

ORIGINAL ARTICLE

Stereoscopically Observing Manipulative Actions

S. Ferri¹, K. Pauwels², G. Rizzolatti¹, and G. A. Orban¹¹Department of Neuroscience, University of Parma, 43125 Parma, Italy and ²Computer Vision and Active Perception Laboratory, School of Computer Science and Communication, KTH, 10044 Stockholm, Sweden

Address correspondence to Guy A. Orban, Department Neuroscience, Via Volturno, 39, 43125 Parma, Italy. Email: guy.orban@kuleuven.be

Abstract

The purpose of this study was to investigate the contribution of stereopsis to the processing of observed manipulative actions. To this end, we first combined the factors “stimulus type” (action, static control, and dynamic control), “stereopsis” (present, absent) and “viewpoint” (frontal, lateral) into a single design. Four sites in premotor, retro-insular (2) and parietal cortex operated specifically when actions were viewed stereoscopically and frontally. A second experiment clarified that the stereo-action-specific regions were driven by actions moving out of the frontoparallel plane, an effect amplified by frontal viewing in premotor cortex. Analysis of single voxels and their discriminatory power showed that the representation of action in the stereo-action-specific areas was more accurate when stereopsis was active. Further analyses showed that the 4 stereo-action-specific sites form a closed network converging onto the premotor node, which connects to parietal and occipitotemporal regions outside the network. Several of the specific sites are known to process vestibular signals, suggesting that the network combines observed actions in peripersonal space with gravitational signals. These findings have wider implications for the function of premotor cortex and the role of stereopsis in human behavior.

Key words: action observation, cerebral cortex, gravity, human fMRI, space, stereopsis

Introduction

The displacement of eyes to a more frontal position in the heads of primates, thus allowing stereopsis, is a key feature of mammalian evolution (Barton 2004; Heesy et al. 2011). Yet investigations into the processing of stereo signals in the brain have until now concentrated on the perception of objects (Tsao et al. 2003; Likova and Tyler 2007; Parker 2007; Orban 2011), with very little known about the mechanisms that underlie the role of stereopsis in perceiving observed actions. Because the human body is nonrigid, stereopsis is crucial for recovering positions and speeds of body parts for actions moving out of the frontoparallel (FP) plane of the observer, since monocular characteristics, such as size changes, are informative only for rigid bodies (Regan and Gray 2001). It is significant that most interactions with conspecifics occur in face-to-face situations and involve such out-of-the-FP-plane (OFPP) action. The plane in which actions unfold in the observers' visual field depends not only on his viewpoint

but also on the axis along which the actor moves his hand. Therefore, to clarify the distinction between the axis in space, linked to the actor, and the direction of the action for the observer, we use the term OFPP when referring to the observer and the terms x- and z-axes to characterize actions in the peripersonal space of the actor. Physics thus predicts that stereopsis is essential for processing observed actions, but only OFPP actions, and stereopsis should therefore improve the fidelity of any such action representation. Although it seems obvious that viewpoint and stereopsis interact, stereopsis was not manipulated in the very few single-cell studies investigating the effect of viewpoint on premotor neuron activity (Caggiano et al. 2011).

Thus far only a single functional imaging study, from our group, has investigated the role of stereopsis in processing observed actions. This study (Jastorff et al. 2016) reported that stereopsis increased activation in premotor and, to some degree, parietal nodes of the action observation network. However, this study, which used biological motion (BM) stimuli to portray

diverse complex actions, left many questions unanswered, in that multiple viewpoints and action types were intermingled, and lacked any background or scene against which the actions took place. In the present study, we utilize videos portraying actions belonging to a well-defined and well-studied class: object manipulation (Binkofski et al. 1999; Iacoboni et al. 1999; Buccino et al. 2001; Culham et al. 2003; Gazzola et al. 2007; Grafton and Hamilton 2007; Jastorff et al. 2010; Mukamel et al. 2010; Gallivan et al. 2016). The actions were shown against a background from 2 different allocentric viewpoints: frontal and lateral. To control for the lower-order effects of disparity, we also included static and dynamic control conditions. Hence, the main experiment followed a $3 \times 2 \times 2$ factorial design, with “stimulus type,” “stereopsis,” and “viewpoint” as factors.

We have introduced a novel analysis complementing traditional univariate analyses, used here in the factorial main experiment and in the previous study (Jastorff et al. 2016). While classical analyses can identify areas involved by indicating areas where stereopsis enhances action processing, they cannot, however, provide information concerning the accuracy of observed action representations; much less demonstrate that stereopsis increases accuracy. To probe directly the action representation in those areas, we developed a novel single-voxel analysis using d' as a measure of neural distance between observed actions. If voxels “discriminate” better between the observed actions when stereopsis is enabled, we expect d' of single voxels to increase in stereo conditions. This analysis may prove a welcome alternative to multivoxel pattern analysis (MVPA) (Norman et al. 2006), a machine learning technique which indicates that a representation is present, but not which representation, as the voxels driving the classification are unknown (Bulthe et al. 2015; Wardle et al. 2015). In addition, MVPA frequently returns classification performances exceeding chance level by only a few percent (Gallivan et al. 2013; Eger et al. 2015; Pilgramm et al. 2016).

The main experiment revealed a previously unknown, stereo-action-specific network, centered on a premotor node, describing how actions unfold in space, including gravitational effects. Additional experiments bore out the prediction of the physics, that the activity in this network would depend on the observed actions being OFPP for the observer and that these actions are more accurately represented by this network when stereopsis is enabled. These results provide new insights into why the visual processing of perceived actions should involve premotor cortex and expands the known role of stereopsis in visual perception.

Materials and Methods

Participants

Twenty-four right-handed, healthy volunteers participated in the present study. This study included a behavioral screening and 3 brain-imaging experiments. All subjects participated in the behavioral and the first main experiment. Data from one subject were discarded because of excessive head motion in the first scanning session, hence the statistical analysis of the first experiment included only 23 subjects (11 females; mean age: 26.2 years, range: 22–30). Sixteen of these (9 females; mean age: 25.3 years, range: 22–30) also participated in the second functional magnetic resonance imaging (fMRI) experiment. Nine of those 16 further participated in the third imaging experiment (5 females; mean age: 24.7 years, range: 22–30). All participants had normal or corrected-to-normal vision. Before admitting subjects to the study, they were tested in the behavioral screening session to assess their stereoscopic vision (see below). This study was approved

by the Ethical Committee of Parma Province, and all volunteers gave written, informed consent prior to undergoing scanning sessions, in accordance with the Helsinki Declaration.

Behavioral Screening

As stereopsis tends to be variable within the normal population, participants were screened in a single behavioral session before scanning. The participants were presented with 32 blocks, lasting 20.8 s and containing 8 action videos presented with or without stereopsis. The videos were identical to those used in the main fMRI experiment (see below for a complete description of the stimuli). The stimuli were presented on a high-definition 3D monitor (Samsung SynchMaster TA950) under control of E-prime software, with the order of blocks randomized. The participants, sitting on a chair at a distance of 1.25 m from the screen, viewed stimuli with polarized glasses. At the end of each block subjects reported verbally whether the videos were presented stereoscopically or not. Only the 24 subjects (out of 34 tested) giving 100% correct responses were admitted to the scanning sessions.

Stimuli and Designs

We used a series of 3 imaging experiments. The first main experiment was used to investigate interactions between stimulus type, stereopsis, and viewpoint. We then conducted 2 subsequent experiments with a simpler factorial structure that allowed us to deconstruct the precise nature of the interactions among these factors. Crucially, the main effects and interactions of the first experiment were used to define regions of interest in the subsequent experiments (in terms of regional summaries or averages). This allowed us to finesse the multiple comparisons problem and provide a detailed picture of the functional anatomy of stereoscopic action observation.

In all 3 imaging experiments, videos lasted 2.6 s, corresponding to 65 frames at a frame rate of 25 frames/s. For the videos presented stereoscopically this framerate applied to each of the 2 eyes, as left and right images were recorded on a single large frame. The size of the video images ($18.4^\circ \times 20.8^\circ$) was approximately double that employed in our previous action studies (Jastorff et al. 2010; Abdollahi et al. 2013; Ferri, Peeters, et al. 2015; Ferri, Rizzolatti, et al. 2015), in order to obtain a sufficient range of depths in the stereoscopic videos.

The aim of the first, “main experiment” was to investigate the effects of stereopsis and viewpoint on the processing of observed manipulative actions. To this end, we created video clips showing an actor (male or female) sitting on a chair close to a table, and manipulating an object placed on the table with the right hand. Four exemplars of the action class “manipulation” were presented: pushing, grasping, dragging, and displacing an object, the first 3 being identical to those used in the Jastorff et al. (2010), Abdollahi et al. (2013), and Ferri, Rizzolatti, et al. (2015) studies. The video started with the actor’s right hand resting on the table, next to the object, so that the action was performed using only the fingers and the wrist, with no arm movement. Most of the actions were directed along the z-axis of the actor’s peripersonal space. All 4 actions were performed by a male and female actor on a small red cube and a large blue ball, generating 4 versions of each exemplar. In order to compare the effects of the 2 viewpoints, all actions were recorded by 2 identical camcorders (Panasonic HCX 900), one positioned to the actor’s left side (lateral viewpoint) and the other in front (frontal viewpoint). To enhance the depth range of the recordings and to equalize it in videos taken from the 2 viewpoints, we took the following

measures: (1) the distance between the camcorder and the object on the table was fixed at 1.2 m for both viewpoints, yielding a full view of the subject's upper body (Fig. 1); (2) the distance between the object and the background was set to 1.3 m for the 2 viewpoints; (3) a heterogeneous pattern similar to a wood texture served as the background for both viewpoints; (4) the recordings were synchronized by using the same remote control for both camcorders. To create stereoscopic videos, both camcorders were fitted with dual lenses for 3D recording (Panasonic VW-CLT2), positioned next to each other in the horizontal plane, thus recording the same actor from 2 slightly different positions. Each camcorder therefore produced pairs of frames: those seen by the right lens and those seen by the left (Fig. 1).

The 3 design factors yielded 12 conditions, since the third factor, stimulus type included 3 levels: action videos, static, and dynamic controls. These control stimuli were needed for the figural and motion components that are inherent to actions, but also to control for lower-order static and dynamic disparity. For convenience, we will refer to the conditions in which stereopsis is present as "stereo-conditions" and those in which it was absent as "mono-conditions" even if all stimuli were displayed binocularly. This terminology indicates that while stereopsis is available as the depth cue in the stereo-conditions, only monocular depth cues are present in the mono-conditions. Hence, the 12 conditions of the first experiment included: Stereo-action frontal, mono-action frontal, stereo-static-control frontal, mono-static-control frontal, stereo-dynamic-control frontal, mono-dynamic-control frontal; stereo-action lateral, mono-action lateral, stereo-static-control lateral, mono-static-control lateral, stereo-dynamic-control lateral, and mono-dynamic-control lateral.

These stimuli were post-processed using Matlab software as follows. First, all the frames of each stereoscopic video were split vertically into halves, separating the left and right eye images, so that 2 videos were obtained from each original. Second, the left and right videos were each resized to $18.4^\circ \times 20.8^\circ$. Third, the mono- and stereo-action videos were created by pairing the identical images (RE-RE or LE-LE) or the left and right images (RE-LE), respectively. Fourth, the static control and dynamic scrambled controls were derived for all action videos. In the static controls, the first, middle, or last frame was selected from both mono and stereo-action videos to capture the shape of the actor at different stages of the action. These static controls not only control for body shape, but also lower-order factors such as spatial frequency and color. To control for the motion component, we created dynamic scrambled control videos by applying the procedure of Ferri, Rizzolatti, et al. (2015) to the videos shown to each eye, a process involving the extraction of optic flow (Pauwels and Van Hulle 2009) followed by temporal scrambling. In summary, the 4 experimental conditions (stereo on/off and both viewpoints) were all derived from the same physical action performed by the actor, with the 2 control conditions matched to each of these 4 experimental videos. Finally, to assess the visual nature of the fMRI signals, we included a fixation condition as the explicit baseline. In this condition, a rectangle having the same size and average luminance as the videos was shown.

The aim of the second experiment was to differentiate the effect of viewpoint from that of the axis of motion in space. In the first experiment these 2 factors remained confounded, as almost all actions unfolded along the z-axis of the actor and were thus

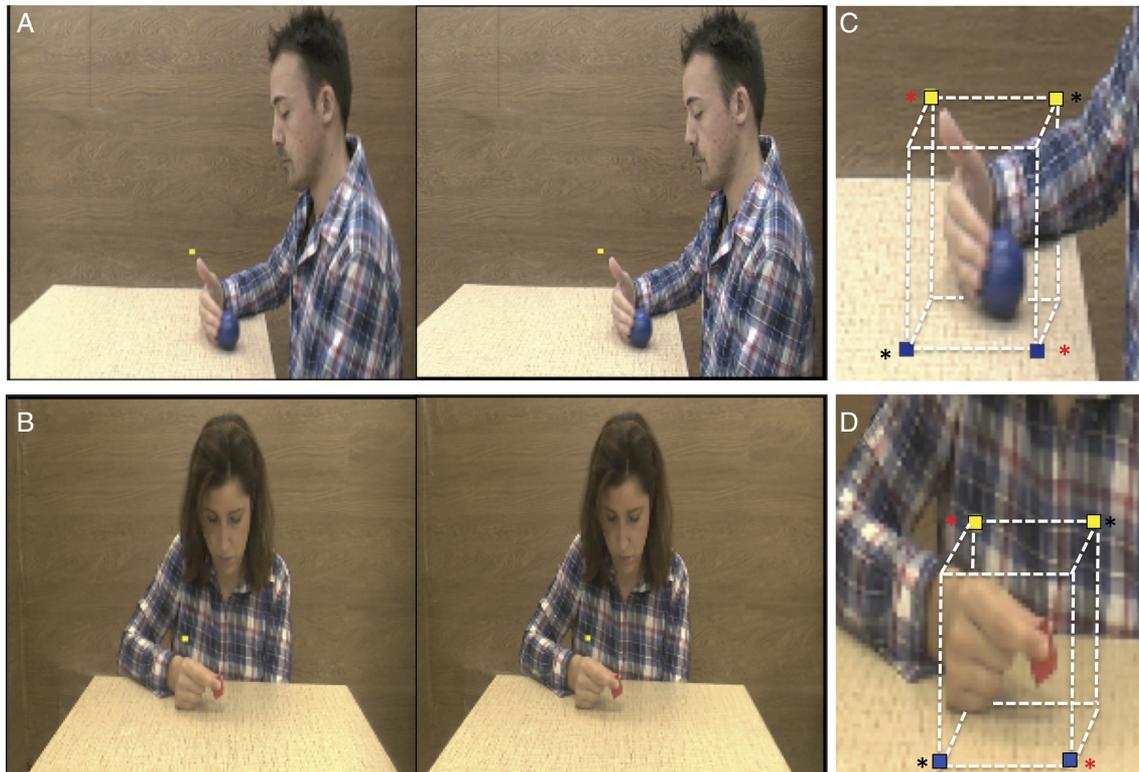


Figure 1. (A,B) Stimuli: stereo-pair of static frames taken from the stereo videos: lateral view (A) and frontal view (B); (C,D) positions of the fixation points in lateral (C) and frontal (D) runs of the first main experiment: yellow, behind the object, blue, in front; in (A,B) yellow dots: fixation point; in (C,D) red asterisks indicate the 2 fixation points used in half of the subjects, the black asterisks those used in the other half during the first session of experiment 1; Fixation points were swapped during the second session, thus all subjects used all 4 fixation points in experiment 1; in individual runs upper and lower fixation points were alternated.

