self-motion was reduced from 16.9 m/s to 7 m/s. To keep the number of poles presented in a block and their individual screen durations the same as for SM2, the initial fade-in distance of poles was reduced. The slower self-motion also resulted in a smaller portion of the full course being traversed, and the pole locations being physically closer together in the simulated world. Consequently, the retinal image size profile over time of the poles were very similar to those in SM2, and the main difference the participant might notice was the slower rate of optic flow being generated by travelling over the ground plane texture. On average the minimum size of a pole, just after it faded in, was 0.86 deg high * 0.38 deg wide. Immediately prior to fadeout the retinal size of poles had grown to 3.32 deg high and 1.46 deg wide. The height and width of a pole averaged across the time it remained on the screen was 1.53 deg * 0.69 deg.

To compare pVIP activation levels in response to simulated object-motion with those to self-motion we included the Object-motion (OMo) experimental condition in which the viewpoint remained static and an object moved across the ground plane towards it. The object 'slalomed' towards the viewpoint in such a way that the relative motion between it and the viewpoint was identical to that in SM2. To implement this in the rendering software we repeated SM2 but with the ground plane object attached to the moving viewpoint so that it remained entirely fixed and stationary on the screen. This method guaranteed that the low-level visual properties of the pole object were identical in OMo and SM2; the only visual difference between OMo and SM2 was that the ground plane texture remained static in OMo, while in SM2 it was the source of optic flow indicating self-motion. During OMo, the participants' joystick movements tracked the lateral motion of the object rather than the rate of change of heading. This produced time courses of joystick position values that were the mirror image of those in SM2.

Given previous reports that VIP/pVIP might be particularly responsive to the presence of objects in the space close to the face we included an Object Appearance (OAp) experimental condition in which the fade-in location of the pole objects was near the point of observation, producing a percept of looming. To maintain the match between OAp and the other conditions in terms of durations for which objects remained on the screen and also low-level visual properties, OAp was a modification of SM2 in which the viewpoint travelled backwards over the same course. This created visual conditions such as those experienced by a passenger in a car moving forwards looks out of the rear window. In OAp the fade in of objects was rapid (0.5 s) in order to produce the looming/appearance percept. The fade-out in OAp occurred when the pole was in the distance and took 1 s. Optic flow was the same in OAp and SM2, except that the expansion component in SM2 was replaced by contraction. Note that in OAp the objects were stationary and so there was no actual motion of an object across the ground plane of the sort present in OMo. However, object-motion in the sense of surprising visual motion that the observer can't attribute to their own self-motion was perceived momentarily each time an object faded rapidly into view close to the simulated point of observation. Shortly after the fade-in the percept changed to one in which the visual motion of the object was predictable on the basis of self-motion, at which point the impression of object-motion ended. As in SM1-4 the joystick task in OAp was to track the rate of change of heading using the optic flow in the display.

Of the seven alternating experimental conditions in the block design, three included expanding optic flow, one presented contracting flow (OAp), one a slower rate of expanding flow (SM4), one no flow (OMo), and rest also had no optic flow. If the BOLD response to the optic flow were to adapt and be selective for direction and speed of flow then this could introduce differential rebound from adaptation effects across the experiment. However, our previous experience piloting this type of experimental stimulus indicated that rapid adaptation occurs for simulated travel in a straight line, or if the rate of path curvature is constant, but when the rate of path curvature is constantly changing and alternates between curving to the left and right as in the stimuli used here the BOLD signal does not decline as a function of block length.

## 2.4. MRI data acquisition

Imaging data were acquired at the University of Reading (Centre for Integrative Neuroscience and Neurodynamics) using a Siemens Allegra 3T scanner. Functional scans with 3 mm isotropic voxels consisted of repeated single-shot echo-planar imaging (EPI): echo time (TE) = 30 msec, flip angle (FA) = 90°, matrix size = 64 x 64, field of view (FOV) = $192 \times 192 \, \text{mm}^2$, with slice order descending and interleaved, 50 slices (0 inter-slice gap), slice thickness = 3.0 mm, and repetition time (TR) = 2670 msec. A detailed T-1 weighted MPRAGE anatomical image (resolution1 $\text{mm}^3$) was acquired (TR = 2020 msec, TE = 2.52 msec, FA = 9, FOV = $250 \times 250 \, \text{mm}^2$, 176 slices, no gap, total scan time = 4 min and 34 s). A gradient field map was acquired and used to unwarp the EPI scans (50 slices, voxel size = 3 $\text{mm}^3$, TR = 529 msec, TE1 = 4.92 msec, TE2 = 7.38 msec).

## 2.5. MRI data analysis

fMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The following pre-statistics processing was applied: distortion correction using BO unwarping, motion correction using MCFLIRT (Jenkinson et al., 2002); slice-timing correction using Fourier-space time-series phase-shifting; non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian kernel of FWHM 6 mm; grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; high pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 25.0 s for the CMA localiser scans and with sigma = 62.5 s for the experimental scan). Registration to high resolution structural and standard space MNI template images was carried out using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002). Registration from high resolution structural to standard space was then further refined using FNIRT nonlinear registration (Andersson et al., 2007).

The BOLD response was modelled using the GLM and a design matrix of explanatory variables (EVs) derived from the time course of the experimental stimuli, convolved with the standard FEAT double gamma HRF function. The 1 s information screens were modelled by a separate explanatory variable in the design matrix. Temporal derivatives of the EVs were also included in the design matrix. EVs were high pass filtered in the same way as the data. Time-series statistical analysis of individual participant data was carried out using FILM with local autocorrelation correction (Woolrich et al., 2001). Mixed effects statistical analysis to determine group average activations of first level contrasts was conducted with FSL using a 2nd level design matrix and FLAME 1 and 2 (Woolrich et al., 2004). Where we report whole brain conjunction analysis at the group level this was implemented by creating a 3rd level design matrix, to which we applied a fixed effects analysis and contrast masking in which the same thresholds were applied to each of the first level contrasts included in the conjunction. Details of the thresholds applied to Z (Gaussianized T/F) statistic images in each whole brain group analysis are given in the Results section.

## 2.6. Definition of pVIP regions of interest

For each participant, separate first level models were run for the visual, tactile, and auditory scans that made up the localiser. To implement the conjunction across the three scans a 2nd level fixed effects model bringing together the three 'stimulation – rest' COPE images from the first level was created for each participant. Thresholded activation for one of the scans was then contrast masked – at the same threshold – by the other two, which isolated voxels active in all three first level contrasts. Following this, all active voxels in the vicinity of the intraparietal sulcus were included in the ROI masks. In practice, despite the auditory stimulus producing detectable activation in pVIP in a group analysis, we found that several individual participants had no detectable activation

for the auditory stimulus in the region of pVIP, and as noted above there were also three participants where a technical problem had prevented the auditory stimulus being presented. Therefore, we used a 2-way conjunction of tactile and visual stimulation to localise pVIP ROI's in individuals. It has frequently been noted that individual participants vary considerably in the statistical contrast to noise ratio they exhibit, especially in passive stimulation as used here. Therefore, following Genovese et al. (2002) in each individual the conjunction was first run with a liberal threshold of p < 0.1 uncorrected for multiple comparisons applied to both the visual and tactile components, and then the voxelwise threshold was raised until appreciable random structure was no longer evident in the activation images. The thresholds selected for each participant using this method are reported in the Results. Note that for a voxel to be falsely declared active in the conjunction analysis, and so incorrectly included in the ROI due to random fluctuations of the fMRI signal, such random events would have to occur in both the visual and tactile scans: the probability of this is the product of the threshold applied to the visual contrast and the threshold applied to the tactile contrast. Once pVIP regions of interest were defined, percent BOLD signal change for the six experimental conditions was extracted using FEATQUERY.

## 2.7. Notes on the analyses of variance applied to percent BOLD signal change in pVIP ROI's and joystick task data

So that we could run ROI based ANOVA's that included hemisphere as a factor we replaced missing BOLD signal change data from those brain hemispheres in which we could not identify the pVIP ROI with the mean signal change in the hemisphere concerned, e.g. if a left hemisphere ROI had not been identified for a given participant then the signal change value was replaced by the mean of the identified left hemisphere ROIs. However, our findings do not depend upon this policy: we repeated our ANOVA analyses using listwise deletion of missing data, which reduced the number of participants included to 14, and this made no difference to any of the results reported here.

For ANOVAs conducted on both the data from the joystick tracking task and the percent BOLD signal change from pVIP ROI's, where Mauchly's test of sphericity was significant (p < 0.05) we report the Greenhouse Geisser corrected F tests.

## 3. Results

### 3.1. Conceptual replication of Bremmer's multisensory pVIP localiser: group analysis

The results of the three components of the multisensory localiser are shown in Fig. 2; the activation maps were thresholded using an initial voxelwise cut off of Z = 2.3, followed by a cluster threshold of 0.05 (corrected for multiple comparisons). Visual activation is shown in red, tactile in green, and auditory in blue. The three colours are mixed to indicate areas responding to two modalities. Brain areas surviving the three-way conjunction of visual motion, tactile motion, and auditory motion are presented in Fig. 3; only a very small activation cluster was found in the vicinity of the expected location of pVIP (5 voxels in the right ventral intraparietal sulcus centred on coordinates x = 28, y = −47, z = 46). This was due to the weak activation in this region produced by the auditory scan; lowering the voxelwise statistical threshold applied to the individual components of the conjunction to 0.05 uncorrected for multiple comparisons produced bilateral activation close to where it was found by Bremmer et al. (2001). Table 1 provides a comparison of pVIP peak coordinates from the three-way multisensory conjunction analysis performed at the more liberal threshold to those from previously published studies. To facilitate comparison of our coordinates with those of other studies we calculated the Euclidean distances in stereotaxic space between the different peak locations. The resulting distance matrix is presented in Table 2. It shows that the activation peaks from our three-way conjunction were closer to those of Bremmer et al. (2001) than
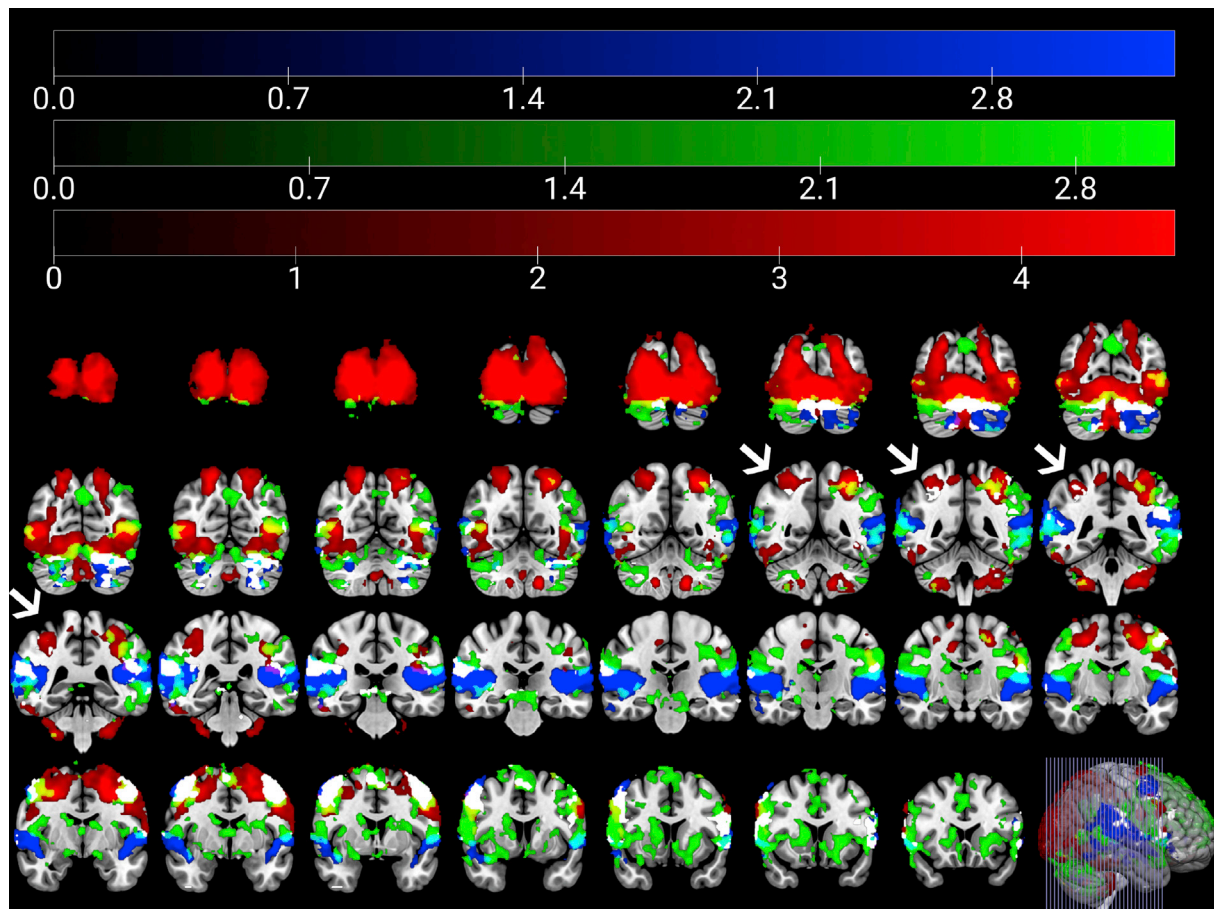
**Fig. 2.** Group level activity produced by each of the modalities of the localiser; visual = red, tactile = green, auditory = blue. Arrows highlight the ventral intraparietal sulcus; pVIP was expected to be found bilaterally in the fundus of this sulcus. Areas activated by visual and tactile = yellow, visual and auditory = purple, tactile and auditory = cyan. Areas activated by all three localisers are shown in a lighter colour, but are highlighted more clearly in Fig. 3. Slices cover the region of the brain from the occipital pole (y = −102) up to y = 18 in 4 mm steps.

to those of studies that localised VIP using different methods; this suggests that our replication of Bremmer's study was successful, although this finding should be interpreted in the context of the weaker auditory activation in pVIP and the stronger multisensory activation we found elsewhere in the brain.

### 3.2. Functional localization of pVIP region of interest in individual participants

Because the auditory stimulus failed to activate pVIP in some participants, and was missing due to a technical failure for 3 others, we defined individual participant regions of interest using the two-way conjunction of 'visual motion – rest' with 'tactile motion – rest'. In this way, we identified pVIP ROI in 17/19 left hemispheres and 15/19 right hemispheres.

The locations of and extent to which the individual participant ROI overlap anatomically are presented in Fig. 4. The volumes of the ROI, and the voxelwise Z thresholds used to define them are given on a per participant basis in Table 3.

### 3.3. Comparing responses to simulated self-motion with responses to object-motion in the pVIP ROI

Fig. 5 compares the BOLD response of the pVIP ROI's in the four experimental conditions in which all visual motion was caused by simulated self-motion with the two in which object-motion is salient. It is apparent in the figure that the BOLD response is greater for the two

object-motion conditions. We tested this statistically with a 2 (brain hemisphere) by 6 (experimental condition) repeated ANOVA, followed by a linear contrast to compare the two object-motion conditions with the self-motion conditions. There was a highly significant main effect of experimental condition (F(5, 32.8) = 18.11, $p < 0.001$), but no main effect of hemisphere (F(1,18) = 0.05 $p = .82$), or interaction between hemisphere and experimental condition (F(2.43, 43.71) = .73, $p = .51$). The linear contrast between the object-motion and self-motion only conditions was highly significant (F(1,18) = 31.19, $p < 0.001$). The increased response in the object-motion conditions can't be explained by low-level visual differences between the stimuli because the physical motion of the object in OMo was identical to that in SM2, while overall low-level motion energy was much lower than in any other condition due to the absence of optic flow in OMo. Likewise, in OAp, the low-level properties of the stimulus were nearly identical to those in SM2 apart from replacing the outflow in the optic flow with inflow and the expansion of the image of the object with contraction.

Our hypothesis that pVIP is specialised for detecting motion of objects predicted that presence versus absence of stationary objects during simulated self-motion would not influence the BOLD signal. Initially this prediction appeared to be confirmed by the contrast of SM2 with SM1, which revealed a non-significantly decreased BOLD response in pVIP (F(1,18) = 1.70, $p = .21$). However, the results for SM3, in which the object was anchored to a point on the ground plane more laterally separated from the point of observation suggest the response to objects in pVIP is more complex than a simple moving versus static object dichotomy. Inspection of Fig. 5 reveals that the BOLD signal in SM3 is
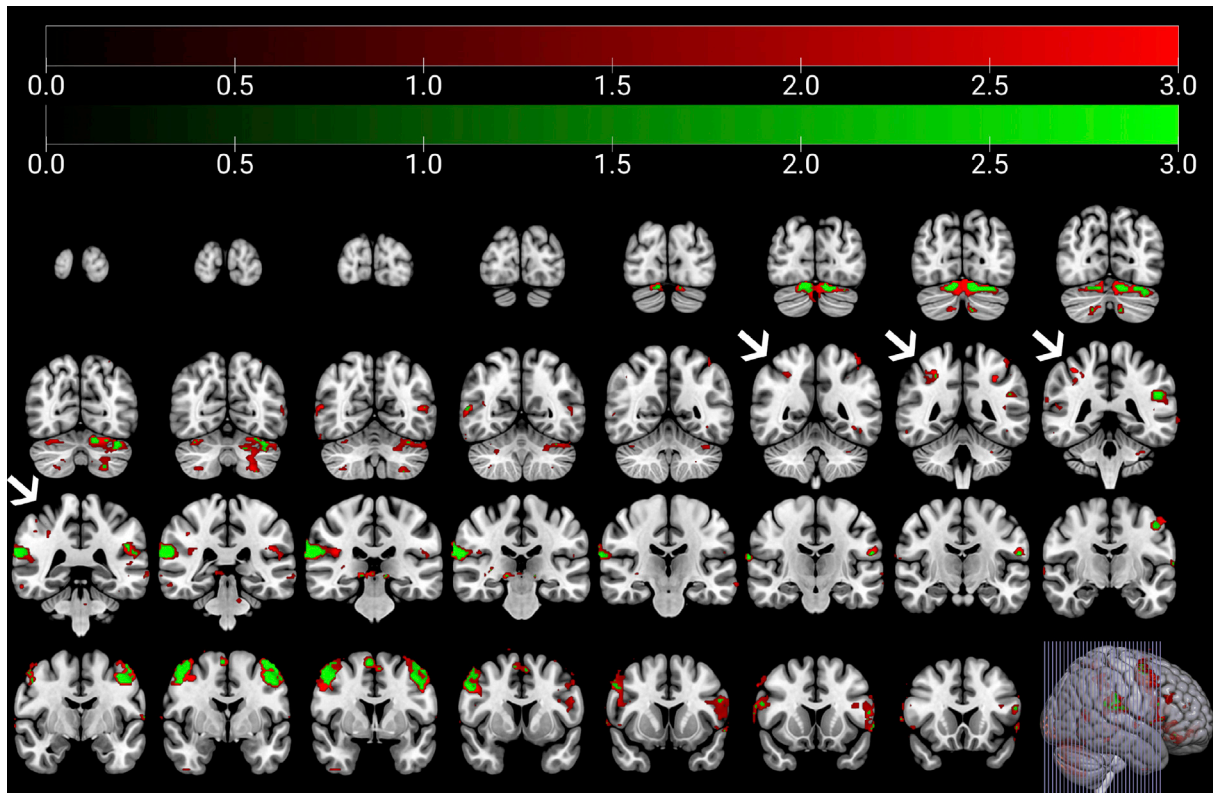
**Fig. 3.** Group level activity found in the 3-way conjunction of visual, tactile, and auditory localiser scans. Arrows highlight the ventral intraparietal sulcus; pVIP was expected to be found bilaterally in the fundus of this sulcus. Green areas were active at the more conservative threshold, while red areas were only active at the more liberal threshold (see Section 3.1 for details). Slices cover the region of the brain from the occipital pole (y = −102) up to y = 18 in 4 mm steps.

lower than either of the other conditions that included stationary objects (SM2 and SM4) and is also lower than SM1 in which no objects were present. These observations were confirmed by significant linear contrasts of SM3 versus SM2 and SM4 (F(1,18) = 5.05, $p$ = .04) and of SM3 with SM1 (F(1,18) = 16.88, $p$ < 0.001). Taken together, these three findings suggest the possibility of an inhibition of neurons in pVIP during self-motion by distant stationary objects, which is released as the observer moves closer to them.