OFPP actions for the frontal viewpoint but not for lateral. Therefore, in this second “axis” experiment the 4 manipulative stereo-actions of the first experiment were performed either along the z-axis or the x-axis of the actor’s peripersonal space and were viewed either frontally or laterally in a 2×2 design. Both camcorders (lateral and frontal viewpoints) were thus used to record stereoscopic video clips of the same 4 action exemplars used in the main experiment (grasping, dragging, pushing, and displacing), and these actions were directed along either the z or x axis of the actor (Fig. 2). Two actors (one male and one female) manipulated either a small red cube or a large blue ball, producing 4 versions of each action along a given axis, for a total of 16 action videos per condition. During the recording, made in the same setting as the main experiment, we followed the precautionary measures described above for the main experiment.

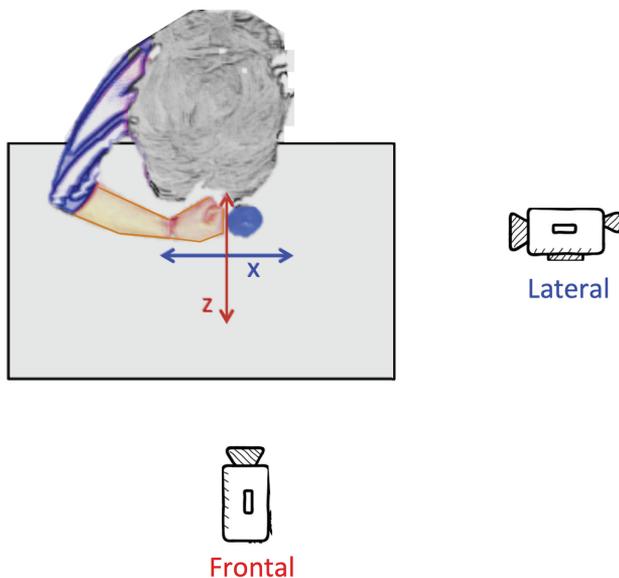


Figure 2. Conditions of the axis control experiment generated by 2 axes of action (z, x) in peripersonal space of the actor and 2 viewpoints (lateral, frontal).

The design of the second experiment yielded 4 experimental conditions: stereo-actions viewed frontally and directed along the actor’s z-axis (Frontal-Z), stereo-actions viewed frontally and directed along the x-axis (Frontal-X), stereo-actions viewed laterally and directed along the z-axis (Lateral-Z), stereo-actions viewed laterally and directed along the x-axis (Lateral-X), plus a fixation condition, as the explicit baseline.

The aim of the third experiment was to investigate whether single voxels in activation sites identified in the main experiment could better discriminate between action exemplars in stereo-conditions than in mono-conditions. In this “discrimination” experiment, we used the stereo and mono-action videos from the main experiment, but seen only from the frontal viewpoint. The experimental design thus included 8 experimental conditions (stereo-grasping, stereo-pushing, stereo-dragging, stereo-displacing, mono-grasping, mono-pushing, mono-dragging, mono-displacing), with a fixation condition as the explicit baseline.

Experimental Procedures

The first main experiment followed a $3 \times 2 \times 2$ design with factors type of stimulus (action videos and static and dynamic controls), stereopsis (stereo, mono), and viewpoint (frontal, lateral). Each of the two 8-run sessions included 4 runs with either frontal- or lateral-viewpoint videos, reducing the design to 2×3 for single runs (Fig. 3). To half the subjects, the frontal view was shown in odd-numbered runs in the first session and even runs in the second; for the remaining half, this sequence was reversed. Within a run, the 5 conditions (stereo-action, stereo-control [static or dynamic], mono-action, mono-control [static or dynamic], fixation) were presented in blocks lasting 20.8 s and repeated 4 times, defining 4 cycles of 5 blocks (Fig. 3), for a total per-run duration of 416 s. Each experimental block included 8 videos, corresponding to 4 action exemplars and both genders. The factor object was spread over successive cycles; hence in each run all 16 action video clips (4 exemplars, 2 genders, and 2 objects) for any given condition were shown twice. Because we wished to control for both static shape and motion, 2 cycles of a given run included a stereo- and a mono-block with 8 static frames, while the other 2 cycles included blocks consisting of 8 dynamic scrambled videos. Both

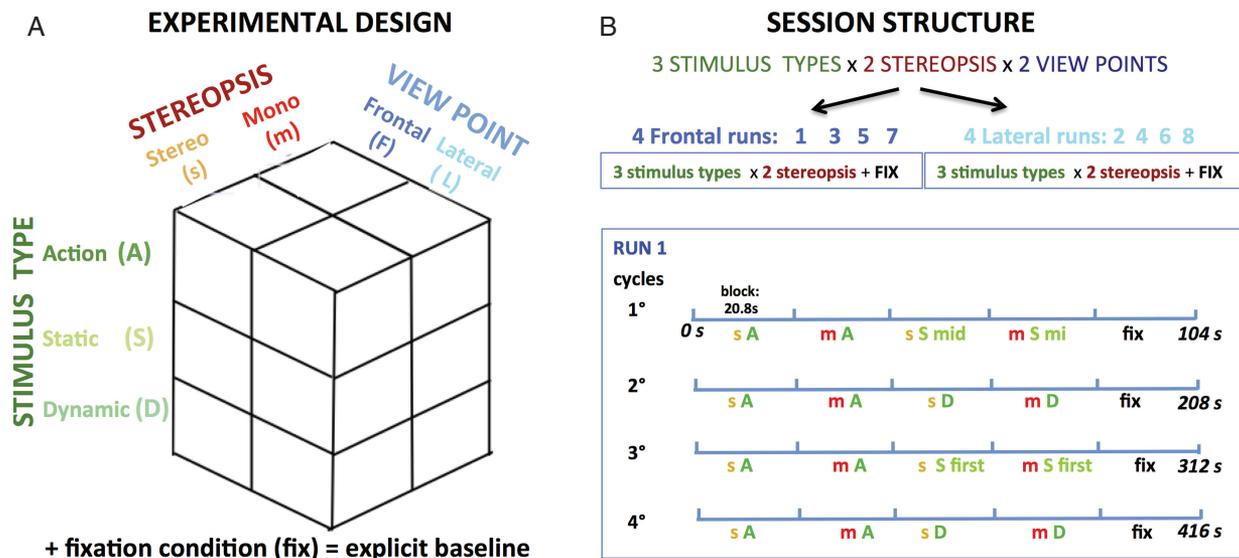


Figure 3. Experimental design (A) and session plus run structure (B) of the first main experiment. Runs included 4 cycles of all 5 conditions in identical order. Across cycles, control conditions alternated, as well as the static frame used; A: action, S: static control, D: dynamic control, mid: middle, fix: fixation, s: stereo, m: mono.

the order of the videos within a block and the order of the blocks themselves were pseudorandomly selected and counterbalanced across runs and participants (with the fixation condition always presented last in the cycle). Moreover, variations within a given condition were equalized across one or several runs. Because the mono-stimuli were generated by the 2 identical left (LE-LE) or right (RE-RE) images taken from the original 3D videos (see above), LE-LE or RE-RE videos were interleaved within each block. In each static control block, a different frame, either the first, middle or last, was selected for all 8 stimuli. These 3 possibilities alternated randomly over the static blocks of the different runs (Fig. 3).

The second axis experiment followed of a 2×2 design with factors viewpoint (frontal and lateral) and direction of the action (along the z or x axis relative to the actor), while the factors stereopsis and stimulus type remained constant, as only stereoscopes were presented. A single MRI session included 8 runs, in which the 5 conditions (frontal-Z, frontal-X, lateral-Z, lateral-X, fixation) were presented in 20.8 s blocks repeated 4 times, yielding 4 cycles, for a total run duration of 416 s, as in the main experiment. Each experimental block included 8 videos, corresponding to 4 action exemplars and both genders. Within a run, all video clips for a given condition ($n = 16$: 4 exemplars, 2 genders, and 2 objects) were shown twice. Both the videos within a block and the order of the blocks were pseudorandomly selected (with the fixation condition always presented last in a cycle), and counterbalanced across runs and participants.

The third discrimination experiment followed a 2×4 design with stereopsis (stereo-mono) and action exemplars (grasping, dragging, pushing, displacing) as factors. Moreover, a fixation condition was added to the 8 experimental conditions, as in the previous experiments. These 9 conditions were presented in 10.4-s mini-blocks repeated 7 times, for a total run duration of 655.2 s. Each mini-block included all 4 versions of the action exemplar (2 genders \times 2 objects). Again, both the videos within a mini-block and the order of the mini-blocks were selected pseudorandomly (with fixation always presented last in a cycle), and counterbalanced across runs and participants. Ten runs were collected in a single session.

In all 3 experiments, the subjects lay motionless in the bore of the scanner while observing videos and maintaining gaze on a stationary fixation point positioned above or below the hand of the actor. To avoid the fixation point having a constant position across runs and subjects, which would bias the lateralization of the activations, we used 4 different fixation point positions relative to the hand in all runs of the 3 experiments: 2 positions above and behind the hand (yellow in Fig. 1C,D); and 2 below and in front of the hand (blue in Fig. 1C,D). Because the data of the first main and second axis experiments were analyzed using univariate statistics, requiring averaging across subjects, fixation-point positions were counterbalanced across subjects in these experiments. In half the subjects, the fixation point was located either up and right (in 4 runs) or down and left (other 4 runs) relative to the hand (black asterisks in Fig. 1C,D), using the other 2 positions (up left and down right) in the other subjects (red asterisks in Fig. 1C,D). In the first experiment, the fixation points were exchanged in the second session. In the discrimination control experiment the 4 positions were instead counterbalanced across the 10 runs in each subject, as it relied on an individual subject statistical analysis.

Data Collection

Participants lay supine in the bore of the scanner. Visual stimuli were presented in the FP plane at the same distance at which

they were recorded, by means of a head-mounted display (60 Hz refresh rate) with a resolution of 800 horizontal \times 600 vertical pixels (Resonance Technology, Northridge, CA) for each eye. The display was controlled by an ATI Radeon 2400 DX dual output video card (AMD, Sun Valley, CA), which allowed a stretched desktop presentation onto the horizontal plane (1600 horizontal pixels), corresponding to the images of the 2 eyes. Thus the display presented horizontal binocular disparities (range -8 minarc to 8 minarc, see [Supplementary Fig. 1](#)) by providing separate images for each eye, without shuttering. Sound-attenuating headphones were used to muffle scanner noise. The presentation of the stimuli and the recording of the onset of each block were controlled by E-Prime software (Psychology Software Tools, Sharpsburg, PA). To reduce the amount of head motion during scanning, the subjects' head was padded with PolyScan™ vinyl-coated cushions. To ensure that the subjects fixated properly, an eye tracker system (60 Hz, Resonance technology, Northridge, CA) monitored eye movements throughout the duration of each run. This is important for proper interpretation of the data, because several cortical regions, including parietal cortex, display saccade-related BOLD responses ([Kimmig et al. 2001](#); [Koyama et al. 2004](#)). Moreover, this tracking system included 2 cameras, positioned on the left and on the right sides of the goggles, recording the positions of both eyes simultaneously. This binocular system allowed us to explicitly control for vergence eye movements, by measuring the standard deviation (SD) of binocular differences in eye position. Again, this is an important control since vergence eye movements have been shown to activate regions in the precentral sulcus, as do saccades ([Alvarez et al. 2010](#)).

Scanning was performed using a 3T MR scanner (GE Discovery MR750, Milwaukee, ILL) with an 8-parallel-channel receiver coil, located in the University Hospital of the University of Parma. Functional images were acquired using gradient-echo-planar imaging with the following parameters: 49 horizontal slices (2.5 mm slice thickness; 0.25 mm gap), repetition time (TR) = 3 s, time of echo (TE) = 30 ms, flip angle = 90° , 96×96 matrix with FOV 240 (2.5×2.5 mm in plane resolution), and ASSET factor of 2. The 49 slices contained in each volume covered the entire brain. Each run started with the collection of 4 dummy volumes to ensure that the MR signal had reached a steady state. A 3-dimensional (3D) high-resolution T_1 -weighted IR-prepared fast SPGR (Bravo) image, covering the entire brain, was acquired in the first scanning session and used as anatomical reference. Its acquisition parameters were as follows: TE/TR 3.7/9.2 ms; inversion time 650 ms, flip-angle 128, acceleration factor (ARC) 2; 186 sagittal slices acquired with $1 \times 1 \times 1$ mm³ resolution. A single scanning session required about 90 min. In the sessions of the first main experiment and the second axis experiment, 1120 volumes (140 volumes \times 8 runs) were collected. 2520 volumes (252 volumes \times 10 runs) were collected in the third discrimination experiment.

Statistical Analysis

First Main Experiment

Data analysis was performed using the SPM8 software package (Wellcome Department of Cognitive Neurology, London, UK) running under Matlab (The Mathworks, Natick, MA). The 16 runs of the 2 sessions were preprocessed in a single stage including: (1) realignment of the images, (2) coregistration of the anatomical and mean functional images, (3) spatial normalization (estimating both the optimum 12-parameter affine transformation and nonlinear deformations) of all images to a standard stereotaxic space (MNI) with a voxel size of $2 \times 2 \times 2$ mm and (4) smoothing

the resulting images with an isotropic Gaussian kernel of 6 mm. Data from one subject were discarded because >10% of the volumes were corrupted, either because the signal strength varied >1.5% from mean value, or because scan-to-scan movement exceeds 0.5 mm per TR in any of the 6 realignment parameters (ArtRepair in SPM8). Because the frontal and lateral viewpoint videos were included in different runs (see above and Fig. 3), we created 2 separated general linear models (GLMs) for each subject to model the conditions of 8 runs. The design matrix of these GLMs was composed of 13 regressors: 7 modeling the conditions used (stereo- and mono-actions, stereo- and mono-static controls, stereo- and mono-dynamic controls, baseline), and 6 from the realignment process (3 translations and 3 rotations). All condition-specific regressors were convolved with the canonical hemodynamic response function.

We also included the disparity changes between frames of the videos as a variable of no interest, because they were not matched in stereo-actions and dynamic controls. Indeed, static and dynamic control conditions were intended to control for static and dynamic disparities, respectively. Hence, we evaluated how closely the action videos and control conditions were matched in these respects (Supplementary Figs 1–3), assessing disparity and change in disparity in the stimuli in the following way. For each frame (Supplementary Fig. 1A,D) a disparity map (Supplementary Fig. 1B,E) was calculated from the video image pairs, using the semi-global matching algorithm (Hirschmuller 2008) as implemented in the OpenCV-library (Bradski 2000). For each pixel in the left image, this computer-vision algorithm finds the corresponding pixel in the right image by minimizing a global cost function including intensity differences and global smoothness in the estimated disparity map. The change in disparity between successive frames of an action video or dynamic control was simply defined as the absolute value of differences in the disparity maps of the frames (Supplementary Fig. 1C,F). We defined a $7^\circ \times 8^\circ$ rectangle (see Supplementary Fig. 1C,F) and spatially averaged the disparities and disparity changes over the pixels within this rectangle, yielding a single change index for each video frame. This rectangle was fixed with respect to the video frame and included most hand trajectories for all fixation point positions. Hence it corresponds to the region of the display where disparities may differ between action videos and static controls and also where dynamic disparities should be non-zero. The disparity and change in disparity associated with each condition was then simply the average of all frames presented in that condition. Supplementary Figure 2 indicates that the dynamic disparities were much greater for the action videos than the dynamic controls. Hence, dynamic disparity had to be included as a variable of no interest. In contrast, the disparities of action videos and static controls was well-matched for both viewpoints (Supplementary Fig. 3) so here, disparity was not included as variable of no interest.

For both GLMs, we performed a 2-level random effect analysis (Holmes and Friston 1998). At the first level, each condition-specific regressor (2 actions, 4 controls, and one fixation condition) was contrasted with the implicit baseline, yielding 7 contrast images, keeping GLMs separate. The second level, performed on 23 subjects, comprised one model that included 14 regressors corresponding to the 7 contrast images from the first level for both frontal and lateral viewpoint GLMs. In this way we could compute main effect and interaction maps using all 3 factors (stimulus type, stereopsis, and viewpoint) in the experimental design. Main effect maps show activations due to any given factor independently of the others. We computed 4 such main-effect maps: one for stimulus type (positive for action), one for

stereopsis (positive for stereo) and 2 for viewpoint (one positive for frontal and one opposite for lateral). Moreover, in order to ensure that reported differences in activity cannot be explained by lower-order factors also present in control conditions, all main-effect T maps were obtained by taking the conjunction (conjunction null, Nichols et al. 2005) between 2 main effects obtained independently using static and dynamic controls. In addition, each main effect map was inclusively masked by the image contrasting all levels of that variable versus fixation at $P < 0.01$ uncorrected (e.g., for action main effect the mask was obtained by contrasting all action conditions vs. fixation condition). The sites of the main effects were subjected to small-volume correction using either the phAIP confidence ellipses of Jastorff et al. (2010) and Georgieva et al. (2009), or a precentral sulcus (PCS) ROI of 10 mm radius centered on coordinates from previous studies (Ferri, Peeters, et al. 2015; Ferri, Rizzolatti, et al. 2015). To evaluate whether one factor was dependent on the levels of the other factors, we also calculated 2- and 3-way interaction maps. As for the main effects, maps of all interactions were obtained by taking the conjunction of the static and dynamic control interactions and were inclusively masked with an image contrasting the condition with positive levels of the factors versus fixation at $P < 0.01$ (e.g., the 3-way interaction positive for frontal viewpoint, action and stereo, was masked with stereo-action-frontal vs. fixation). All maps were thresholded at $P < 0.001$ uncorrected level. We considered only clusters for which the size exceeded 10 voxels and the maxima reached $P < 0.05$ FWE corrected level, either in the whole brain or after small-volume correction. To summarize the effect sizes (i.e., parameter estimates) subtending the interactions, we applied the same GLM to region-of-interest averages identified in the SPM. Strictly speaking, this is redundant because the implicit GLM in the SPM analysis is simply being reproduced using linear mixtures of data. However, repeating the analysis provides useful regional summaries and is consistent with our subsequent ROI analysis of the 2 control experiments.

To evaluate how stereopsis in the action-frontal condition affected the effective connectivity of the regions significant in the 3-way interaction map, we applied a psycho-physiological interaction (PPI) analysis (Friston et al. 1997; Gitelman et al. 2003), using stereopsis as the experimental factor. To define seed regions in individual subjects, we contrasted the stereo-action-frontal condition with the average of its control conditions, and retained the voxels reaching $P < 0.05$ and having local maxima (LM) within 15 mm of the group LM. The threshold for the PPI was set at $P < 0.001$ and a cluster size of 10 voxels.

Both the SPMs and PPI maps were projected (enclosing voxel projection) onto flattened left and right hemispheres of the human PALS B12 atlas (Van Essen 2005, <http://sumsdb.wustl.edu:8081/sums/directory.do?id=636032>) using the Caret software package (Van Essen et al. 2001, <http://brainvis.wustl.edu/caret>). Profiles of 3-way interaction and action main effect sites were computed by averaging the voxels reaching $P < 0.001$ uncorrected in the contrast, with the masking described above. These profiles were computed independently for frontal and lateral runs in a split analysis, using 2 runs (one from each session) to define the ROI and 6 runs to compute activity in the different conditions. These profiles were analyzed with analysis of variances (ANOVAs) and Fishers least significant difference was used as a post hoc test.

Second Axis Experiment

Data analysis was performed using SPM8. The preprocessing of the 8 runs included the same steps as in the main experiment (see above). The design matrix of GLM was composed of 11

regressors: 5 modeling the conditions used (Frontal_Z, Frontal_X, Lateral_Z, Lateral_X, fixation) and 6 from the realignment process (3 translations and 3 rotations). We did not include disparity change as a variable of no interest, since the experimental design did not make use of control conditions.

The statistical analysis was restricted to the four 3-way interaction sites from the first main experiment. The data of the main experiment were used to define ROIs in individual subjects. First the interaction positive for action and stereo, was computed for each subject, using only the frontal-viewpoint runs. We used this contrast because the 3-way interaction could not be computed at the single subject level since the 2-viewpoint levels were in different runs. Second, we looked for the individual LM in the interaction nearest to the four 3-way interaction sites obtained in the main experiment. As for the PPI analysis, we accepted only local maxima within 15 mm of the original group site. Finally, we defined 27-voxel ROIs around the individual maxima.

Third Discrimination Experiment

Data analysis was performed using SPM8 and Matlab. The preprocessing of the 10 runs included the same steps as in the main experiment (see above), except for smoothing, since the

analysis was performed at the voxel level (Serences et al. 2009). The design matrix of the GLM was composed of 15 regressors: 9 modeling the conditions (all frontal: stereo-and mono-grasping, stereo-and mono-pushing, stereo-and mono-dropping, stereo-and mono-displacing, fixation) and 6 taken from the realignment process. We did not include disparity change as a variable of no interest, since the experimental design did not include control conditions.

Using a novel approach, we incorporated the sensitivity index or d' from signal detection theory (Green and Swets 1966) to investigate the representation of observed action at the single-voxel level. In signal detection theory, d' provides a measure of the separation between the means of the signal (S) and noise (N) distributions, compared against the standard deviation of the distributions: $d' = (\text{mean S} - \text{mean N}) / \text{square root} [(SD_S^2 + SD_N^2) / 2]$ (Green and Swets 1966; Solis and Doupe 1997; Theunissen and Doupe 1998; Person and Perkel 2007). Here we applied the same statistic to each voxel, comparing one observed action, considered as a signal (S), to another observed action, considered as noise (N) in the following way. For each voxel we calculated the d' values for the 6 possible pairs comparing 4 observed actions, using either the stereo-actions or the mono-actions (Fig. 4).

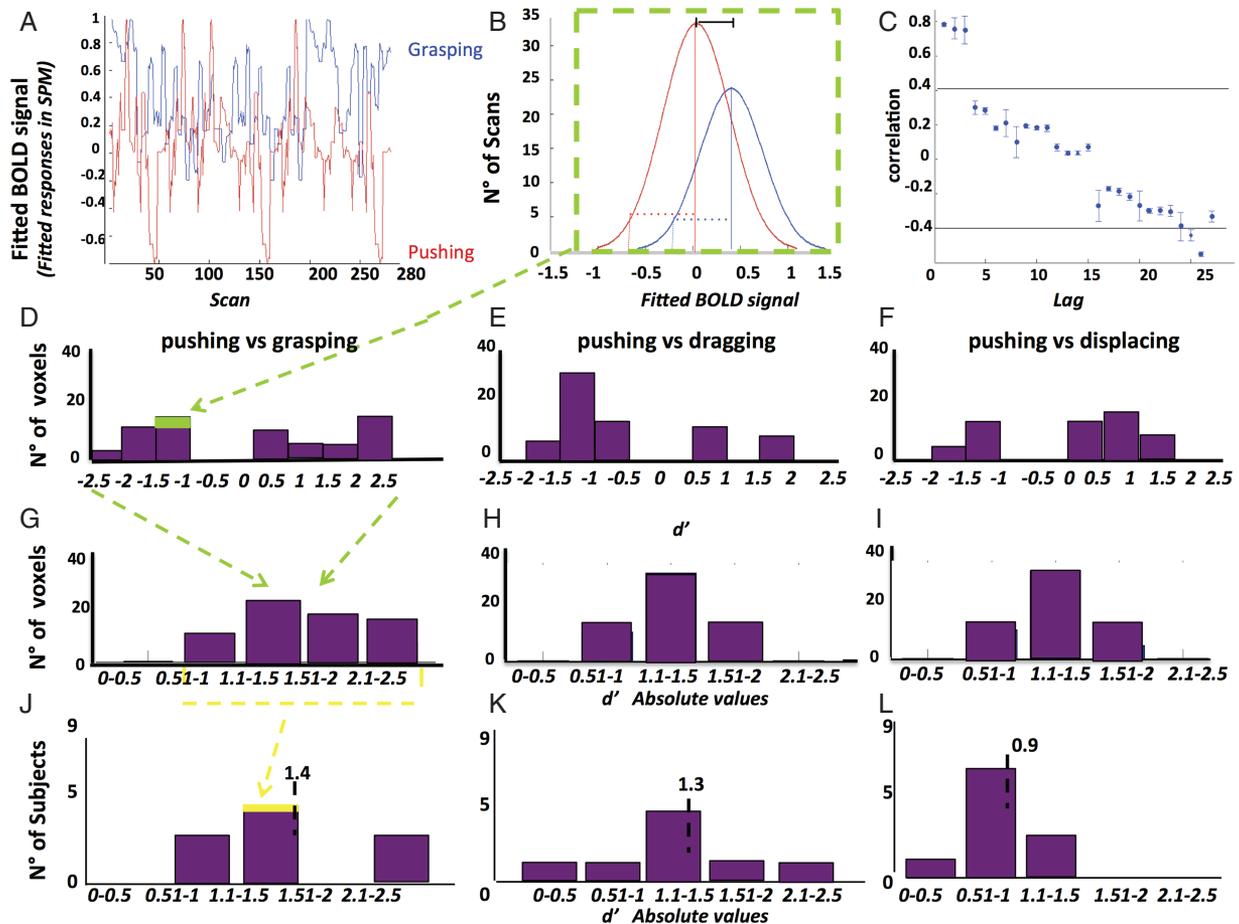


Figure 4. Definition of d' . (A,C) single voxel (in PCG of subject 1, stereo conditions): (A) fitted BOLD activity of one voxel for observing pushing (red) and grasping (blue) as a function of scan number in the sessions (concatenated); (B) distributions of fitted BOLD of the voxel for the 2 actions (pushing and grasping) compared in the d' which compares difference of means (full line) to the standard deviations (stippled line); (C) mean (over runs) autocorrelation function of the voxel for grasping compared with the confidence limits for zero correlation. This function is used to correct the standard deviation in B. Notice that the lag takes different values in ms as scans for a given action were blocked together in miniblocks of 4 TRs. Vertical lines: SE, horizontal lines: mean confidence limits (D–L) PCG ROI in stereo conditions. (D–I): distribution of d' (D–F) and its absolute value (G–I) in subject 1 for 3 pairs: pushing–grasping (C,F), pushing–dragging (D,G), and pushing–displacing (E,H); (I–K) distribution of mean d' in PCG for 9 subjects: pushing–grasping (I), pushing–dragging (J), and pushing–displacing (K). (J–L) The average d' across the 9 subjects (value indicated by vertical dashed line and number) is used in Figure 9.

Each run yielded 28 fitted Bold values (fitted responses in SPM) corresponding to the 28 scans taken during the presentation of a single action. Because 10 runs were sampled, the d' of each voxel for an action pair was based on 280 samples (Fig. 4A,B). As the d' calculation assumes that the samples are independent, we tested the autocorrelation in the data points from single runs, which differed significantly from zero. Hence the computation of the standard deviations entering the d' calculus was broken down in a within-run part, which was corrected using the mean (over 10 runs) autocorrelation coefficients for each action and voxel (Fig. 4C), and a between-run component. These d' values could assume positive and negative signs, depending on the levels of activity evoked by the 2 observed actions (Fig. 4D–F). Since the distributions in a ROI were symmetrical, and we were interested only in differences in activity, we took the absolute value of the d' before averaging over voxels in a ROI (Fig. 4G–I). The distribution of this average d' across subjects is summarized by an inter-subject mean (Fig. 4J–L).