Finally, a slower simulated self-motion condition in the presence of stationary objects condition (SM4) was included to explore whether pVIP responses might be modulated by the speed of self-motion, but contrasting this with the condition that had the closest low-level visual match (SM2) produced no evidence of that (F(1,18) = 1.94, $p$ = .18).

**Table 1**

Peak pVIP stereotaxic coordinates from this study and previous studies. Where coordinates from other studies were reported in the Talairach space they were converted to the MNI space using the method of Lacadie et al. (2008).

|  | X | Y | Z |
| --- | --- | --- | --- |
| pVIP object-motion | −34 | −47.9 | 49.6 |
|  | 36 | −45.6 | 55.6 |
| pVIP localiser | −34 | −44 | 42 |
|  | 38 | −40 | 52 |
| Bremmer et al. (2001) | −40 | −43 | 45 |
|  | 37 | −48 | 50 |
| Quinlan and Culham, 2007 | −27 | −56 | 55 |
|  | 25 | −54 | 58 |
| Wall and Smith (2008) | −26 | −56 | 43 |
|  | 25 | −51 | 43 |
| Cardin and Smith, 2010 | −25 | −59 | 55 |
|  | 25 | −60 | 43 |
| Sereno and Huang (2006) | 30 | −43 | 59 |
|  | −28 | −46 | 58 |

### 3.4. Overlap between multisensory activation and object-motion related activation in the intraparietal sulcus: pVIPobject

Our finding that responses are stronger for simulated object-motion than self-motion in pVIP localised using the multisensory response criteria suggests that object-motion selectivity could form part of a functional localiser for pVIP, in conjunction with the multisensory criteria. To test this possibility we carried out three group level, whole brain, conjunction analyses; a 2-way conjunction to localise object-motion selectivity, a 2-way conjunction to localise multisensory regions, and then a combined 4-way conjunction to localise regions that showed both multisensory properties and object-motion selectivity. The first of the two inputs to the object-motion conjunction analysis was OAp – SM2. The second was OMo – SM3 in which both conditions include retinal motion of an object, but only in the latter case was the retinal motion perceived as object-motion. Furthermore, this contrast provides the advantage that it is unlikely voxels will be declared active due to low-level visual drive because the total motion energy in the subtracted control stimulus (SM3) is far greater than in OMo due to the presence of ground plane optic flow. The two inputs to the multisensory conjunction were from our conceptual replication of Bremmer's multisensory localiser study (visual motion – rest and tactile motion – rest). The combined 4-way conjunction was carried out separately across the four contrasts described above. The individual contrasts submitted to the conjunction analyses were thresholded using an initial voxelwise cut off of Z = 3, followed by a cluster threshold of $p$ < 0.05 (corrected for multiple comparisons).

The results of the 2-way conjunctions are shown in Figs. 6 and 7, and both of these show activations in the region of pVIP, which is more extensive in the case of the object-motion conjunction. Fig. 7 shows that a number of other brain regions also showed object-motion selectivity, which are anatomically identified in section 3.5 below. To shed light on

**Table 2**
Matrices of Euclidean distances (mm) between pVIP activation coordinates in the left and right hemispheres of the brain, comparing centres of gravity from this study with the published pVIP coordinates from previous studies. The mean of the distances between each coordinate and the others is given in the rightmost column.

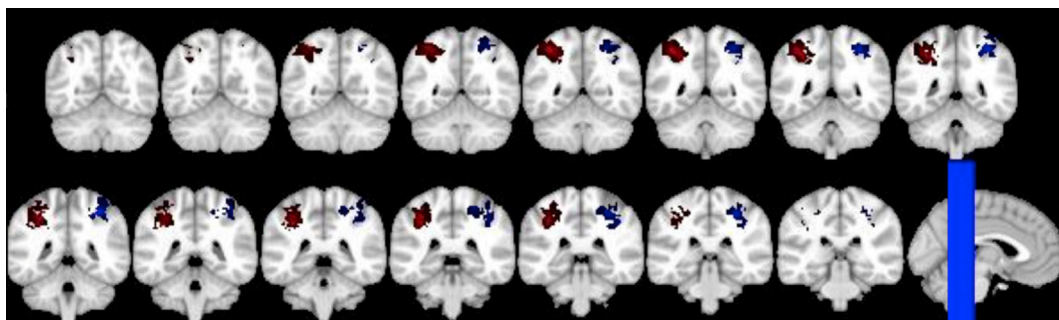| Left hemisphere | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | pVIP obj. | pVIP loc. | Bremmer | Quinlan | Wall | Cardin | Sereno | Mean |
| pVIP object-motion | | 8.5 | 9.0 | 12.0 | 13.2 | 15.3 | 10.5 | *11.6* |
| pVIP localiser | 8.5 | | 6.8 | 19.0 | 14.5 | 21.8 | 17.2 | *14.1* |
| Bremmer et al. (2001) | 9.0 | 6.8 | | 20.9 | 19.2 | 24.1 | 17.9 | *16.0* |
| Quinlan and Culham, 2007 | 12.0 | 19.0 | 20.9 | | 12.0 | 3.6 | 10.5 | *13.5* |
| Wall and Smith (2008) | 13.2 | 14.5 | 19.2 | 12.0 | | 12.4 | 18.1 | *14.3* |
| Cardin and Smith, 2010 | 15.3 | 21.8 | 24.1 | 3.6 | 12.4 | | 13.7 | *15.4* |
| Sereno and Huang (2006) | 10.5 | 17.2 | 17.9 | 10.5 | 18.1 | 13.7 | | *14.7* |
| **Right hemisphere** | | | | | | | | |
| | pVIP obj. | pVIP loc. | Bremmer | Quinlan | Wall | Cardin | Sereno | *Mean* |
| pVIP object-motion | | 7.0 | 6.2 | 14.0 | 17.6 | 22.1 | 7.4 | *13.4* |
| pVIP localiser | 7.0 | | 8.3 | 20.0 | 19.3 | 25.5 | 11.0 | *16.0* |
| Bremmer et al. (2001) | 6.2 | 8.3 | | 15.6 | 14.2 | 18.4 | 12.4 | *12.5* |
| Quinlan and Culham, 2007 | 14.0 | 20.0 | 15.6 | | 15.3 | 16.2 | 12.1 | *16.2* |
| Wall and Smith (2008) | 17.6 | 19.3 | 14.2 | 15.3 | | 9.0 | 18.6 | *15.1* |
| Cardin and Smith, 2010 | 22.1 | 25.5 | 18.4 | 16.2 | 9.0 | | 23.9 | *18.2* |
| Sereno and Huang (2006) | 7.4 | 11.0 | 12.4 | 12.1 | 18.6 | 23.9 | | *14.2* |



**Fig. 4.** Locations of pVIP ROI of individual participants registered to the MNI template brain; brighter colours indicate that a particular voxel was activated in more participants. Left hemisphere ROI are shown in red and right hemisphere ROI are shown in blue. The blue rectangle on the sagittal slice indicates the positions of the coronal slices, which cover the region from y = −57 to y = −29 in 2 mm steps. The sizes of ROI from individual participants are given in Table 3.

the possibility that our object-motion conjunction was confounded by neural activity associated with possible attentional effects of the presence of object-motion, Fig. 7 also displays the results of an automated meta-analysis of 1831 fMRI studies of attention, which we downloaded from https://neurosynth.org/analyses/terms/attention/(Yarkoni et al., 2011). We selected the recommended association test option on the Neurosynth website, and the results were thresholded with the default false discovery rate criterion of 0.01. Inspection of Fig. 7 reveals moderate overlap between the object-motion selective regions and the attention network.

The results of the 4-way conjunction analysis confirmed the existence of a region within the ventral intraparietal sulcus containing neurons that show stronger responses to simulated object-motion than self-motion and which is also multisensory (Fig. 8, top 3 panels, orange and red); here we will refer to the localised region as pVIPobject. This activation was slightly more medial and superior in the intraparietal sulcus to that produced by our conceptual replication of Bremmer's study using a 3-way conjunction of visual, tactile, and auditory response (Fig. 8, top 3 panels, shown in light blue). The reason for the difference in activation location may be related to the weakly activating auditory condition being included only in the 3-way replication localiser and the more liberal statistical threshold that its' inclusion required. Despite this small difference, the MNI coordinates of both of the intraparietal sulcus pVIPobject activations shown in Fig. 8 are closer to the multisensory pVIP identified by Bremmer et al. (2001) than they are to the pVIP coordinates of other authors (see Table 2).