At the suggestion of an anonymous reviewer, we have run a canonical variate analysis (CVA), an established multivariate approach (Friston et al. 1995) implemented in SPM8, to test differences between action exemplars. The d' calculation and the CVA have some similarities, insofar as both are related to linear discriminant analysis, and in the opinion of the reviewer, CVA might replace the d' analysis. The preprocessing and the GLM design were the same as for the d' computation, to make the analyses as comparable as possible. CVA was carried out at the single-subject level using 6 contrasts, comparing the action exemplars pairwise, but only in the stereo conditions (stereo-grasping vs. stereo-dragging, stereo grasping vs. stereo pushing, stereo grasping vs. stereo displacing, stereo pushing vs. stereo dragging, stereo pushing vs. stereo displacing, stereo dragging vs. stereo displacing). Thus the 6 contrasts are the same as those used to compute the d' values. A single volume of 220 voxels, corresponding to the voxels of the 4 stereo-action-specific (SAS) regions, was used as a mask. CVA uses the generalized eigenvalue solution to the treatment (here contrasts between actions) and the residual sum of squares and products of a GLM. The eigenvalues have a χ^2 distribution, which allows testing the null hypothesis that the mapping is of D or greater dimensionality. These 2 results (dimensionality, and χ^2) were tabulated for each subject and action pair. Since we use 6 contrasts for each of the 9 subjects, we correct for 6 comparisons for each individual. If the χ^2 did not reach the

significance level in the first analysis, CVA was repeated a second time using each SAS region separately as a mask.

Results

Behavioral Data: Fixation

In the factorial main experiment, subjects ($n = 23$) averaged 7.7 saccades/min ($SD = 1.4$) in the frontal runs, and 8.6 saccades/min ($SD = 0.8$) in the lateral runs, with no significant differences among conditions ($F_{6,15} = 0.8$; $P > 0.8$, $F_{6,15} = 0.4$; $P > 0.9$, frontal and lateral, respectively). The standard deviation of binocular differences in horizontal position, providing a measure of vergence eye movements (see Materials and Methods), averaged 0.87° and 0.76° in frontal and lateral runs respectively, with no significant differences across conditions ($F_{6,15} = 0.08$; $P > 0.9$; $F_{6,15} = 0.05$; $P > 0.9$).

Subjects ($n = 16$) fixated well during the 8 runs of the axis control session. The average across subjects was 9.3 saccades/min ($SD = 1.2$), without any significant differences across conditions ($F_{4,10} = 0.3$; $P > 0.9$). The standard deviation of binocular position differences, estimating vergence eye movements (see methods), averaged 0.53° with no significant differences across conditions ($F_{4,10} = 0.08$; $P > 0.9$). All subjects ($n = 9$) also fixated well during the 10 runs of the discrimination control session, averaging 10.1 saccades/min ($SD = 1.1$), with no significant differences across conditions ($F_{8,1} = 0.2$; $P > 0.9$). The standard deviation of binocular position differences averaged 0.13° with no significant differences across conditions ($F_{12,1} = 0.07$; $P > 0.9$).

Random Effect Analysis of Factorial Main Experiment

The principal aim of this study was to investigate the contributions of stereopsis and the observer's point of view to the processing of observed manipulative actions. Hence, the first main experiment followed a $3 \times 2 \times 2$ design (Figs 1 and 3) with factors stimulus type, stereopsis, and viewpoint. Table 1 and Supplementary Table 1 list the local maxima and sizes of the sites reaching significance level ($P < 0.001$ uncorrected) in the random-effect analysis. Most of the sites yielding a main effect or interaction, however, reached FWE significance in the whole brain (bold) or after SVC (small volume correction, underline). This was particularly true for the main effect of action and the 3-way interaction

Table 1 Coordinates (x, y, z) of local maxima and cluster size (number of voxels at $P < 0.001$) of sites yielding significant (corrected) main effect of action and 3-way interaction in the first main experiment:

Contrast	Action main effect action > control				Three-way interaction Stereo \times action \times frontal	
	RH		LH		LH	
	x	y z	x	y z	x y z	Size
1 pITS–pMTG	-40	-68 6	44	-64 4		
2 DIPSM					-22 -54 54	52
3 phAIP–DIPSA			<u>36</u>	<u>-38</u> <u>52</u>		<u>40</u>
4 phAIP	-34	-42 48				
5 PFop					-56 -32 20	119
6 PFcm posterior	-50	-40 24				
7 dorsal PCS	<u>-26</u>	<u>-14</u> <u>56</u>				
8 PCG					-42 -12 48	49

Note: Bold characters represent FWE corrected level (whole brain) at $P < 0.05$; underlined characters represent FWE corrected level at $P < 0.05$ after small-volume correction. pITS, posterior inferior temporal sulcus; pMTG, posterior middle temporal gyrus; DIPSM, dorsal intraparietal sulcus Medial; phAIP, putative human anterior intraparietal; DIPSA, dorsal intraparietal sulcus anterior; preCS, precentral sulcus; PCG, precentral gyrus.

(Table 1). Since most of the sites yielded by the main effect of viewpoint and the 2-way interactions (Supplementary Table 1) were also present in the 3-way interaction, only sites displaying FWE significance for the latter interaction and for the action main effect are described further.

Main Effect of Action

The action main effect contrasted the action conditions with the control conditions (static and dynamic), independently of viewpoint and disparity. As Figure 5A–D shows, the main effect of action resulted in activations in both hemispheres, but was biased toward the left hemisphere in parietal and premotor cortex, typical of action–observation studies (Buccino et al. 2001; Wheaton et al. 2004; Jastorff et al. 2010; Ferri, Rizzolatti, et al. 2015).

In the left hemisphere, activation was present at all 3 levels of the action observation’s network: occipitotemporal, posterior-parietal, and premotor (Fig 5A–C). The lateral occipitotemporal cortex (LOT) site (367 voxels) began at the posterior part of inferior temporal sulcus (pITS), corresponding to the retinotopic areas MT and pMSTv (Abdollahi et al. 2014) that are part of the human MT/V5 cluster (Kolster et al. 2010), and extended to the posterior part of middle temporal gyrus (MTG). The local maximum ($-40 -68 6$; $t = 5.56$; $P < 0.05$ FWE corrected), was located next to the rostral border of area MT/V5 (1 in Fig. 5C). Because EBA overlaps with the MT cluster (Ferri et al. 2013), the posterior

part of the LOTC activation overlapped EBA. The cluster (4 in Fig. 5C) in rostral intraparietal sulcus (IPS) corresponded to the dorsal part of phAIP (Jastorff et al. 2010), considered the homolog of the anterior part of monkey AIP (Vanduffel et al. 2014). Its local maximum ($-34 -42 48$; $t = 3.64$) reached the $P < 0.05$ FWE corrected level with small-volume correction using the phAIP ellipse, defined by Jastorff et al. (2010) and Georgieva et al. (2009), as a priori ROI. It was also close to the average coordinates ($-31, -41, 56$) of the human parietal grasp region of Konen et al. (2013). The main effect also activated another parietal site ($-50 -40 24$; $t = 5.27$, $P < 0.05$ FWE corrected) located in the caudo-dorsal end of the sylvian sulcus. Based on the inferior parietal lobule (IPL) parcellation proposed by Caspers et al. (2006), it corresponds to the dorsal part of PFcm (6 in Fig. 5C). The premotor activation (7 in Fig. 5C) was located in the dorsal part of the left PCS, at the level of the fourth–sixth mini-ROIs of Jastorff et al. (2010). Its local maximum ($-26 -14 56$; $t = 3.9$ $P < 0.001$ uncorrected) reached the $P < 0.05$ FWE significance level after small-volume correction using the premotor site from the manipulation activation map (Ferri, Rizzolatti, et al. 2015).

The right hemisphere was activated only in occipitotemporal and posterior parietal regions, symmetrically to those sites on the left (Fig. 5D). The parietal site (36, 38, 52, $t = 3.9$), straddling phAIP and DIPSA (3 in Fig. 5D), the homologs of anterior and posterior parts of AIP (Vanduffel et al. 2014) reached FWE

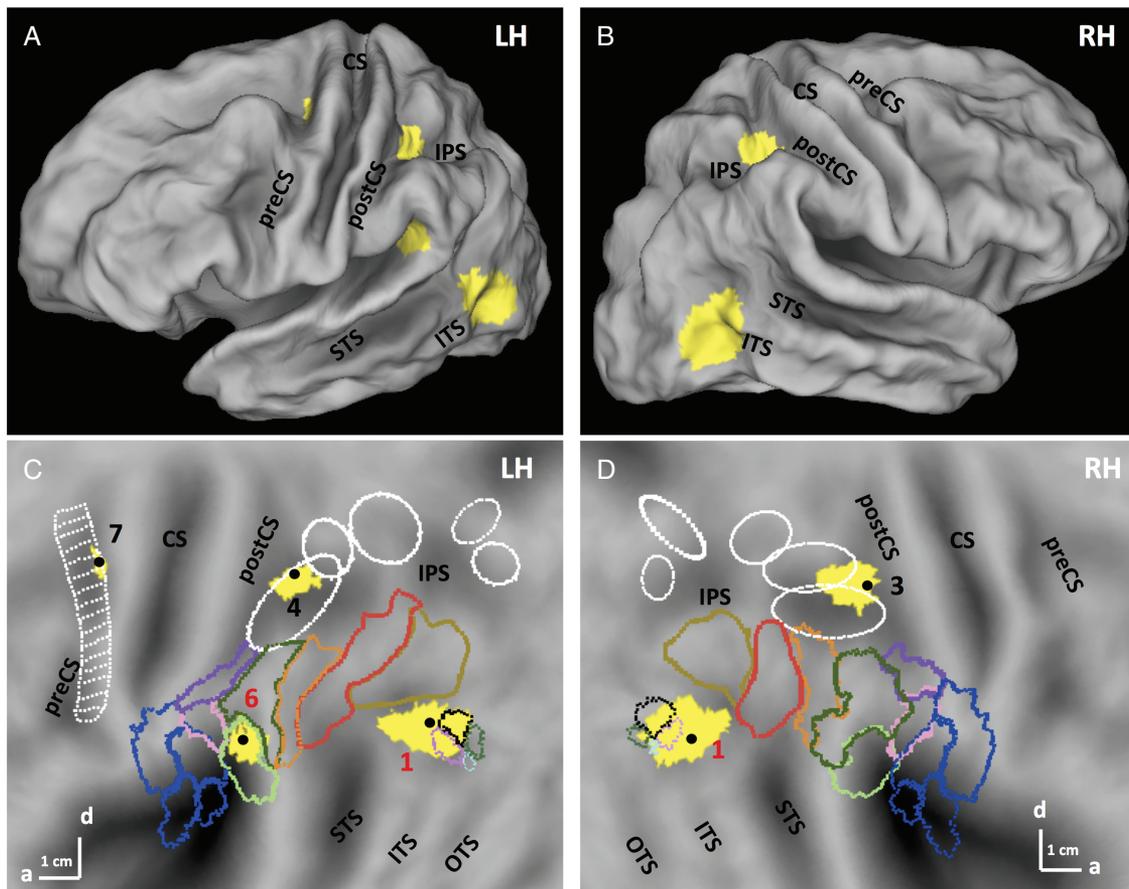


Figure 5. SPMs showing the significant (LM FWE corrected $P < 0.05$) sites (yellow) for the action main effect of the first experiment in left and right hemisphere on folded brain (A,B) and flat maps (C,D). Colored outlines correspond to cytoarchitectonically defined regions: opercular regions: blue (Eickhoff et al. 2006), and IPL regions: other colors (pink: PFop, light green: PFcm, Caspers et al. 2006). White ellipses are confidence ellipses for phAIP, DIPSA, DIPSM, POIPS, and VIPS (from rostral to caudal) and white ladder-like outline: premotor mini-ROIs, from Jastorff et al. (2010). Colored outlines on the post ITS indicate the MT cluster (black: MT, purple: pMSTv, green pV4t, light blue: pFST, Abdollahi et al. 2014). Numbers: see Table 1. preCS, precentral sulcus, CS, central sulcus, postCS, postcentral sulcus, IPS, intraparietal sulcus, STS, superior temporal sulcus, ITS, inferior temporal sulcus, OTS, occipitotemporal sulcus.

$P < 0.05$ after SVC using the phAIP ellipse (Jastorff et al. 2010), as a priori ROI.

The main effect of action thus corresponds rather well to the classical manipulative-action observation network described by Jastorff et al. (2010), Abdollahi et al. (2013), and Ferri, Rizzolatti, et al. (2015). The main differences are the absence here of the second LOTC activation site typically found in OTS, and the presence of a PF activation not reported in those earlier studies.

Three-way Interaction

In the present study, the 3-way interaction taken in the direction action > controls, stereo > mono and frontal > lateral yielded 3 significant activation sites, all of which reached FWE-corrected level (Table 2 and Fig. 6A). The interaction in the opposite directions yielded no activations. The first interaction site (8 in Fig. 6A, $-42, -12, 48, t = 5.7; P < 0.01$ FWE corrected) was located in the precentral gyrus (PCG), adjacent and ventral to the action main effect site. While the latter site overlapped mini-ROIs 4 to 6 of Jastorff et al. (2010), the interaction site was located caudal to mini-ROI

9. The second site (2 in Fig. 6A, $-22 -54 54, t = 5.2; P < 0.05$ FWE corrected) was located in the medial-dorsal part of IPS, inside the dorsal IPS medial (DIPSM) motion-sensitive region (Sunaert et al. 1999), and was more caudal than the phAIP action main effect site. The profiles of the interaction sites (Fig. 7) indicate that, basically, they were activated in only one of the 12 conditions: stereo-actions observed in frontal view. This is strikingly different from the profiles of the action main effect regions (Supplementary Fig. 4). The profiles of the premotor and parietal sites were relatively similar (Fig. 7A,B), although the DIPSM profile showed weak responses to mono-action in lateral view, corresponding to its location on the edge of the activation map for observing manipulation in the Ferri, Rizzolatti, et al. (2015) study.