**Table 3**
Thresholds used to define pVIP ROI's in individual participants. ROI volumes in mm$^3$ are also given.

| Participant | Z threshold | Volume left (mm$^3$) | Volume right (mm$^3$) |
| --- | --- | --- | --- |
| **1** | 1.6 | 1134 | 554 |
| **2** | 1.6 | 797 | 716 |
| **3** | 1.6 | 1607 | – |
| **4** | 1.6 | 540 | 446 |
| **5** | 1.6 | – | 68 |
| **6** | – | – | – |
| **7** | 2 | 3983 | 446 |
| **8** | 1.6 | 41 | 81 |
| **9** | 1.6 | 108 | 999 |
| **10** | 4 | 2282 | 2012 |
| **11** | 2.5 | 1404 | 1161 |
| **12** | 2 | 311 | 189 |
| **13** | 2 | 796.5 | 1161 |
| **14** | 1.6 | 41 | – |
| **15** | 2.55 | 284 | 148.5 |
| **16** | 5 | 1863 | 3119 |
| **17** | 4.5 | 2484 | 243 |
| **18** | 4 | 986 | 567 |
| **19** | 1.6 | 527 | – |

### 3.5. Object-motion selectivity in other brain regions with multisensory responses

Our 4-way conjunction analysis located a number of brain regions, other than pVIPobject, which were responsive to visual and tactile motion and selective for simulated object-motion over self-motion. These are anatomically labelled in Fig. 8 (panels 4–9), and their stereotaxic
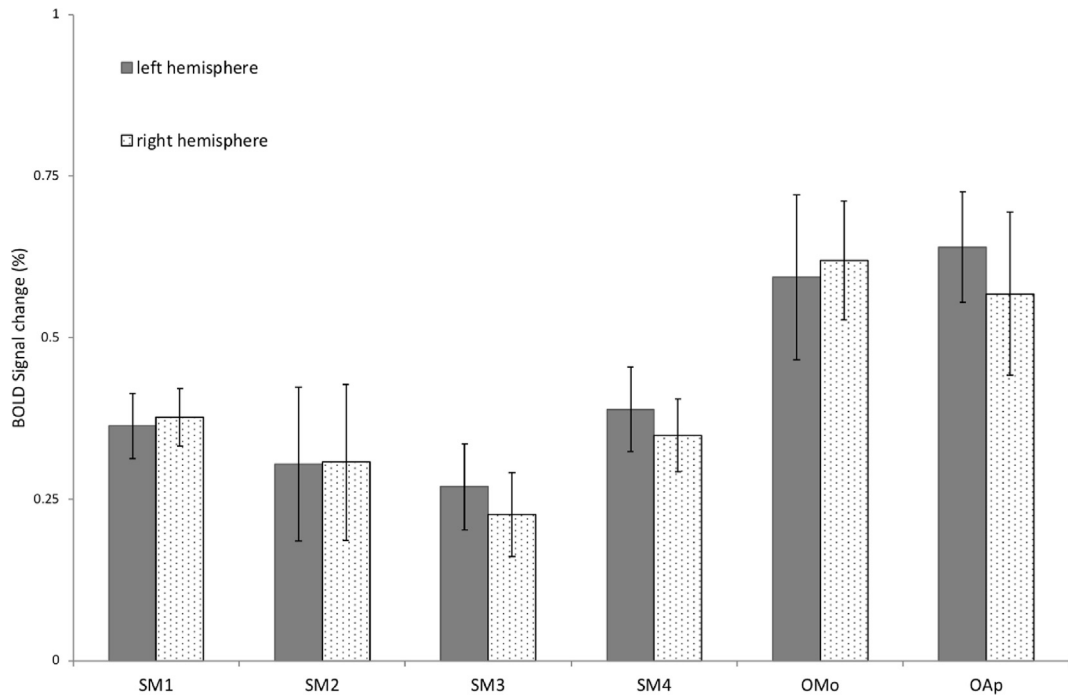
**Fig. 5.** Percent BOLD signal change relative to resting baseline in pVIP ROI's for each of the visual conditions. SM1 = simulated self-motion over a ground plane; SM2 = addition of objects anchored to ground plane; SM3 = objects more remote from viewpoint; SM4 = reduced speed of self-motion in presence of objects; OMo = no self-motion, pole moves towards observer; OAp = objects appear close to observer during self-motion producing looming percept. Error bars indicate 95% confidence intervals.
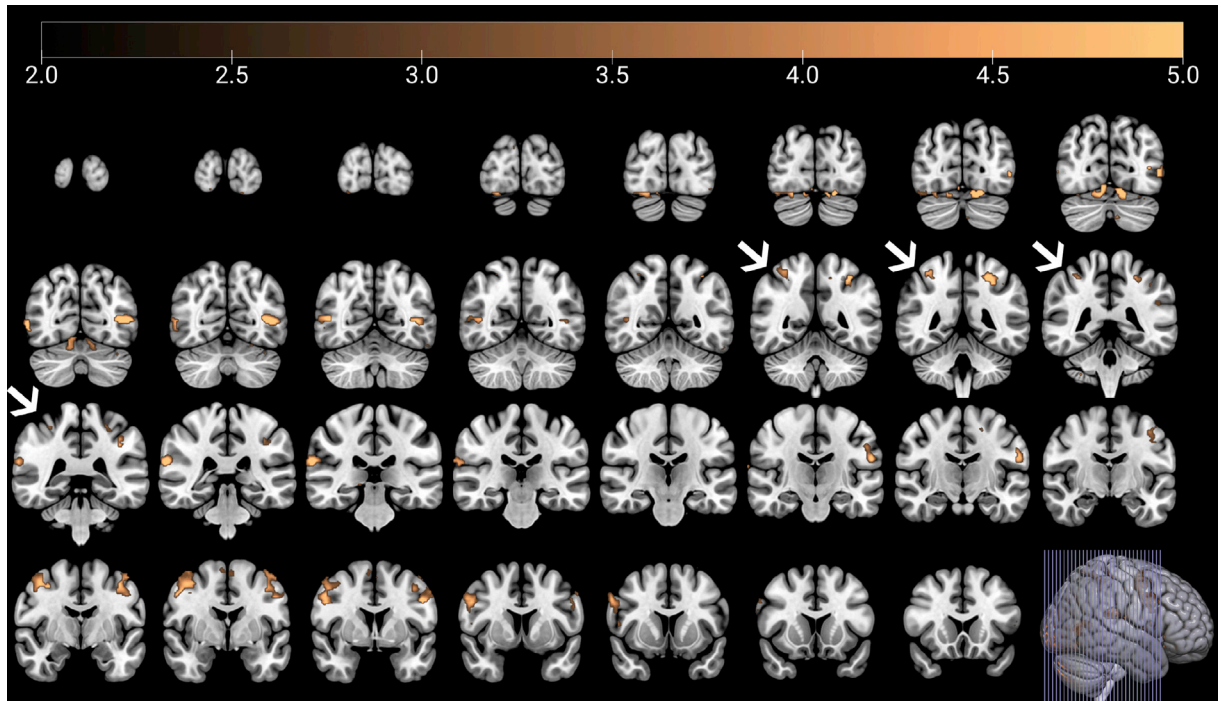


**Fig. 6.** Group level activity found in the 2-way conjunction of visual, and tactile localiser scans. Arrows highlight the ventral intraparietal sulcus; pVIP was expected to be found bilaterally in the fundus of this sulcus. Slices cover the region of the brain from the occipital pole (y = −102) up to y = 18 in 4 mm steps.

coordinates are provided in Table 4. We consider these activations further in the Discussion.

### 3.6. Behavioural task analysis

To verify that participants could accurately track changes in heading direction, the joystick x-position data were analysed using the same method as Billington et al. (2010). In summary, the rate of change of the joystick position (60 samples per second) was rescaled to match the rate of change of angular heading, allowing the two time-series to be plotted on the same axes. The temporal accuracy of tracking was quantified using cross-correlations of the actual heading with the joystick position across a
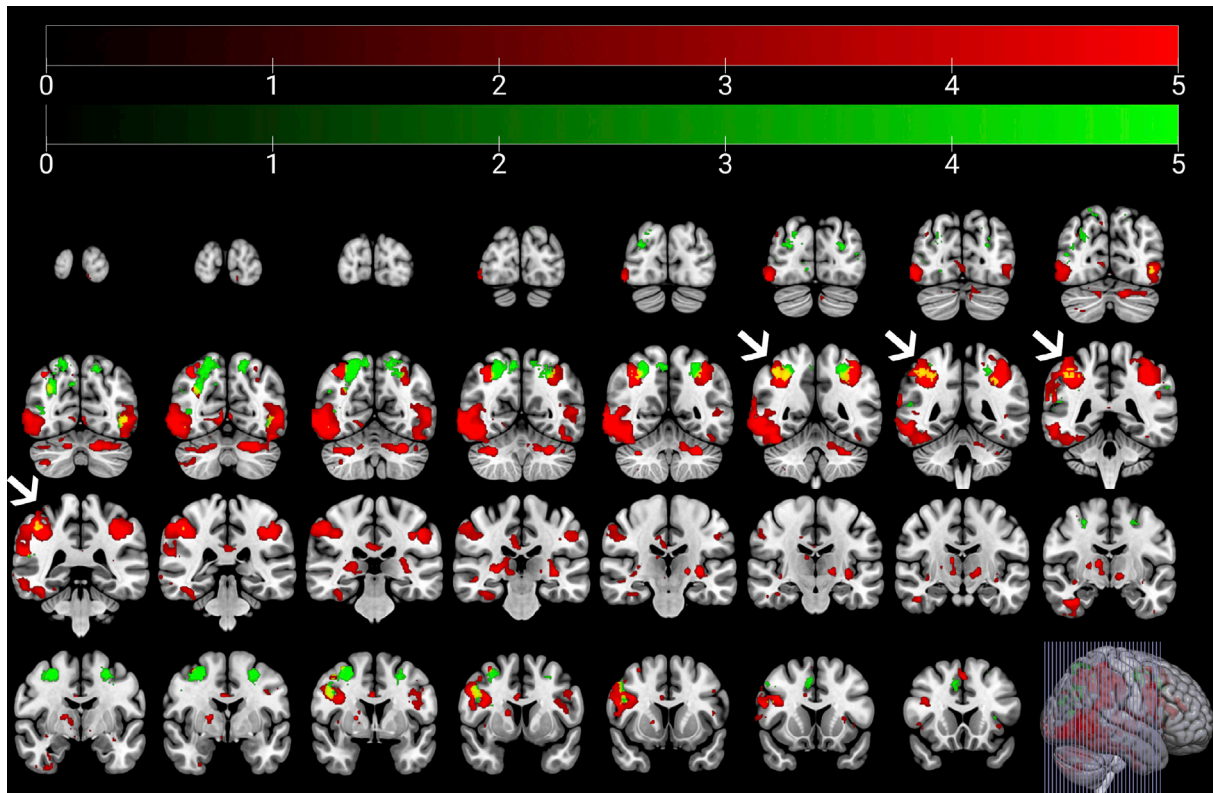
**Fig. 7.** Group level activity found in the 2-way object-motion conjunction analysis is shown in red. The attention network is shown in green for comparison; yellow indicates overlap between the two (see Section 3.4 for details). Arrows highlight the ventral intraparietal sulcus; pVIP was expected to be found bilaterally in the fundus of this sulcus. Slices cover the region of the brain from the occipital pole (y = −102) up to y = 18 in 4 mm steps.

range of plausible lags. The maximum cross-correlation lag provides an estimate of the temporal tracking lag between the joystick position and what is displayed on the screen at that moment. Note however, that this method of calculating the lag does not correct for the delays produced by the joystick and the recording computer, which are likely to add about 100 msec. The spatial accuracy of tracking the heading changes was assessed by the $R^2$ value of the fit of the joystick positions at the chosen lag to the heading changes.