The third site (5 in Fig. 6A) was located at the level of Sylvian fissure, in retro-insular (RI) cortex, considered vestibular cortex in humans (Lopez et al. 2012). This region extended from cytoarchitectonic area PPop (Caspers et al. 2006) rostro-dorsally to the ventral part of PFcm (Caspers et al. 2006). Its local maximum, $-56 -32 20$, located in PPop, reached the FWE corrected level ($t = 5.6$;

Table 2. PPI of stereo-action-specific regions: significant ($P < 0.001$ uncorrected, cluster size 10) sites (all in left hemisphere)

Seeds	PCG ($-42 -12 48$)	DIPSM ($-22 -54 54$)	dorsal RI ($-56 -32 20$)	ventral RI ($-40 -38 20$)
<i>Targets</i>				
1. Lingual gyrus	$-24 -72 -4$			$-24 -70 -6$
2. <u>MT cluster</u>	$-32 -74 -2$			
3. <u>post STS</u>	$-54 -54 14$			
4. DIPSM	$-24 -60 52$		$-26 -62 54$	
5. <u>phAIP</u>	$-38 -52 40$			
6. ant PF		$-64 -30 36$		
7. dorsal RI	$-56 -38 20$			
8. ventral RI		$-44 -36 22$		
9. PCG		$-40 -6 48$	$-40 -2 48$	$-40 -4 50$
10. BA 44			$-54 18 10$	$-56 18 8$
11. ant IFS		$-30 42 8$		

Note: Italicized characters represent SAS region; underlined characters represent action main effect region. STS, superior temporal sulcus, IFS, inferior frontal sulcus.

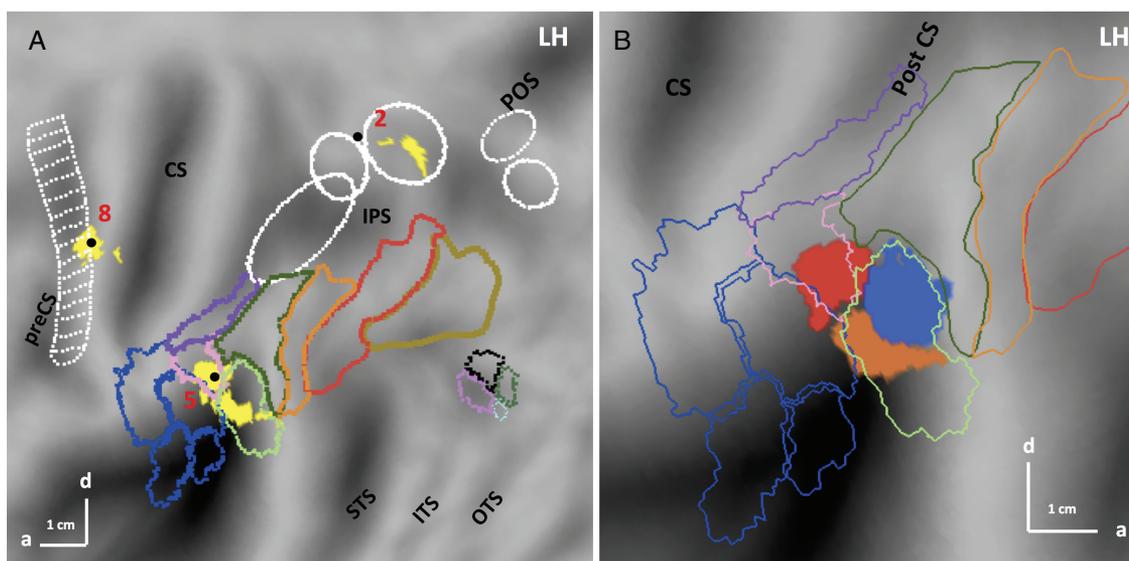


Figure 6. (A) SPM showing the significant (LM $P < 0.05$ FWE corrected) sites (yellow) for the 3-way interaction between viewpoint (frontal = positive), stereopsis (stereo = positive), and stimulus type (action = positive) in the first experiment on flatmap of left hemisphere; (B) enlargement of left flatmap showing dorsal (red), ventral (orange) and posterior (blue) RI regions. Numbers see Table 1. Same conventions as Figure 5.

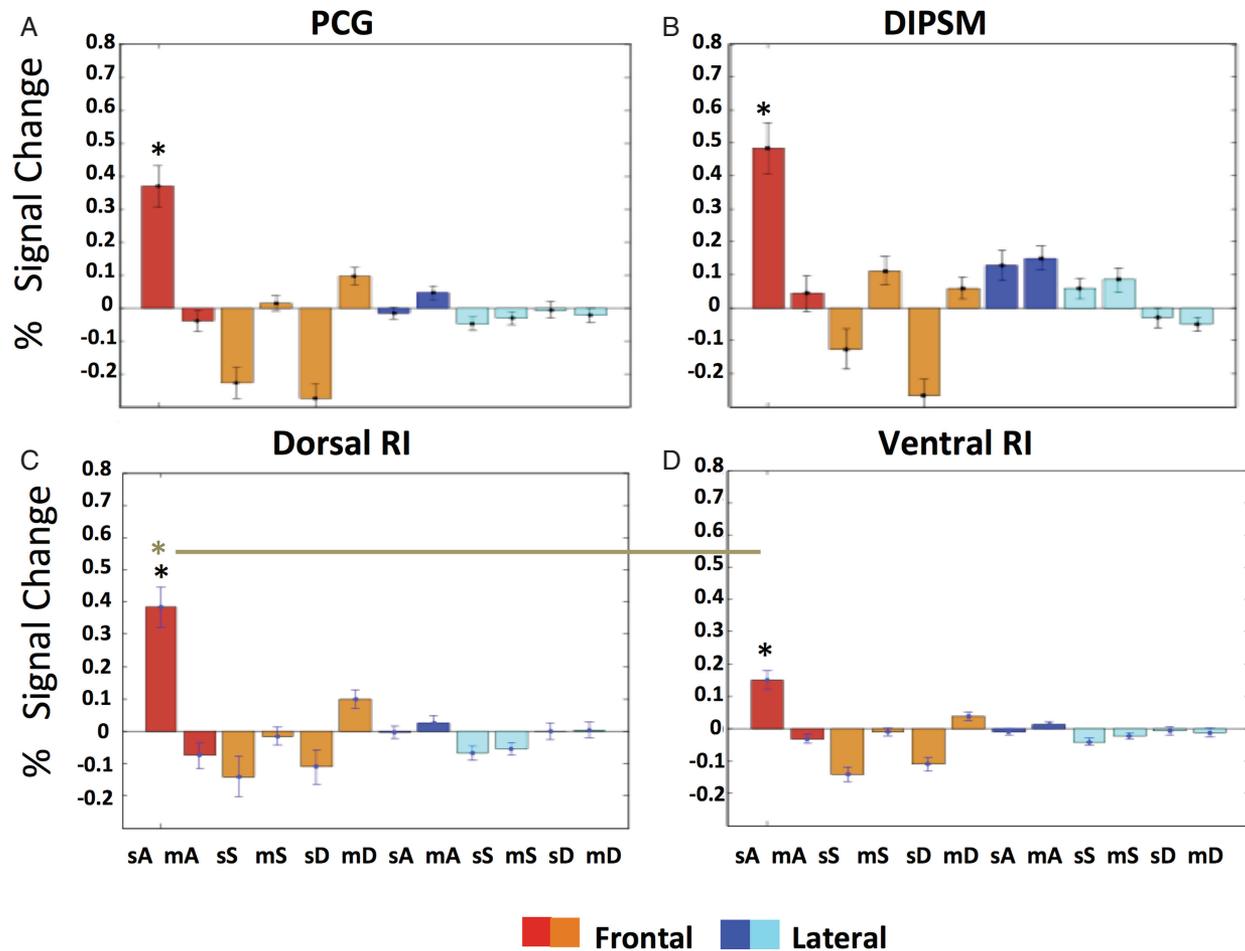


Figure 7. Activity profiles of stereo-action-specific regions: PCG (A), DIPSM (B), dorsal (C), and ventral RI (D) in the first experiment. Color code indicates viewpoint. A: action (red, dark blue), S: static control (orange, light blue), D: dynamic control (orange, light blue). Since these profiles were either obtained by split analysis (see Materials and Methods) or were computed for separate anatomical entities, they confirm the random effect analysis. Three-way ANOVA's for individual profiles: PCG. main effect VP: $F_{1,22} = 0.02$ $P > 0.9$; main effect STIT: $F_{2,44} = 26.9$ $P < 0.001$; main effect STE: $F_{1,22} = 16.5$ $P < 0.001$; VP \times STIT: $F_{2,44} = 14.9$ $P < 0.001$; VP \times STE: $F_{1,22} = 7.4$ $P < 0.01$; STIT \times STE: $F_{2,44} = 28.5$ $P < 0.001$; VP \times STIT \times STE: $F_{2,44} = 33.7$ $P < 0.001$. DIPSM. main effect VP: $F_{1,22} = 0.05$ $P > 0.8$; main effect STIT: $F_{2,44} = 31.1$ $P < 0.001$; main effect STE: $F_{1,22} = 4.5$ $P < 0.05$; VP \times STIT: $F_{2,44} = 10.4$ $P < 0.001$; VP \times STE: $F_{1,22} = 1.8$ $P > 0.2$; STIT \times STE: $F_{2,44} = 26.9$ $P < 0.001$; VP \times STIT \times STE: $F_{2,44} = 27.4$ $P < 0.001$. Dorsal RI. main effect VP: $F_{1,22} = 0.5$ $P > 0.5$; main effect STIT: $F_{2,44} = 31.8$ $P < 0.001$; main effect STE: $F_{1,22} = 15.7$ $P < 0.001$; VP \times STIT: $F_{2,44} = 12.7$ $P < 0.001$; VP \times STE: $F_{1,22} = 9$ $P < 0.01$; STIT \times STE: $F_{2,44} = 28.3$ $P < 0.001$; VP \times STIT \times STE: $F_{2,44} = 28.9$ $P < 0.001$. Ventral RI. main effect VP: $F_{1,22} = 0.3$ $P > 0.6$; main effect STIT: $F_{2,44} = 23.3$ $P < 0.001$; main effect STE: $F_{1,22} = 22.8$ $P < 0.001$; VP \times STIT: $F_{2,44} = 11.7$ $P < 0.001$; VP \times STE: $F_{1,22} = 8.4$ $P < 0.01$; STIT \times STE: $F_{2,44} = 26.9$ $P < 0.001$; VP \times STIT \times STE: $F_{2,44} = 30.6$ $P < 0.001$. Black asterisks indicate that stereo-action frontal condition differed from all other conditions in post hoc tests. Two-way ANOVA comparing the 2 RI profiles: 2ROIs (dorsal RI, ventral RI) \times 12 Conditions. main effect ROI: $F_{1,22} = 0.75$ $P > 0.3$; main effect condition: $F_{11,242} = 22.7$ $P < 0.001$; ROI \times condition: $F_{11,242} = 18.7$ $P < 0.001$. Brown asterisk and line indicate that the stereo-action frontal condition differed between the dorsal and ventral RI in post hoc test. STIT, stimulus type; STE, stereopsis, VP, viewpoint; m, mono; s, stereo, A, action, S, static, D, dynamic.

$P < 0.01$ FWE corrected). This activation is located in RI cortex, but rostro-ventral to the action main effect site. Figure 6B shows that the 2 functional sites did not overlap.

Since the RI interaction site was almost evenly divided between 2 distinct anatomical areas, we considered splitting it into 2 distinct functional ROIs: PPop, designated dorsal RI, with the maximum located at $-56 -32 20$, and a size of 70 voxels, and PFCm, designated ventral RI, with a maximum located at $-40 -38 20$ ($t = 5.3$; $P < 0.05$ FWE corrected) and comprising 49 voxels. Their profiles, shown in Figure 7C,D, were significantly different. Using statistical analysis to confirm the partitioning of the RI cluster, we applied a 2 (ROI) \times 12 (condition) ANOVA to the profiles of dorsal and ventral RI regions. The 2-way interaction was highly significant ($F_{1,242} = 18.7$; $P < 0.001$); moreover a post hoc test showed that the response in the stereo-action-frontal condition was stronger in the dorsal part than in the ventral ($P < 0.001$). This means that this partition divides the RI interaction site into functionally different regions, due mainly to the activation

by stereo-action viewed frontally being stronger in the dorsal portion. Thus we tentatively considered the 2 functional activations seen in RI as distinct sectors, themselves different from the action main effect site, which we designate posterior RI (Fig. 6B).

In summary, the main effect of action revealed, as expected, occipitotemporal, parietal, and premotor regions. The 3-way interaction revealed 4 sites: PCG, DIPSM, dorsal, and ventral Retro-Insula, which are specific for the processing of observed manipulative stereo-actions seen from a frontal perspective. We refer to these sites as stereo-action-specific (SAS) sites.

Further Characterization of the Stereo-Action-specific Regions

Second Axis Experiment

The first main experiment has shown that the 4 stereo-action-specific regions were more active for the frontal viewpoint. As noted earlier, 3 of the 4 manipulative action exemplars used

(grasping, dragging pushing) moved along the z-axis of the actor's peripersonal space. Since this axis is orthogonal to the FP plane of the observer in the frontal but not the lateral view (Fig. 2), viewpoint specificity was difficult to interpret, as viewpoint and OFPP direction were confounded. Therefore, in the second axis experiment we presented the 4 action exemplars of the first main experiment, but only in stereo, moving either along the z or x axis of the actor, and viewed either frontally or laterally in a 2×2 design (Fig. 2). If the frontal viewpoint effect in the first main experiment was due only to the axis of motion being OFPP, we should observe an interaction between the factors axis of motion and viewpoint.

Figure 8 shows the results of the ROI analysis in 16 subjects, with ROIs defined by the 3-way interaction in the main experiment. As per our prediction, the activity in all 4 SAS regions showed an interaction between axis of motion and viewpoint: the activity was stronger for frontal viewpoint when actions were directed along the z axis of the actor; the opposite proved true for the lateral viewpoint. This was confirmed by significant interactions in the four 2-way ANOVAs (all $P < 0.01$, see legend Fig. 8). However, the profiles of the 4 areas also showed some differences: the activity in PCG is stronger for frontally-viewed Z axis, while the 2 RI regions show the opposite trend, with the laterally viewed X axis evoking somewhat more activity and the DIPSM site showing a nearly perfect reversal. The 3-way interactions computed in a 4 (ROIs) \times 2 (viewpoints) \times 2 (directions of action) ANOVA reached significance ($F_{3,45} = 3.1$ $P < 0.05$), showing that the 4 ROIs were functionally different. Moreover, a post hoc test ($P < 0.001$, asterisk in Fig. 8) showed that the activity in PCG evoked by frontally viewed Z axis was the strongest of the 16 conditions. These results show that in all SAS regions, the stereo effect occurred only for OFPP actions, as predicted by physics. In

addition, however, left premotor cortex exhibited a frontal viewpoint preference: its activity was boosted when stereo-actions were observed from the frontal perspective.