Fig. 9 (top panel) compares the grand average time course of joystick responses during each experimental condition with the displayed heading. From this it is apparent that participants produced joystick time courses that were an accurate reflection of the heading changes with a lag of around 0.75 s. Conditions SM1-4 produced very similar patterns of response, while in OAp the joystick response amplitude was reduced relative to SM1-4, resulting in worse spatial tracking accuracy. We speculate that this occurred because participants are far more experienced with the expanding flow fields caused by travelling forwards than they are with contracting flow fields caused by travelling backwards. In the OMo condition the time course of joystick responses included deflections that are not present in the displayed heading, or in the joystick responses made in the other experimental conditions. These deflections occurred only in OMo because it was the one experimental condition in which information to control joystick movements became unavailable intermittently – this occurred in the period between the occlusion of a pole and the next one appearing.

The two lower panels of Fig. 9 plot the average lag and spatial tracking accuracy in each experimental condition. A one-way repeated measures ANOVA confirmed that there were no significant differences in lag between conditions. However, spatial accuracy was significantly lower in OAp and OMo than the other conditions; this was confirmed by a one-way ANOVA ($F(2.20, 39.53) = 10.64$, $p < .001$), followed by a significant linear contrast comparing OAp and OMo together against SM1-4 ($p < .001$). There was no significant difference between OAp and OMo in

terms of spatial accuracy ($t(18) = 1.01$, $p = .326$).

## 4. Discussion

VIP/pVIP is a functionally complex brain region which receives input from several sensory modalities, contains a somatosensory representation of the face that is yoked to a representation of visual space near the face, and is clearly responsive to moving objects that approach the face. But neurons in this region also respond to optic flow and encode information about heading direction that would be relevant during locomotion, even though locomotion is a 'far visual space' activity that does not generally involve objects in very close proximity to the face. In this study we used the region's multisensory property to localise pVIP and then used well controlled visual stimuli to probe its responses to simulated self-motion with and without objects present. We also compared the case where an object is perceived to move with that where it rests on the ground-plane during self-motion. Our results reveal a much stronger response to visual motion if the visual context implies that the motion is caused by object-motion than if the same motion is perceived as being due to simulated self-motion. Critically, this pattern can't be explained in terms of the low-level statistics of visual motion. Given that detecting object-motion cannot proceed in isolation from processing self-motion we suggest that the purpose of the self-motion signals that are present in VIP/pVIP is to allow 'parsing' of the flow field into self-motion and object-motion components. Thus, we conclude that the functional role of VIP/pVIP in perception is related to object-motion, while perception of self-motion is supported by one or more of the other brain regions that have been shown to contain optic-flow tuned neurons. Importantly, our conclusion reached on the basis of human data can also explain the results of macaque deactivation studies, which have found that deactivating VIP does not influence heading perception while deactivating MST does (Gu et al., 2012; Chen et al., 2016).
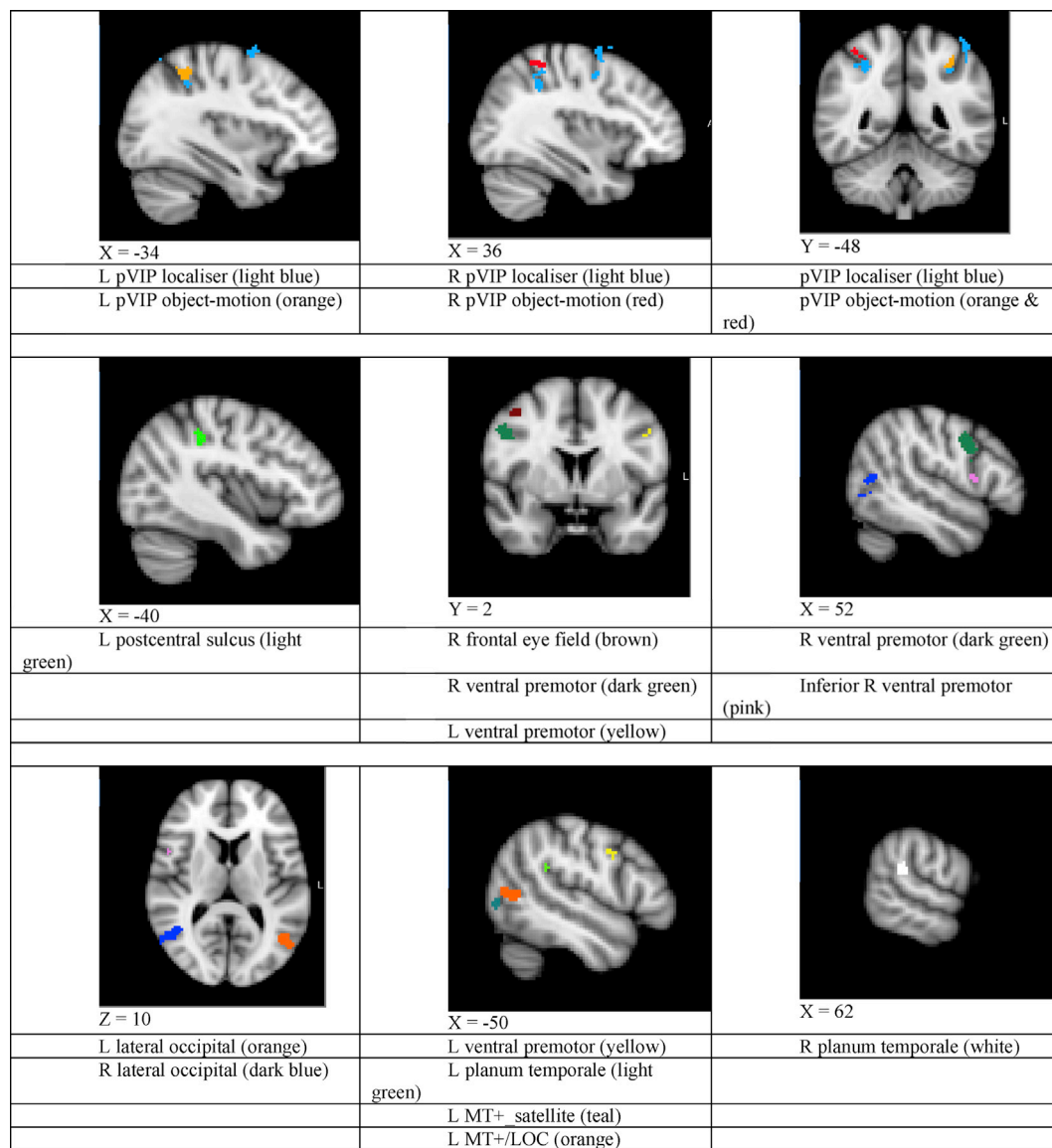
**Fig. 8.** Top three panels show group level pVIP activation produced by the multisensory localiser (conjunction of visual motion, tactile motion, and auditory motion, with individual components of the conjunction thresholded at $p < 0.05$ uncorrected for multiple comparisons), as well as pVIP activation produced by the 4-way conjunction of object-motion selectivity and multisensory response (visual motion and tactile motion), with individual components of the conjunction thresholded using an initial voxelwise cut off of Z = 3, followed by a cluster threshold of $p < 0.05$ (corrected for multiple comparisons). The bottom six panels present the other brain regions activated by the latter conjunction.

**Table 4**

Stereotaxic coordinates of centres of gravity of other brain regions with similar functional responses to pVIPobject (both object-motion selective and multisensory).

|                               | X   | Y   | Z  |
|-------------------------------|-----|-----|----|
| L postcentral sulcus          | −41 | −37 | 41 |
| R ventral premotor            | 53  | 7   | 33 |
| R ventral premotor (inferior) | 52  | 10  | 10 |
| R frontal eye field           | 42  | 1   | 51 |
| L ventral premotor            | −50 | 3   | 35 |
| L lateral occipital           | −48 | −66 | 9  |
| L lateral occipital (inferior)| −49 | −76 | 1  |
| R lateral occipital           | 52  | −64 | 7  |
| L planum temporale            | −51 | −42 | 26 |
| R planum temporale            | 62  | −36 | 21 |

Although our data show much greater responsiveness in pVIP for moving than stationary environmental objects, and the presence of stationary objects on or near the path traversed during simulated self-motion does not produce a response higher than the optic flow created by simulated self-motion alone, we did make a novel observation that suggests stationary objects can influence responses in pVIP under some circumstances. Specifically, when more distant objects were presented (SM3) pVIP activation levels were lower than for either the absence of objects, or objects lying on the path travelled. Future studies should investigate in more detail what this finding suggests – a possible inhibition of neurons in pVIP by distant stationary objects during self-motion, which is released if the observer moves closer to them. It would be particularly informative for future studies to measure responses in pVIP when simulated object-motion occurs at different distances from the viewpoint as well as in the presence and absence of simulated self-motion and the presence and absence of stationary objects. There are suggestive parallels to this finding in the macaque VIP literature, where it has been found that those VIP neurons that are disparity tuned tend to prefer
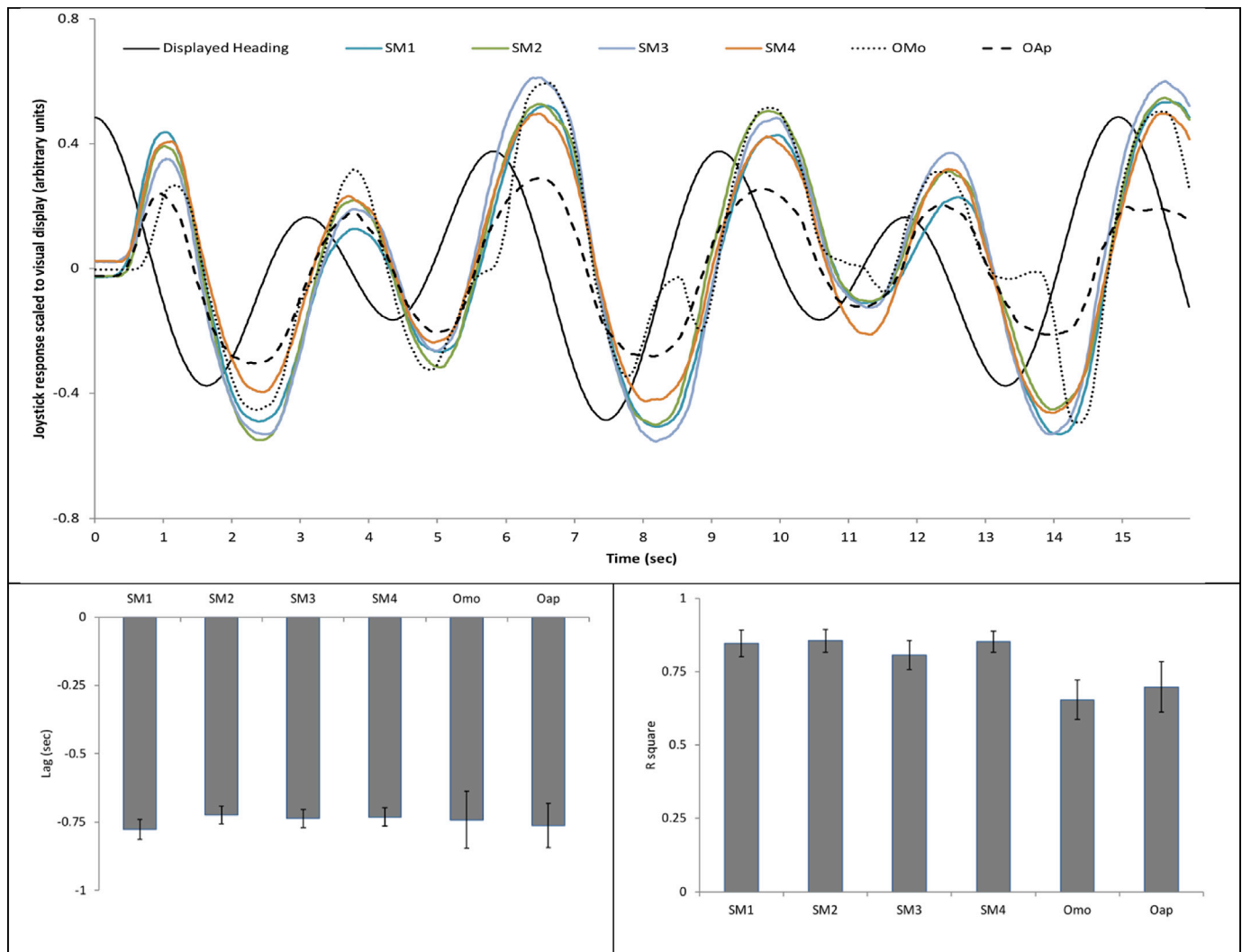
**Fig. 9.** Top panel: joystick movements made under the six experimental conditions in response to the displayed rate of change of heading, or in OMo, the motion path of the object. See Section 3.6 for details. Lower left panel: mean lags between displayed heading and joystick responses. Lower right panel: mean spatial tracking accuracy of joystick responses relative to the displayed heading ($R^2$). Error bars indicate 95% confidence intervals.

crossed disparities that are indicative of an object nearby, and also that macaque VIP has stronger BOLD responses for real objects placed near the animal than far space (Bremmer et al., 2013; Clery et al., 2018).

Previous localisations of pVIP have reported somewhat different anatomical locations, reflected in the stereotaxic coordinates reported in Table 1, and there has consequently been some debate as to whether the different studies of pVIP are focusing on the same functional region (Huang et al., 2017). Our aim was to test responses to object-motion in the putative human homologue of macaque VIP, and so we followed the multisensory localization procedure of Bremmer et al. (2001). The region we localised using this procedure was found in the fundus of the intraparietal sulcus, and in both hemispheres its group level stereotaxic coordinates were closer to those reported by Bremmer et al. (2001) than to those from other pVIP studies (Table 2). However, using our stimuli multisensory activity was stronger elsewhere in the brain than in pVIP.

Previously, Bartels et al. (2007) questioned Bremmer's multisensory localization of pVIP because a type of conjunction analysis was used that potentially allows voxels active in a single modality to be declared active; we used a more conservative method of conjunction analysis not subject to this weakness, and therefore our replication clears up this methodological doubt, and shows that it is possible to locate a multisensory response in the human ventral intraparietal sulcus. However, the auditory response was not sufficiently robust to be used as an efficient

localiser in individual participants; future studies could explore different auditory stimuli, but given the very robust response to object-motion we found in pVIP, we believe that a localiser based on that functional property will prove more efficient.

The pVIPobject region we defined based on its responsiveness to optic flow, tactile stimulation of the face, and a greater response for environmental motion than simulated self-motion is located on the anterior wall of the intraparietal sulcus, slightly superior to the region defined by the multisensory localiser (see Fig. 8). Despite the difference in the activation location, the object-motion selective region was still located closer to the pVIP defined by Bremmer than to the pVIP locations of other studies (Table 2). The slightly different location of the object-motion selective region probably reflects our omission of the inconsistent and statistically weak auditory localiser results from the conjunction analysis used to define it. That omitting one of the three sensory modalities from the conjunction analysis slightly shifted the activation focus in the intraparietal sulcus suggests potential subdivisions within pVIP in which individual sensory modalities have relatively greater dominance. Such subdivisions would be consistent with what is known about macaque VIP; patches with relatively greater dominance of one or two of the three sensory modalities have recently been identified in macaque VIP (Guipponi et al., 2013), and a previous architectonic parcellation study identified lateral and medial subdivisions within macaque VIP (Lewis

and Van Essen, 2000a). The possibility of subdivisions in pVIP could also potentially explain the differing pVIP locations reported by different studies, most of which use stimuli confined to a single sensory modality. Further studies are required to test this suggestion.

Whole brain conjunction analysis showed that several brain regions apart from pVIPobject combine responsiveness to optic flow, tactile stimulation of the face, and a greater response for simulated environmental motion than self-motion. One of these was found unilaterally in a position only slightly anterior, lateral, and inferior to our left pVIP location. However, this region is not confusable with pVIP because, as Fig. 8 shows, it is located in the postcentral sulcus rather than the intraparietal sulcus.

Another area showing sensitivity to all of optic flow, air moving over the face, and a greater response to visual motion when it is perceived environmental in origin was the ventral premotor cortex (bilateral, though right dominant); this is the same region which showed a right dominant activation in the study of polymodal motion processing carried out by Bremmer et al. (2001). This area is the projection zone of area VIP in monkeys (Luppino et al., 1999), so it is not surprising that it also shows enhanced responses for environmental motion. Nearby, we also found a smaller activation cluster directly inferior to the main ventral premotor cluster in the right hemisphere, as well as activation superior and lateral to the main ventral premotor cluster that is likely to correspond to the Frontal Eye Field (FEF).

Thirdly, the 4-way conjunction analysis also revealed a bilateral activation in the lateral occipital cortex, and an additional smaller cluster located inferior and posterior to the main left hemisphere cluster. Given the location, one possibility is that this activation corresponds to the lateral occipital complex (LOC), which has tactile responses (Amedi et al., 2001, 2002) and is known to be specialised for processing of objects, so is likely to be more responsive when objects move. Consistent with the possibility that object-motion in our stimuli drove a response in LOC, specific responses to motion in LOC were found during free viewing of a movie by Bartels et al. (2007). However, it is also possible given the nearby location that the activation we found corresponds to part of the motion complex, MT+; Bartel's et al. (2007) found specific responses for object-motion rather than self-induced flow during free viewing of a movie in this region. A future study of simulated self-motion versus object-motion perception should include functional localisers for MT+ and LOC to distinguish between these two possibilities.

If the activation we found does correspond to MT + rather than LOC then functional considerations suggest that MST is the most likely sub region of MT + that could explain this because it is responsive to tactile stimulation, while MT is much less responsive. Furthermore, MST also has strong reciprocal connections with VIP (Lewis and Van Essen, 2000b). However, it can't be ruled out that the type of tactile stimulation we used provoked visual imagery, which would indirectly activate MT (Beauchamp et al., 2007). Also raising doubt, the stereotaxic coordinates of our activation are located lateral and superior relative to the average of published coordinates for MST (Dukelow et al., 2001; Cardin et al., 2012; Pitzalis et al., 2013), and so it is not possible to conclude that our activation corresponds to MST. Nonetheless, it is worth highlighting that recent primate studies have examined the ability of two subdivisions of MST (MSTd and MSTl) to dissociate movements of large objects in the frontoparallel plane from visual motion due to self-motion; of the two, MSTd performed better (Sasaki, et al, 2017, 2019). Although the visual stimuli used in these studies were very different to those used here, this does highlight the value of future studies of object-motion processing including specific localisers for the subdivisions of MT + as well as the LOC in order to facilitate cross-species comparison.

The final region active in our 4-way conjunction analysis was located at the posterior end of the planum temporale in the right hemisphere, with a corresponding but much smaller activation in the left hemisphere. The weaker left hemisphere activation fell within Wernicke's area, and the stronger activation is the homologous right hemisphere region. This region has previously been found to be responsive to visual optic flow

similar to that used in our localiser (Antal et al., 2008). Although we could not find previous reports of activation by tactile stimulation of the face similar to that used in our study at this location, Bremmer et al. (2001) reported activation somewhat anterior to this location for the 3-way conjunction of auditory, visual, and tactile motion. This region has also previously been found to take part in the integration of auditory and vestibular signals, and is involved in perception of the movement of auditory stimuli (Eikhoff et al., 2006; Krumbholz et al., 2005; Pavani et al., 2002), so it is unsurprising that we found a response to environmental motion there.