Single-voxel Analysis of the Discrimination Experiment

In the first 2 experiments, results were obtained with classical univariate methods based on the logic that only suprathreshold activity can detect an underlying neural process. Although very useful, these methods may be too coarse for testing whether the increased activity in the stereo-action-specific regions corresponded to more accurate representations of the actions, as one would expect from theory. To address this question, we investigated activity in single voxels and tested how well they could distinguish between 2 observed actions. In the discrimination control experiment we presented the stereo and mono-conditions of the 4 action exemplars (from the previous 2 experiments) separately. We then introduced a new statistic, the d' , based on a comparison between means and standard deviations of fitted BOLD activity evoked by 2 action-conditions at the voxel level (see Fig. 4). This statistic, used as sensitivity metric in perceptual studies, is mathematically a distance and here it can be thought of as a neural distance, measuring differences in MR activity between conditions. Hence, it allowed us to test whether or not single voxels discriminated better between 2 observed actions in stereo- than in mono-conditions.

The top row of Figure 9 shows the activity profiles of the 4 SAS regions for the third discrimination experiment. As expected from the previous 2 experiments, activation by the observation of manipulation exemplars is stronger in stereo-conditions than in mono. Repeated-measures ANOVA yielded a significant main effect of stereopsis after correction for 4 comparisons in PCG ($F_{1,8} = 10.9$, $P < 0.01$) and the DIPSM sector ($F_{1,8} = 16.9$,

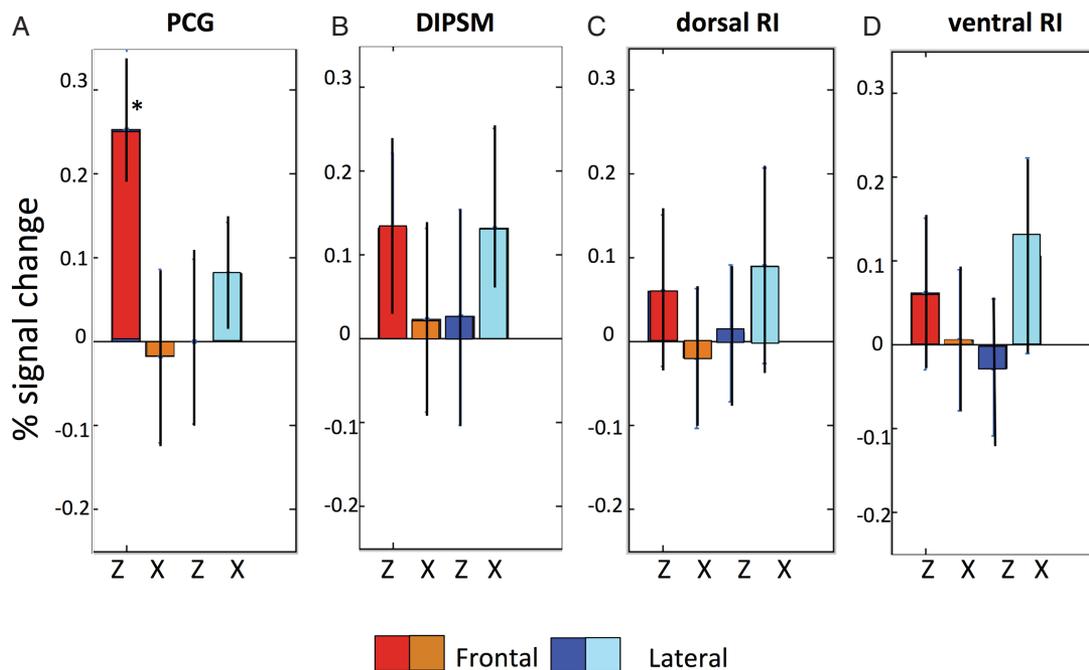


Figure 8. Activity profiles of the 4 stereo-action-specific regions in the second axis experiment. Same conventions as Figure 7. Two-way ANOVAs on individual profiles: PCG. Main effect VP: $F_{1,15} = 19.8$ $P < 0.001$; main effect AX: $F_{1,15} = 20.1$ $P < 0.001$; VP \times AX: $F_{1,15} = 27.7$ $P < 0.001$. DIPSM. Main effect VP: $F_{1,15} = 0.007$ $P > 0.9$; main effect AX: $F_{1,15} = 0.008$ $P > 0.9$; VP \times AX: $F_{1,15} = 8.3$ $P < 0.01$. Dorsal RI. Main effect VP: $F_{1,15} = 1.4$ $P > 0.2$; main effect AX: $F_{1,15} = 0.08$ $P > 0.9$; VP \times AX: $F_{1,15} = 7.8$ $P < 0.01$. Ventral RI. Main effect VP: $F_{1,15} = 0.2$ $P > 0.6$; main effect AX: $F_{1,15} = 3.9$ $P > 0.07$; VP \times AX: $F_{1,15} = 8.2$ $P < 0.01$. Three-way ANOVA comparing profiles: main effect ROI: $F_{3,45} = 1.4$ $P > 0.2$; main effect VP: $F_{1,15} = 1.3$ $P > 0.2$; main effect AX: $F_{1,15} = 1.1$ $P > 0.3$; ROI \times VP: $F_{3,45} = 9.8$; $P < 0.001$; ROI \times AX: $F_{3,45} = 15.8$ $P < 0.001$; VP \times AX: $F_{1,15} = 18.9$ $P < 0.001$; ROI \times VP \times AX: $F_{3,45} = 3.1$ $P < 0.05$. Red asterisk indicates that the frontal Z condition in PCG differs from the same condition in the 3 other ROIs. VP, viewpoint, AX, axis.

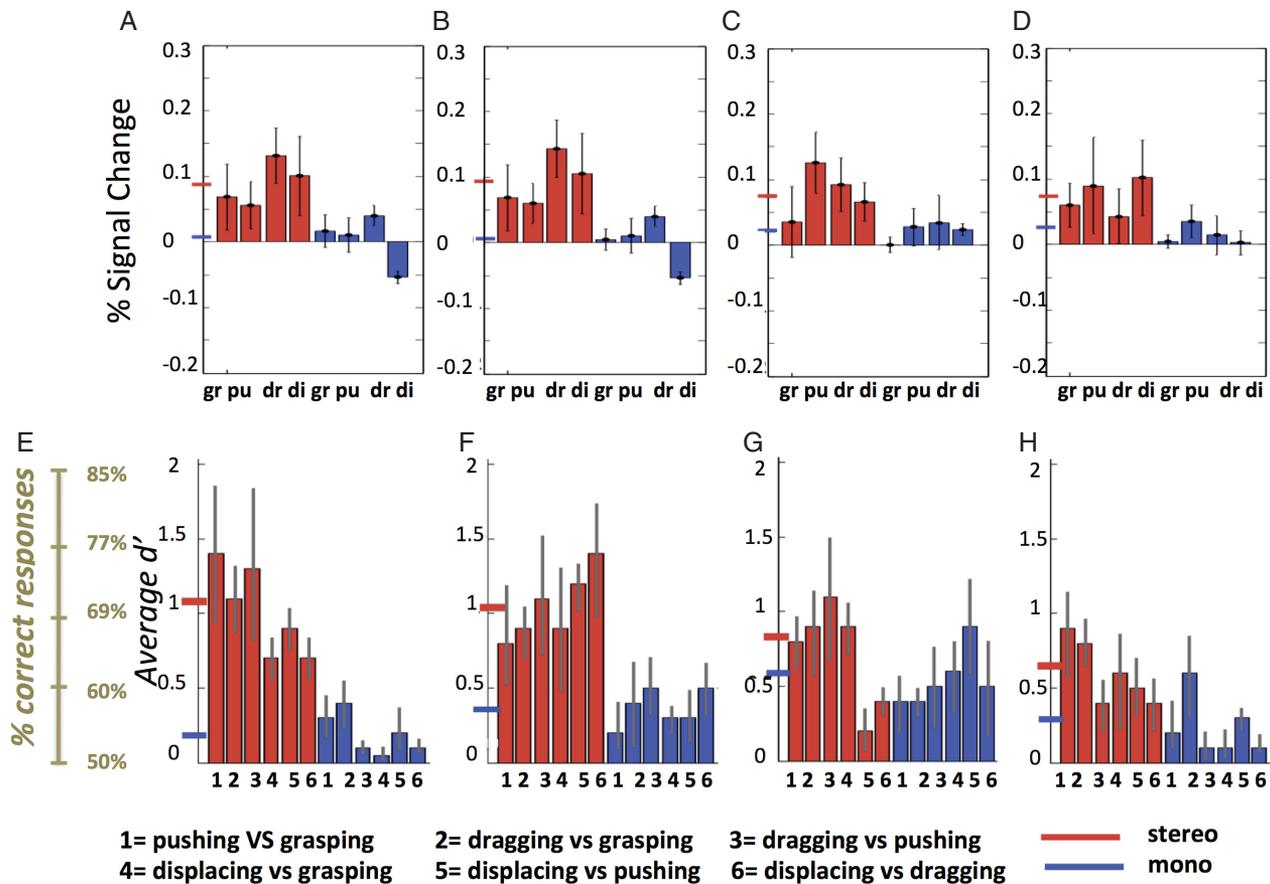


Figure 9. Activity profiles (A–D) and d' (E–H) for stereo- (red) and mono- (blue) actions of the stereo-action-specific regions (left PCG: A,E; left DIPSM: B,F; left dorsal RI: C,G; and left ventral RI: D,H) across the 9 subjects in the discrimination experiment. Means and SE are indicated, as well as corresponding % correct (for unbiased observers) in left lower corner. gr, grasping, pu, pushing, dr, dragging, di, displacing. Short horizontal lines on y axis: grand averages for stereo (red) and mono-conditions (blue). Repeated-measure ANOVAs for d' distributions: main effect of stereopsis: $F_{1,18} = 9.1$ ($P < 0.01$) in PCG, $F_{1,18} = 8.8$ ($P < 0.01$) in DIPSM sector, $F_{1,18} = 1.9$ ($P > 0.15$) in dorsal RI and $F_{1,18} = 1.4$ ($P > 0.45$) in ventral RI. Main effect of action pairs and interactions all $P > 0.3$. The distribution across subjects of d' values for pairs 1, 3, and 5 in PCG are shown in Figure 4J–L. The lowest d' values across subjects and action pairs equaled 0.45 in PCG and 0.43 in DIPSM.

$P < 0.005$). Other main effects and interactions were not significant. Six d' values could be calculated between the 4 exemplars tested and are shown in the bottom row of Figure 9. To a greater extent than activity levels, the d' increased in stereo-conditions compared with mono. Again the increase was larger in PCG and the DIPSM sector. Indeed in the repeated-measure ANOVAs, the main effects of stereopsis were significant after correction for 4 comparisons only for these 2 sites (see Fig. legend). The grand average d' values in the stereo-conditions ranged from 0.6 to 0.7 in the RI sites to 1 in PCG and the DIPSM sector. The latter value corresponds in an unbiased observer to 70% correct. Recent psychophysical studies (Platonov and Orban 2015) indicate that human observers are indeed unbiased with regard to action discrimination, thus the representation of observed manipulative stereo-actions is quite accurate. The main effect of action-pair and the interactions were all nonsignificant. The d' values for the different action pairs are relatively similar in PCG or the DIPSM sector, ranging between 0.7 and 1.4. Thus, the accuracy of action-exemplar representation in these 2 regions not only improved with stereo, but was also relatively homogeneous. The effects of stereopsis on the main effect regions were very different (Supplementary Fig. 5). The activity profiles show that the mean activity of the ROIs changed little, as expected from the first

experiment. On the other hand, the d' values generally increased slightly reaching on average 0.5 in parietal and premotor sites for stereo-conditions.

The combined increases in activation and in d' with stereopsis in PCG and the DIPSM sector could suggest that these 2 measures are not independent. If this were true, there would be no justification for the additional d' analysis. We investigated this issue at the single-voxel level in the 4 SAS sites (Supplementary Fig. S6). The relationships between d' and the mean activity of the voxel for the 2 actions entering the d' were best fitted by linear functions. The correlation coefficients, although significant, were equal to 0.05 or smaller, indicating that <1% of the d' variance was explained by the activation level of the voxel (Supplementary Fig. 6). In univariate analyses the most relevant measure is a contrast between conditions, hence we also related the d' values to the difference in mean activity and in SD (Supplementary Figure 7). As one might expect from the definition of d' , it was related, if nonlinearly, to the difference in activity between the 2 actions compared in the voxel in any of the 4 SAS regions (Supplementary Fig. 7A–D). Difference in activity accounted for 50% of the variance in PCG and the DIPSM sector and for 30–40% of d' variance in the RI regions. On the other hand d' correlated little with the difference in standard deviation, which accounted for

<1% of the variance (Supplementary Fig. 7E–H). Hence, we conclude that d' analysis is largely independent of activation levels and thus complementary to univariate analysis of activity levels.

Finally, one anonymous reviewer suggested that CVA, a more established technique (Friston et al. 1995), might yield results similar to the d' analysis, insofar as both are related to linear discriminant analysis. Canonical variance analysis does indeed allow for testing contrasts between action exemplars in the stereo-conditions, thus addressing the accuracy of action representations. However, CVA is a multivariate analysis that maximizes the ratio of the variance explained by the experimental design relative to the variance due to noise. It is thus a global method that tests not only activation effects, but also functional connectivity, resulting in eigen-images, that is, sets of voxels, in which contrasts are significant. As implemented in SPM8, one has to mask the resulting eigen-values with the set of regions of interest, here the SAS regions, and test whether the dimensionality within the mask is larger than zero, indicating that at least one eigen-image is significant. The results for the various action pairs in individual subjects are listed in Supplementary Table 2. The χ^2 reached significance for dimensionality 1 in all subjects for most action pairs, but in only 8/9 subjects for grasping–pushing and in 5/9 subjects for dragging–displacing. After correction for 6 comparisons, dragging–displacing was not significant for any subject and grasping–pushing in only 4. Further testing with masks reduced to single SAS regions indicated that the lack of significance was not linked to any particular SAS regions. This is rather different from the d' analysis results. Indeed, the d' was similarly high for pushing–grasping in all SAS regions (1 in Fig. 9), and for dragging–displacing in PCG and the DIPSM sector (6 in Fig. 9). Thus the CVA analysis is not only very different in nature and outcome, but also seems less sensitive than the d' analysis introduced here.