We used two different visual stimuli to produce a percept of object-motion, and in one of these (OAp) a radial contraction component in the optic flow indicated that the viewpoint was travelling backwards. This raises the possibility that the increased response in pVIP to this specific stimulus was caused by simulation of travelling backwards rather than by the appearance of an object near the viewpoint; our current experiment did not include the necessary control condition to rule this out, which would require the simulated viewpoint to travel backwards in the absence of objects. However, one of our previous studies did include conditions in which optic flow indicative of both simulated forwards and backwards travel could be compared (in the absence of object-motion), and no activation difference was found between the two conditions in the vicinity of pVIP (Billington et al., 2010). Furthermore, a PET study that compared inward and outward radial flow found that both types of flow activated the same brain regions, but that activation for inward flow was weaker (Ptito et al., 2001). These findings of these studies rule out the possibility that the increased response to OAp in pVIP was due to the simulated backwards motion of the viewpoint. Consistent with this, the number of VIP neurons tuned to expanding flow fields and inhibited by contracting ones is roughly double the number tuned to contraction (Bremmer et al., 2002a). However, in a human psychophysical study of the ability to detect optic flow caused by forwards and backwards postural sway, both directions were equally detectable (Fitzpatrick and McCloskey, 1994). A second feature present in OAp but not the other experimental conditions was the intermittent rapid fade-in of the visually large objects, while in other conditions the equivalent fade-in events were visually smaller and less salient. It is possible that this difference drove the high BOLD signal change for OAp the pVIP ROI (Fig. 5), although the brief perceptual correlate of this difference was of a visual event not caused by self-motion – this was the percept we were seeking to achieve but is not possible to disentangle this from the low-level difference with the other experimental conditions. Finally, even if some voxels were specifically activated by the contracting flow in OAp, or by the intermittent rapid fade-in of a visually large object, these factors would not influence the voxels highlighted by our four-way conjunction analysis used to identify object-motion/multisensory selectivity in the brain because the other inputs to the conjunction analysis did not include those features.

Rather than including a fixation cross, which would require participants to perform the effortful task of suppressing their spontaneous urges to make saccades and pursuit eye movements in response to our complex visual stimuli, we used naturalistic viewing conditions. This has a number of advantages and disadvantages, and possible implications for interpretation of our results, which warrant discussion. Allowing participants to directly view the parts of the flow-field they found most informative for performing the heading tracking task makes our results more generalisable to how the brain might respond to visual stimuli in everyday life; while staring at a fixation cross mis-locates the flow field in retinotopically organised visual areas relative to the ecological context. Secondly, introducing a fixed visual reference point into the stimulus creates a strong percept of relative motion that we were keen to avoid, particularly in our baseline condition SM1; this condition simulated self-movement in the absence of objects, yet a fixation cross *is* an object in the scene, and including it would have produced confounding neural activity associated with the presence of objects. However, a potential disadvantage of naturalistic viewing is that the type or quantity of eye movements

made is likely to differ between experimental conditions, and because planning and executing eye movements is a significant source of neural activity this could account for observed differences in BOLD signal between conditions. In our experiment it is highly likely that participants performed a mixture of visual tracking of objects and of features of the flow field such as the FOE, but the relative balance between these two activities would have varied between conditions. At one extreme, SM1 would contain no tracking of objects because none were present, and at the other OAp would be most dominated by tracking of object-motion because no other visual motion was present. The other conditions, SM2-4 and OMo would contain a more balanced mixture of the two types of tracking because both object-motion and optic flow from the ground plane were present. The question arises whether the differential BOLD signal in pVIP shown in Fig. 5 could have been driven by such eye movement differences? While we can't entirely rule out this possibility on the basis of the current data set, it does seem unlikely since a model of pVIP function based on the pattern of eye movement differences between conditions just described would predict incorrectly that OAp would have a different signal change in pVIP to all other conditions, and also that SM1 would differ systematically from SM2-4. It is also worth considering what might happen in a version of our experiment that included strict fixation conditions. Firstly, it is effortful for participants to fixate in naturalistic scenes, and this effort would produce neural activity associated with suppression, which could potentially confound results if the effort required is unequal between conditions. Secondly, eye movement related activity may persist in these conditions because suppressed eye movements, being essentially covert shifts of spatial attention, also produce neural activity (Beauchamp et al., 2001). An empirical investigation of these complex issues, which included an experimental condition identical to SM1 (but termed 'Flow') was carried out by Field et al. (2007). In that study, SM1 was repeated with and without fixation, as was a condition in which road edges – which influenced measured eye movements in the no fixation condition – were added to the ground plane. A separate localiser for brain regions involved in producing saccadic eye movement was performed. The addition of road edges produced activation in superior parietal lobule (SPL) that was not present in SM1, regardless of whether or not a fixation cross was added; in both cases some but not all of the SPL voxels activated by the road edges were also activated in the eye movement localiser task. While that experiment was not focused on pVIP, its results support the general point that preventing overt eye movements when viewing naturalistic motion stimuli by inclusion of a fixation cross does not in any case prevent neural activity associated with eye movements due to the increase in covert attention shifts and planned but unexecuted saccades that the fixation cross produces.

Related to the issue of eye movements, it could be argued that the fluctuation of attention over time differed between experimental conditions. In OMo, there were brief periods when there was only a static ground plane and no object on the screen and therefore there was no visual information present to guide the joystick tracking movements, which is reflected in the joystick traces presented in Fig. 9. In all the other experimental conditions that contained objects, there was an identical cycle of appearance and disappearance of objects, but dynamic task-relevant visual motion information from the ground plane was still present in the brief object-free periods. It is reasonable to assume that attention was attracted by objects in all conditions, but more so in OMo as object-motion was the only information available to perform the joystick task; therefore, greater fluctuation of attention may have occurred in OMo, driven by the duty cycle of the object appearance and disappearance. On the other hand, in conditions SM2-4 and OAp participants could choose to attend mainly to ground plane flow or to alternate attention between objects and the ground plane, and possibly object appearance and disappearance could act as a distraction to participants who tried to focus on ground plane flow. While we can't rule out that greater salience of object appearance and disappearance in OMo had some effect on pVIP signal change, such attentional fluctuation accounts

don't predict the differences in pVIP signal change between other experimental conditions shown in Fig. 5, e.g. under that account why should signal change be higher in SM1 where there were no objects and so less fluctuation than in condition SM3 where there were objects?

As well as the fluctuation of attention over time that differed between the object-motion and other conditions, joystick tracking accuracy was also slightly worse in those conditions, and this too could point to attention related differences in brain activation. One approach the issue of whether the object-motion selective activations we found in pVIP and other regions can be explained by differences in attention is to compare the locations of the activations with those of regions strongly associated with attention in previous literature; a high degree of overlap would suggest that the activation specific to the experimental conditions containing object-motion in this experiment was driven by attention rather than any specifically object-motion related processing. We did this and the results (Fig. 7) indicated that a moderate proportion of the voxels activated in our object-motion contrasts are part of the attention network, suggesting that a higher proportion had some more specific functional role in object-motion processing. Future studies of object-motion processing in pVIP should aim for closer control over attention and behavioural performance to clarify this issue.

In conclusion, although pVIP is known to be responsive to optic-flow and to carry information about self-motion, we have found that pVIP is more responsive to visual motion that implies environmental object-motion than to that implying self-motion. We propose that the self-motion signals pVIP receives are used there in a subtractive way, to isolate elements of the optic flow that can't be explained by current self-motion. Being independent of the visual input, the vestibular information about self-motion that pVIP receives would be particularly useful for this. To test this idea requires recordings to be made from the brain while study participants undergo the rotations and accelerations that stimulate the vestibular system, at the same time as experimentally controlled visual input. This is not possible with fMRI methods that require the participants to remain still, but might be achieved using near infrared spectroscopy (NIRS) to measure cerebral blood flow (Ferrari and Quaresima, 2012), or by exploiting new developments in wearable magnetoencephalography (MEG) systems that allow freedom for participant movement (Boto et al., 2018).

## Declaration of competing interest

The authors have no competing interests to declare.

## CRediT authorship contribution statement

**David T. Field:** Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing, Supervision. **Nicolò Biagi:** Visualization, Formal analysis. **Laura A. Inman:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2020.116679.

# References

Amedi, A., Malach, R., Hendler, T., Peled, S., Zohary, E., 2001. Visuo-haptic object-related activation in the ventral visual pathway. Nat. Neurosci. 4 (3), 324–330.

Amedi, Amir, Jacobson, G., Hendler, T., Malach, R., Zohary, E., 2002. Convergence of visual and tactile shape processing in the human lateral occipital complex. Cerebr. Cortex 12 (11), 1202–1212. New York, N.Y.: 1991.

Andersson, J L, Jenkinson, M, Smith, S, 2007. Non-linear registration aka Spatial normalisation FMRIB Technial Report TR07JA2. FMRIB Analysis Group of the University of Oxford.

Antal, A., Baudewig, J., Paulus, W., Dechent, P., 2008. The posterior cingulate cortex and planum temporale/parietal operculum are activated by coherent visual motion. Vis. Neurosci. 25 (1), 17–26.

Bartels, A., Zeki, S., Logothetis, N.K., 2007. Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. Cerebr. Cortex 18 (3), 705–717.

Beauchamp, M.S., Petit, L., Ellmore, T.M., Ingeholm, J., Haxby, J.V., 2001. A parametric fMRI study of overt and covert shifts of visuospatial attention. Neuroimage 14 (2), 310–321.

Beauchamp, M.S., Yasar, N.E., Kishan, N., Ro, T., 2007. Human MST but not MT responds to tactile stimulation. J. Neurosci. 27 (31), 8261–8267.

Billington, J., Field, D.T., Wilkie, R.M., Wann, J.P., 2010. An fMRI study of parietal cortex involvement in the visual guidance of locomotion. J. Exp. Psychol. Hum. Percept. Perform. 36 (6), 1495–1507.

Boto, E., Holmes, N., Leggett, J., Roberts, G., Shah, V., Meyer, S.S., et al., 2018. Moving magnetoencephalography towards real-world applications with a wearable system. Nature 555 (7698), 657.

Bremmer, F., Churan, J., Lappe, M., 2017. Heading representations in primates are compressed by saccades. Nat. Commun. 8 (1), 920.

Bremmer, F., Duhamel, J.-R., Hamed, S.B., Graf, W., 2002a. Heading encoding in the macaque ventral intraparietal area (VIP). Eur. J. Neurosci. 16 (8), 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 (8), 1569–1586.