PPI Analysis of the First Main Experiment

The 3-way interaction in the first experiment produced statistically significant activation sites in a sector of DIPSM, 2 retro-insular sites (PPop and PFCm) and in precentral gyrus. As mentioned, these 4 SAS regions, unlike the action main effect areas, are located outside the classical manipulation–observation network (Jastorff et al. 2010; Ferri, Rizzolatti, et al. 2015). To gain further

insight into the nature of the signals processed in the SAS regions, and ultimately their identity, we investigated their effective connectivity using a PPI analysis of data from the first main experiment. In performing this analysis, we had 2 questions in mind. Firstly, how similar is the connectivity of the stereo-action-specific regions and that of the action-observation network. In this latter network, the parietal nodes typically project to both lower (LOTG) and higher (premotor) level cortices, a feature also present in the connectivity of the action main effect areas (see Supplementary Fig. 8 and Table 3). Secondly, we wondered whether and at what level the stereo-action-specific regions might be connected with the areas displaying the action main effect.

The seed regions were obtained in 18, 19, 21 and 18 subjects for DIPSM, PCG (Fig. 10A), dorsal and ventral RI (Fig. 10B) sites, respectively. These are indicated as stars in Figure 10, while the target areas (Table 2) are indicated by color-matched nodes for DIPSM and PCG sites (Fig. 10A) and for dorsal and ventral RI (Fig. 10B) independently. As predicted, the seed regions were effectively interconnected, often with tight overlaps between seed and target voxels. Targets from different seeds also overlap closely. Connections between the 2 RIs are sparse (Fig. 10B): dorsal RI connects only to the PCG and DIPSM sites and to a BA 44 site, while ventral RI projects to PCG, to the same BA 44 site as dorsal RI (both indicated by purple voxels), and to a lingual gyrus site. These connections are consistent with the presumed vestibular nature of the RI sites, as the lingual site has been implicated in visuo-vestibular interactions (Maffei et al. 2010), and the BA 44 site in vestibular processing (Lobel et al. 1999). The DIPSM site (Fig. 10A) projects to 2 SAS regions, PCG and ventral RI, as well as to PF and an inferior frontal sulcus site. Only PCG (Fig. 10A) has more extensive connectivity, targeting the DIPSM site rather precisely and dorsal RI among other SAS regions, as well as phAIP (see Supplementary Fig. 9), several LOTG sites, and the lingual region mentioned above. Hence the stereo-action-specific regions form a rather closed network, as only 9 of its 16 connections are to outside nodes, all in the left hemisphere, compared with 5 of 7 connections of the main effect network, which connects bilaterally (Supplementary Fig. 8 and Table 3). Only PCG, which is linked to all other network nodes, is connected to the lower-level regions of the manipulation-action observation network (phAIP, pSTS).

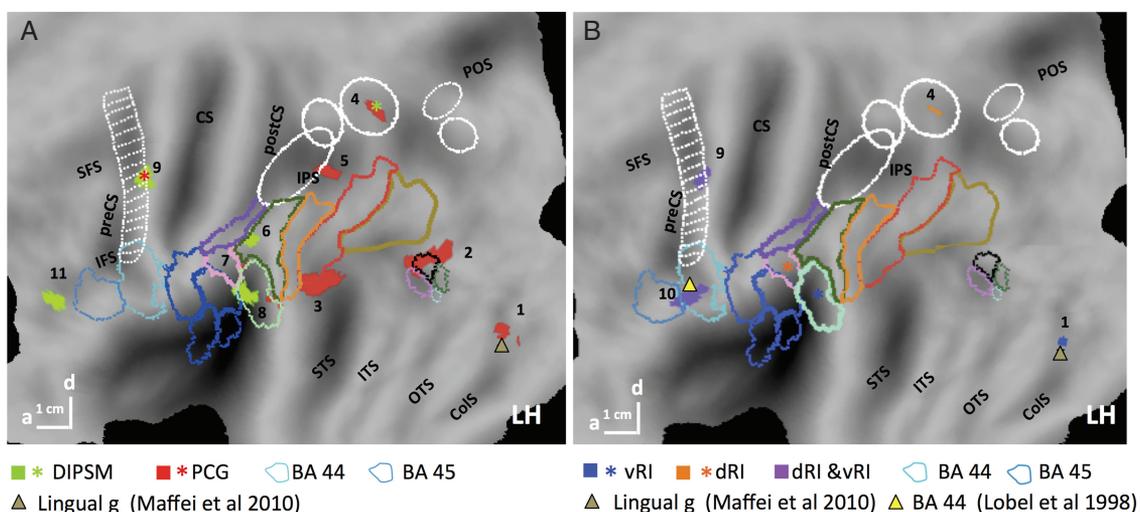


Figure 10. Flat maps showing regions with significant ($P < 0.001$ uncorrected, size > 10 voxels) PPI for PCG (red) and DIPSM (green) seeds (stars, A) and dorsal (orange), and ventral (blue) RI seeds (stars, B). In B, overlap is indicated by purple. Same conventions as Figure 5. Triangles: activation sites from previous studies, as indicated. Numbers: see Table 2.

Thus the PPI analysis shows that the stereo-action-specific regions form a relatively segregated network distinct from the action-observation network. In addition, this analysis provides additional indications that some information circulating within this network is vestibular in nature. This should not be surprising, insofar as stereopsis is needed for exact computation of the kinematics of OFPP observed actions, allowing for more precise comparisons of visual and vestibular signals.

Discussion

Our results show that stereoscopic observation of others' actions activates specific sites in premotor, parietal, and retro-insular cortex, but only for OFPP actions. Importantly, d' analysis indicated that the premotor and parietal specific sites house a more accurate neural representation of observed actions. They also show that action-related stereoscopic signals are processed in a specific network including several visuo-vestibular nodes, suggesting that it provides information about the unfolding of observed actions in space where gravity operates. We will discuss, in succession, the relationship with previous work, the different stereo-action-specific sites and the more general implications of our findings.

Comparison with Previous Studies

Our results are in accord with the recent findings of Jastorff et al. (2016) that stereopsis enhances processing of observed action in premotor and parietal cortex. Importantly, the present study, in addition to uncovering a previously unknown SAS network, provides 3 novel items of information: (1) the stereopsis effects concern only observed actions moving out of the FP plane of the observer, (2) a frontal viewpoint further enhances the stereo effect only at the premotor level, and (3) this increased activity entails a more accurate representation of the observed actions at the single-voxel level. Unlike the Jastorff et al. (2016) study in which the monocular depth cues were limited in the mono-conditions, rich depth information was available in these conditions in the present study. Hence it is unlikely that the present results reflect a nonspecific increase in depth information, rather than the addition of the stereopsis cue. The latter is the only cue able to disambiguate dynamic OFPP changes for nonrigid bodies (Regan and Gray 2001). The Jastorff et al. (2016) study suggested that selectivity for observed stereo-actions emerges at the occipitotemporal cortical level, just as selectivity for observed actions in the FP plane, and that the proportion of 3D-action-selective neurons gradually increases through successive levels of the action observation network, reaching a maximum at the premotor level. The present results are consistent with this view and further imply a large concentration of stereo-action-selective neurons in PCG and in the other SAS regions.

The present study introduced a novel analysis: d' assessed at the single-voxel level. The d' is a statistic whose reliability depends on the number of samples, which in the present study exceeded the 200 samples considered sufficient for avoiding biases in d' estimations (Miller 1996). The d' computation assumes that these 280 samples are independent or that their dependency is taken into account. Our calculation indeed corrects for the autocorrelation within runs measured for each voxel and action separately. d' is a mathematical distance with all properties of distances (e.g., positivity, symmetry, the triangular inequality, and ratio scaling, (Macmillan and Creelman 2005) and hence is a natural choice for expressing neural distance capturing the space separating MR representations of 2 observed actions at

the single-voxel level. Since it has been shown that humans are usually unbiased in discriminating 2 manipulative actions (Platonov and Urban 2015) d' has an implicit interpretation in terms of % correct discrimination. The comparison with mean and differential activation levels and difference in standard deviations indicate that while d' depends on differential activation, as expected from its definition, it is largely independent of mean activation or difference in variability. These results suggest that the d' analysis is valid and complements the univariate analyses, but further work is needed to confirm this first report. The present findings are consistent with a number of previous studies indicating that single voxels carry more information than is generally assumed. Indeed, Serences et al. (2009) have shown that single voxels in early visual areas (V1–3) are tuned to orientation. At higher levels, Harvey et al. (2015, 2013) have demonstrated that single SPL voxels are tuned to numerosity and object size. These studies emphasizing the processing power of single voxels stand in sharp contrast to MVPA which maintains that single voxels carry little information and that more information is available in the activity patterns of a set of voxels. However, MVPA indicates only that a feature might be represented in a region, while d' directly describes the representation of that feature, here observed manipulative actions. d' analysis is not only radically different from MVPA, but it may also be more sensitive, as the percent correct corresponding to the d' values that we report for the stereo conditions (70% and higher) exceed the levels commonly reported for MVPA (often 55–60%). As an additional benefit, d' analysis, unlike MVPA, readily indicates the most sensitive voxels. Like MVPA, however, it assumes that neurons with similar selectivities are clustered (Dubois et al. 2015).

An anonymous reviewer suggested on theoretical grounds that the d' analysis should be similar to the canonical variance analysis (Friston et al. 1995) already available in SPM. CVA, however, is a much more global multivariate analysis and does not at all address the processing power of single voxels. In addition, its output is a significance level, not a distance measure. While CVA confirmed that the action representation in the SAS regions is generally accurate where stereopsis operates, it did not provide measures of this accuracy and could not detect that all 4 action exemplars were equally accurately represented. Such a finding is of theoretical importance, as it relates to the coverage of the observed actions by the cortical representation. Direct comparison of the 2 methods also indicates that CVA may be less sensitive than the d' analysis used here. Indeed, 2 pairs gave nonsignificant results with CVA in the majority of subjects, while one can show by estimating the confidence limits of the d' that it differed significantly from zero for all subjects and for all pairs in PCG and the DIPSM sector. In fact, the lowest d' across subjects and action-pairs is 0.45 in PCG and 0.43 in the DIPSM sector, still significantly different from zero, as the confidence limits for a sample size of 256 span ± 0.22 ($\pm 1.96 \times 0.013^{1/2}$) around the d' of a voxel. Further work will be needed to assess the respective merits of d' analysis, univariate activation and MVPA.

The Lead Node of the Stereo-action-specific Network: PCG

In the present study, the PCG site (Fig. 6) located ventral to the premotor main effect of action site, appears as the main stereo-action-specific region. The PCG site corresponds to the part of area 6 receiving vestibular input (Lobel et al. 1999; Ebata et al. 2004). As shown in Figure 11 it is located in ventral premotor cortex (Tomassini et al. 2007; Schubotz et al. 2010), between pHF5c (Ferri, Peeters, et al. 2015) and M1 (Geyer et al. 1996), and thus should correspond to F4 of the monkey, which has similar

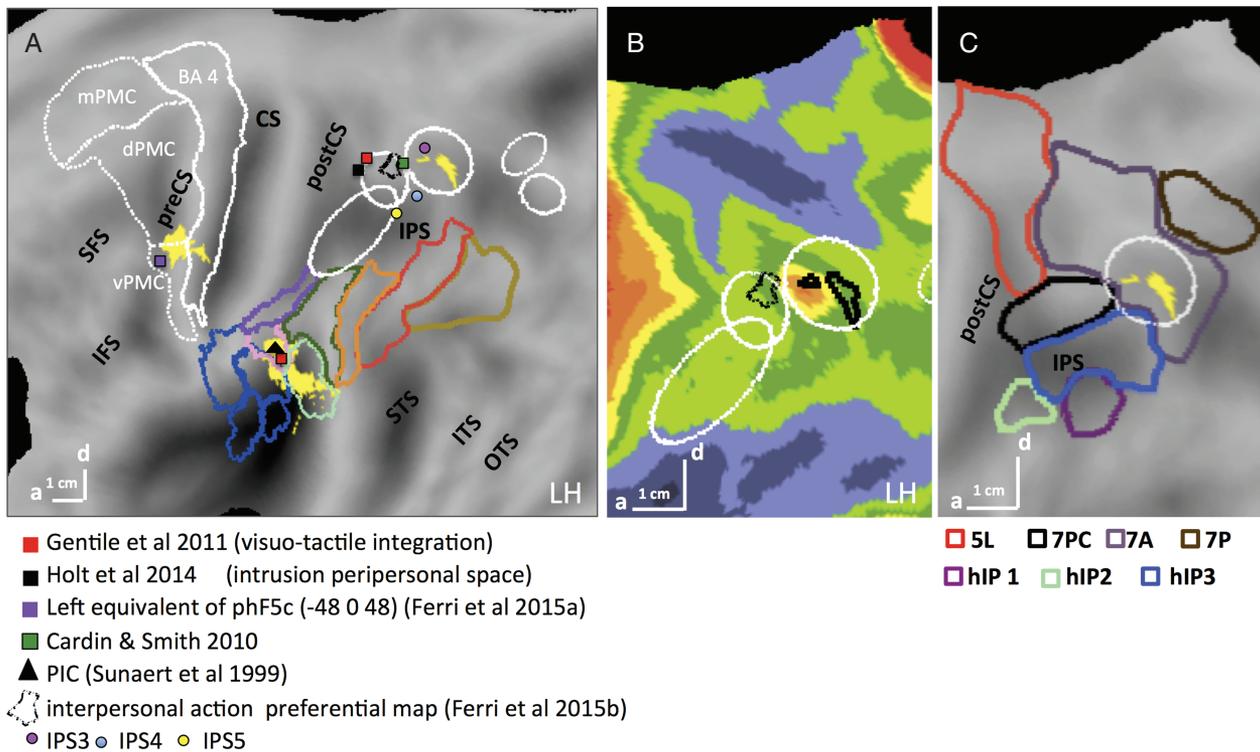


Figure 11. (A) Flat map (partial left hemisphere) showing the stereo-specific regions in relation to subregions of area 6 and area 4 (Geyer et al. 1996; Geyer 2004), and to functionally defined regions (see insets); IPS3-5 centers of retinotopic maps from Konen and Kastner (2008); (B, C) flatmap (partial left hemisphere) showing SAS sector in DIPSM in relation to myelin pattern-B, (Abdollahi et al. 2014)-, and to cytoarchitectonic regions (C) of Choi et al. (2006) and Scheperjans et al. (2008). In B the putative rostral VIP site (thin black lines) from Ferri, Rizzolatti, et al. (2015) and its putative extension in DIPSM (thick black lines) curl around the myelin peak, presumably corresponding to LIPv.

topological relationships and also receives vestibular input (Akbarian et al. 1994; Barbas et al. 1999). Hence, we propose the PCG site as the putative human homolog of monkey F4 (phF4). Because it is connected to the 3 other stereo-action-specific nodes, 2 of which located in what is generally believed to be vestibular cortex in humans, we propose that PCG represents others' actions as they unfold in space under the effect of gravity (see below). This proposal is consistent with its proximity to a site involved in reorienting attention along the z-axis near the body (Chen et al. 2012). PCG is located more caudally relative to the stereo interaction sites described in the previous study (Jastorff et al. 2016), both of which are located in the precentral sulcus. However, the biological motion stimuli shown in that study were diverse whole-body actions, such as dancing, shown with no background from a single viewpoint. Hence it is likely that the representation of stereo-actions in PCG was not activated in that study.