Bremmer, F., Schlack, A., Kaminiarz, A., Hoffmann, K.P., 2013. Encoding of movement in near extrapersonal space in primate area VIP. Front. Behav. Neurosci. 7, 8.

Bremmer, F., Schlack, A., Shah, N.J., Zafiris, O., Kubischik, M., Hoffmann, K.-P., et al., 2001. Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys. Neuron 29 (1), 287–296.

Britten, K.H., van Wezel, Richard J.A., 1998. Electrical microstimulation of cortical area MST biases heading perception in monkeys. Nat. Neurosci. 1 (1), 59–63.

Calabro, F.J., Vaina, L.M., 2012. Interaction of cortical networks mediating object-motion detection by moving observers. Exp. Brain Res. 221 (2), 177–189.

Cardin, V., Hemsworth, L., Smith, A.T., 2012. Adaptation to heading direction dissociates the roles of human MST and V6 in the processing of optic flow. J. Neurophysiol. 108 (3), 794–801.

Cardin, V, Smith, A T, 2010. Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. Cerebral cortex 20 (8), 1964–1973.

Chen, A., DeAngelis, G.C., Angelaki, D.E., 2013a. Functional specializations of the ventral intraparietal area for multisensory heading discrimination. J. Neurosci. 33 (8), 3567–3581.

Chen, A., Gu, Y., Liu, S., DeAngelis, G.C., Angelaki, D.E., 2016. Evidence for a causal contribution of macaque vestibular, but not intraparietal, cortex to heading perception. J. Neurosci. 36 (13), 3789–3798.

Chen, X., DeAngelis, G.C., Angelaki, D.E., 2013b. Eye-Centered representation of optic flow tuning in the ventral intraparietal area. J. Neurosci. 33 (47), 18574–18582.

Cléry, J., Guipponi, O., Odouard, S., Wardak, C., Hamed, S.B., 2015. Impact prediction by looming visual stimuli enhances tactile detection. J. Neurosci. 35 (10), 4179–4189.

Cléry, J., Guipponi, O., Odouard, S., Wardak, C., Hamed, S.B., 2018. Cortical networks for encoding near and far space in the non-human primate. Neuroimage 176, 164–178.

Cléry, J., Guipponi, O., Odouard, S., Pinède, S., Wardak, C., Hamed, S.B., 2017. The prediction of impact of a looming stimulus onto the body is subserved by multisensory integration mechanisms. J. Neurosci. 37 (44), 10656–10670.

Colby, C.L., Duhamel, J.R., Goldberg, M.E., 1993. Ventral intraparietal area of the macaque: anatomic location and visual response properties. J. Neurophysiol. 69 (3), 902–914.

Duffy, C.J., Wurtz, R.H., 1991. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J. Neurophysiol. 65 (6), 1329–1345.

Dukelow, S.P., DeSouza, J.F.X., Culham, J.C., van den Berg, A.V., Menon, R.S., Vilis, T., 2001. Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. J. Neurophysiol. 86 (4), 1991–2000.

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. Hum. Brain Mapp. 27 (7), 611–621.

Ferrari, M., Quaresima, V., 2012. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage 63 (2), 921–935.

Field, D T, Inman, L A, Li, L, 2015. Visual processing of optic flow and motor control in the human posterior cingulate sulcus. Cortex 71, 377–389.

Field, D.T., Wilkie, R.M., Wann, J.P., 2007. Neural systems in the visual control of steering. J. Neurosci. 27 (30), 8002–8010.

Fitzpatrick, R., McCloskey, D.I., 1994. Proprioceptive, visual and vestibular thresholds for the perception of sway during standing in humans. J. Physiol. 478 (1), 173–186.

Furlan, M., Wann, J.P., Smith, A.T., 2014. A representation of changing heading direction in human cortical areas pVIP and CSv. Cerebr. Cortex 24 (11), 2848–2858.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15 (4), 870–878.

Gibson, J.J., 1950. The Perception of the Visual World. Houghton Mifflin, Oxford, England.

Gu, Y., Cheng, Z., Yang, L., DeAngelis, G.C., Angelaki, D.E., 2016. Multisensory convergence of visual and vestibular heading cues in the pursuit area of the frontal eye field. Cerebr. Cortex 26 (9), 3785–3801.

Gu, Y., DeAngelis, G.C., Angelaki, D.E., 2012. Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. J. Neurosci. 32 (7), 2299–2313.

Guipponi, O., Wardak, C., Ibarrola, D., Comte, J.C., Sappey-Marinier, D., Pinède, S., Hamed, S.B., 2013. Multimodal convergence within the intraparietal sulcus of the macaque monkey. J. Neurosci. 33 (9), 4128–4139.

Huang, R.S., Chen, C.F., Sereno, M.I., 2017. Mapping the complex topological organization of the human parietal face area. Neuroimage 163, 459–470.

Jenkinson, M, Bannister, P, Brady, M, Smith, S, 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17 (2), 825–841.

Jenkinson, M, Smith, S, 2001. A global optimisation method for robust affine registration of brain images. Medical image analysis 5 (2), 143–156.

Kaminiarz, A., Schlack, A., Hoffmann, K.P., Lappe, M., Bremmer, F., 2014. Visual selectivity for heading in the macaque ventral intraparietal area. J. Neurophysiol. 112 (10), 2470–2480.

Krumbholz, K., Schönwiesner, M., Rübsamen, R., Zilles, K., Fink, G.R., von Cramon, D.Y., 2005. Hierarchical processing of sound location and motion in the human brainstem and planum temporale. Eur. J. Neurosci. 21 (1), 230–238.

Lacadie, C.M., Fulbright, R.K., Rajeevan, N., Constable, R.T., Papademetris, X., 2008. More accurate Talairach coordinates for neuroimaging using non-linear registration. Neuroimage 42 (2), 717–725.

Lewis, J.W., Van Essen, D.C., 2000a. Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex. J. Comp. Neurol. 428 (1), 79–111.

Lewis, J.W., Van Essen, D.C., 2000b. Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. J. Comp. Neurol. 428 (1), 112–137.

Luppino, G., Murata, A., Govoni, P., Matelli, M., 1999. Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). Exp. Brain Res. 128 (1), 181–187.

Nichols, T, Brett, M, Andersson, J, Wager, T, Poline, J B, 2005. Valid conjunction inference with the minimum statistic. Neuroimage 25 (3), 653–660.

Pavani, F., Macaluso, E., Warren, J.D., Driver, J., Griffiths, T.D., 2002. A common cortical substrate activated by horizontal and vertical sound movement in the human brain. Curr. Biol.: CB 12 (18), 1584–1590.

Pitzalis, S., Sdoia, S., Bultrini, A., Committeri, G., Di Russo, F., Fattori, P., et al., 2013. Selectivity to translational egomotion in human brain motion areas. PloS One 8 (4), e60241.

Ptito, M., Kupers, R., Faubert, J., Gjedde, A., 2001. Cortical representation of inward and outward radial motion in man. Neuroimage 14 (6), 1409–1415.

Quinlan, D.J., Culham, J.C., 2007. fMRI reveals a preference for near viewing in the human parieto-occipital cortex. Neuroimage 36 (1), 167–187.

Raffi, M., Squatrito, S., Maioli, M.G., 2002. Neuronal responses to optic flow in the monkey parietal area PEc. Cerebr. Cortex 12 (6), 639–646.

Rushton, S.K., Niehorster, D.C., Warren, P.A., Li, L., 2018. The primary role of flow processing in the identification of scene-relative object movement. Journal of Neuroscience 38 (7), 1737–1743.

Rushton, S.K., Warren, P.A., 2005. Moving observers, relative retinal motion and the detection of object movement. Current Biology 15 (14), R542–R543.

Sasaki, R., Angelaki, D.E., DeAngelis, G.C., 2017. Dissociation of self-motion and object-motion by linear population decoding that approximates marginalization. J. Neurosci. 37 (46), 11204–11219.

Sasaki, R., Angelaki, D.E., DeAngelis, G.C., 2019. Processing of object-motion and self-motion in the lateral subdivision of the medial superior temporal area in macaques. J. Neurophysiol. 121 (4), 1207–1221.

Schaafsma, S.J., Duysens, J., 1996. Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. J. Neurophysiol. 76 (6), 4056–4068.

Sereno, M.I., Huang, R.-S., 2006. A human parietal face area contains aligned head-centered visual and tactile maps. Nat. Neurosci. 9 (10), 1337–1343.

Siegel, R.M., Read, H.L., 1997. Analysis of optic flow in the monkey parietal area 7a. Cerebr. Cortex 7 (4), 327–346.

Smith, S M, 2002. Fast robust automated brain extraction. Human brain mapping 17 (3), 143–155.

Wall, M.B., Smith, A.T., 2008. The representation of egomotion in the human brain. Curr. Biol. 18 (3), 191–194.

Warren, W.H., Mestre, D.R., Blackwell, A.W., Morris, M.W., 1991. Perception of circular heading from optical flow. J. Exp. Psychol. Hum. Percept. Perform. 17 (1), 28–43.

Warren, P.A., Rushton, S.K., 2009. Optic flow processing for the assessment of object movement during ego movement. Current Biology 19 (18), 1555–1560.

Woolrich, M.W., Behrens, T.E.J., Beckmann, C.F., Jenkinson, M., Smith, S.M., 2004. Multilevel linear modelling for FMRI group analysis using Bayesian inference. Neuroimage 21 (4), 1732–1747.

Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14 (6), 1370–1386.

Yang, Y., Liu, S., Chowdhury, S.A., DeAngelis, G.C., Angelaki, D.E., 2011. Binocular disparity tuning and visual–vestibular congruency of multisensory neurons in macaque parietal cortex. J. Neurosci. 31 (49), 17905–17916.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D., 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat. Methods 8 (8), 665.

Zhang, T., Britten, K.H., 2004. Clustering of selectivity for optic flow in the ventral intraparietal area. Neuroreport 15 (12), 1941.

Zhang, T., Britten, K.H., 2011. Parietal area VIP causally influences heading perception during pursuit eye movements. J. Neurosci. 31 (7), 2569–2575.

Zhang, T., Heuer, H.W., Britten, K.H., 2004. Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. Neuron 42 (6), 993–1001.