The PCG site is the lead node of the stereo-action-specific network providing most of its outputs. It is also distinct from the other SAS nodes in its dependence on a frontal viewpoint, suggesting that it may play a role in face-to-face interactions requiring accurate assessment of actions performed by conspecifics, especially when directed toward the observer. Furthermore, it shares with the DIPSM sector a representation of stereo-actions that is both accurate and homogeneous. This PCG site, although functionally clearly distinct from the dorsal PCS site of the generalized (over stereopsis and viewpoint) action-observation network, is contiguous and ventral to this dorsal site, the location of which matches earlier findings (Ferri, Peeters, et al. 2015; Ferri, Rizzolatti, et al. 2015). The latter site may correspond to

F2vr of the monkey, which is also related to hand and wrist actions (Raos et al. 2004). The spatial arrangements of the PCG and dPCS sites are reminiscent of human and monkey FEF, a ventro-caudal subregion of which is devoted to pursuit in 3D and a target of vestibular input (Ebata et al. 2004).

Other Nodes of the Stereo-action-specific Network

The parietal stereo-action-specific region was located in DIPSM. Its profile indicates only a weak mono-action response, matching its location at the periphery of the manipulation action observation network (Ferri, Rizzolatti, et al. 2015). It is located just behind a region (Fig. 11A) involved in optic flow processing (Cardin and Smith 2010), a characteristic of VIP, which in monkeys is connected reciprocally with F4 (Rizzolatti and Luppino 2001), as the DIPSM site is with PCG (phF4). Extensive studies have attributed 2 functions to monkey VIP (Clery et al. 2015): first, the visuo-tactile processing of intrusion into peripersonal space, especially near the head (Graziano and Cooke 2006; Avillac et al. 2007; Guipponi et al. 2015), a characteristic appearing predominantly in the rostral part of VIP, and second, visuo-vestibular processing of self-motion (Schaaafsma and Duysens 1996; Bremmer et al. 2002; Zhang et al. 2004; Chen et al. 2011a) that is distributed over most of its extent. We have earlier proposed (Ferri, Rizzolatti, et al. 2015) that dorsal DIPSA, just in front of DIPSM, may correspond to rostral VIP, processing intrusions into the peripersonal space of the actor, as suggested by its overlap with visuo-tactile regions (Gentile et al. 2011) and visual intrusion-processing regions (Holt et al. 2014). That proposal implied that only ventral DIPSA corresponds to posterior AIP, rather than the entire DIPSA region as

initially postulated (Durand et al. 2009; Vanduffel et al. 2014). Similarly, we now propose that only the ventral part of DIPSM corresponds to monkey anterior LIP, specifying our earlier interpretation (Orban et al. 2006; Durand et al. 2009; Vanduffel et al. 2014), and that the SAS site in DIPSM corresponds to a more caudal part of VIP. We further propose that its function includes the processing of movements of body parts in peripersonal space, in addition to overall motion of the body in extrapersonal space. Just as postulated for the anterior part of VIP (Ferri, Rizzolatti, et al. 2015), this function can apply to the subject's own peripersonal space or to that of an observed actor. Hence, according to our hypothesis, the signals exchanged by the SAS site in DIPSM and PCG reflect the unfolding of actions in the peripersonal space of the actor. This presumed identification of the DIPSM site as an extension of the human counterpart of VIP fits with its proximity to the myelin concentration near the IPS (Fig. 11B), supposedly corresponding to monkey vLIP (Glasser et al. 2012). The subdivision of DIPSM into 2 functional subregions is consistent with the original report of Sunaert et al. (1999), who described 2 motion-sensitive regions at that level, and the overlap of DIPSM with 2 distinct retinotopic maps (IPS3 and 4, Fig. 11A). It also fits with the cytoarchitectonic parcellation: DIPSM overlaps mainly with 7A (Scheperjans et al. 2008) but also with hIP3 (Fig. 13C) (Choi et al. 2006).

The 2 remaining stereo-action-specific regions are dorsal and ventral RI, located in what is assumed to be human vestibular cortex (Bottini et al. 1994; Kahane et al. 2003; Lopez et al. 2012; zu Eulenburg et al. 2012; Dieterich and Brandt 2015). In fact, the retro-insular region of the vestibular network described by Indovina et al. (2005), supposedly belonging to a network encoding the visual effects of gravity (internal 1 g network), is centered on the border between dorsal and ventral RI (Fig. 11). Dorsal RI overlaps with posterior insular cortex, a visual motion region (Sunaert et al. 1999; Claeys et al. 2003; Frank et al. 2014) putatively corresponding to monkey vestibular posterior sylvian region (Chen et al. 2011b). This area is located next to monkey Parieto-insular vestibular cortex, the core vestibular region (Guldin and Grusser 1998; Chen et al. 2011b), to which ventral RI may then correspond. Indeed, ventral RI was less reactive to 3D action observation (Fig. 7) than its dorsal counterpart, and their d' patterns are somewhat different. Interestingly, the RI region also receives convergent visuo-tactile input (Gentile et al. 2011), suggesting that it integrates multiple sensory channels, as does dDIPSA/DIPSM. We propose that the RI connection to PCG (Fig. 10) provides visual gravity signals from the internal 1 g network described by Indovina et al. (2005). Indeed, the 2 regions displayed only a nonsignificant increase in d' between the observed actions, suggesting a more ancillary role in the network. This raises the issue of vestibular input to the classic action observation network obtained with lateral viewing without stereo. This network generally includes DIPSM (Jastorff et al. 2010; Ferri, Rizzolatti, et al. 2015), thus parts of DIPSM neighboring the SAS sector might provide vestibular input. Lower-level, occipitotemporal regions are a possible alternative, since Maffei et al. (2015) documented gravity effects on biological motion in the pOTS region.

The stereo-action-specific network is entirely left-lateralized. Since it is driven primarily by the frontal view in the factorial main experiment, its lateralization is unlikely to reflect a simple visual asymmetry in the stimulus, as repeatedly reported for the parietal action regions (Wheaton et al. 2004; Jastorff et al. 2010; Ferri, Rizzolatti, et al. 2015). The left bias is more likely to reflect motor dominance in the left hemisphere. This may reflect the fact that in all videos the hand used by the actor was the right hand (Shmuelof and Zohary 2006). Indeed, the effects of stereo

at the premotor and parietal level were much more bilateral in the previous study (Jastorff et al. 2016), in which left and right limbs contributed equally to the actions portrayed in the BM stimuli. Further work is needed to understand the left lateralization of the stereo-action-specific regions.

In conclusion, the stereo-action-specific regions are closely related to regions that also process vestibular signals, unsurprisingly so since stereopsis allows precise assessment of the kinematics of the actions observed, including accelerations. This functional overlap may help explain that vestibular symptoms such as dizziness or nausea have been reported while viewing of 3D movies (Yang et al. 2012; Solimini 2013).

Implications of the Present Study

The findings of the present study indicate that stereopsis improves action representation and activates a network of regions with vestibular inputs, possibly providing a description of how the observed actions unfold in peripersonal space. Such observations have profound theoretical implications. First, there are significant implications regarding the debate about the role of premotor cortex in action observation: whether premotor mirror neurons are critical for action perception (Rizzolatti et al. 2014), or are not (Rogalsky et al. 2013; Vannuscorps and Caramazza 2015). If action perception includes the unfolding of observed actions in space, action perception may indeed require activation of the premotor level. Indeed this level hosts the required signals, deriving from the integration of signals arising from parietal and vestibular cortex. Second, it also sheds light onto the question as to why stereopsis is crucial to primate evolution (Barton 2004; Lambert et al. 2005; Heesy et al. 2011). Traditionally, it is considered essential for the visual control of hand actions, for which motor cortex sends signals directly to spinal motoneurons (Lemon 2008). This is generally understood as visual control of object manipulation (Watt and Bradshaw 2003). The present study indicates that face-to-face interactions and sharing workspaces with conspecifics may represent additional reasons for stereoscopic control of hand actions.

Supplementary material

Supplementary Material can be found at <http://www.cercor.oxfordjournals.org/online>.

Funding

This work is supported by ERC Parietal action 323606 (to GAO), by a grant from Telenet NV (to G.A.O.) and by IUAP VII-11 (to G.R.). Funding to pay the Open Access publication charges for this article was provided by ERC parietalaction (to GAO).

Notes

The authors are indebted to Dr F Lacquaniti and Dr S Raiguel for comments on an earlier version of the manuscript. *Conflict of Interest:* None declared.

References

- Abdollahi RO, Jastorff J, Orban GA. 2013. Common and segregated processing of observed actions in human SPL. *Cereb Cortex*. 23:2734–2753.
- Abdollahi RO, Kolster H, Glasser MF, Robinson EC, Coalson TS, Dierker D, Jenkinson M, Van Essen DC, Orban GA. 2014.

- Correspondences between retinotopic areas and myelin maps in human visual cortex. *Neuroimage*. 99:509–524.
- Akbarian S, Grusser OJ, Guldin WO. 1994. Corticofugal connections between the cerebral cortex and brainstem vestibular nuclei in the macaque monkey. *J Comp Neurol*. 339(3): 421–437.
- Alvarez TL, Alkan Y, Gohel S, Douglas Ward B, Biswal BB. 2010. Functional anatomy of predictive vergence and saccade eye movements in humans: a functional MRI investigation. *Vision Res*. 50:2163–2175.
- Avillac M, Ben Hamed S, Duhamel JR. 2007. Multisensory integration in the ventral intraparietal area of the macaque monkey. *J Neurosci*. 27:1922–1932.
- Barbas H, Ghashghaei H, Dombrowski S, Rempel-Clower N. 1999. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. *J Comp Neurol*. 410:343–367.
- Barton RA. 2004. From The Cover: Binocularity and brain evolution in primates. *Proc Natl Acad Sci USA*. 101:10113–10115.
- Binkofski F, Buccino G, Posse S, Setz RJ, Rizzolatti G, Freund H. 1999. A fronto-parallel circuit for object manipulation in man: evidence from an fMRI-study. *Eur J Neurosci*. 11:3276–3286.
- Bottini G, Sterzi R, Paulesu E, Vallar G, Cappa SF, Erminio F, Passingham RE, Frith CD, Frackowiak RS. 1994. Identification of the central vestibular projections in man: a positron emission tomography activation study. *Exp Brain Res*. 99:164–169.
- Bradski G. 2000. The OpenCV Library. Dr Dobb's J Soft Tools. 25:120–126.
- Bremmer F, Klam F, Duhamel JR, Ben Hamed S, Graf W. 2002. Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Eur J Neurosci*. 16:1569–1586.
- Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ. 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur J Neurosci*. 13:400–404.
- Bulthe J, De Smedt B, Op de Beeck HP. 2015. Visual number beats abstract numerical magnitude: format-dependent representation of Arabic digits and dot patterns in human parietal cortex. *J Cogn Neurosci*. 27:1376–1387.
- Caggiano V, Fogassi L, Rizzolatti G, Pomper JK, Thier P, Giese MA, Casile A. 2011. View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *Curr Biol*. 21:144–148.
- Cardin V, Smith AT. 2010. Sensitivity of human visual and vestibular cortical regions to egomotion-compatible visual stimulation. *Cereb Cortex*. 20:1964–1973.
- Caspers S, Geyer S, Schleicher A, Mohlberg H, Amunts K, Zilles K. 2006. The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. *Neuroimage*. 33:430–448.
- Chen A, DeAngelis GC, Angelaki DE. 2011a. A comparison of vestibular spatiotemporal tuning in macaque parietoinsular vestibular cortex, ventral intraparietal area, and medial superior temporal area. *J Neurosci*. 31:3082–3094.
- Chen A, DeAngelis GC, Angelaki DE. 2011b. Convergence of vestibular and visual self-motion signals in an area of the posterior sylvian fissure. *J Neurosci*. 31:11617–11627.
- Chen Q, Weidner R, Vossel S, Weiss PH, Fink GR. 2012. Neural mechanisms of attentional reorienting in three-dimensional space. *J Neurosci*. 32:13352–13362.
- Choi HJ, Zilles K, Mohlberg H, Schleicher A, Fink GR, Armstrong E, Amunts K. 2006. Cytoarchitectonic identification and probabilistic mapping of two distinct areas within the anterior ventral bank of the human intraparietal sulcus. *J Comp Neurol*. 495:53–69.
- Claeys KG, Lindsey DT, De Schutter E, Orban GA. 2003. A higher order motion region in human inferior parietal lobule: Evidence from fMRI. *Neuron*. 40:631–642.
- Clery J, Guipponi O, Wardak C, Ben Hamed S. 2015. Neuronal bases of peripersonal and extrapersonal spaces, their plasticity and their dynamics: knowns and unknowns. *Neuropsychologia*. 70:313–326.
- Culham JC, Danckert SL, DeSouza JFX, Gati JS, Menon RS, Goodale MA. 2003. Visually guided grasping produces fMRI activation in dorsal but not ventral stream brain areas. *Exp Brain Res*. 153:180–189.
- Dieterich M, Brandt T. 2015. The bilateral central vestibular system: its pathways, functions, and disorders. *Ann N Y Acad Sci*. 1343:10–26.
- Dubois J, de Berker AO, Tsao DY. 2015. Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. *J Neurosci*. 35:2791–2802.
- Durand JB, Peeters R, Norman JF, Todd JT, Orban GA. 2009. Parietal regions processing visual 3D shape extracted from disparity. *Neuroimage*. 46:1114–1126.
- Ebata S, Sugiuchi Y, Izawa Y, Shinomiya K, Shinoda Y. 2004. Vestibular projection to the periarculate cortex in the monkey. *Neurosci Res*. 49:55–68.
- Eger E, Pinel P, Dehaene S, Kleinschmidt A. 2015. Spatially invariant coding of numerical information in functionally defined subregions of human parietal cortex. *Cereb Cortex*. 25:1319–1329.
- Eickhoff SB, Amunts K, Mohlberg H, Zilles K. 2006. The human parietal operculum. II. Stereotaxic maps and correlation with functional imaging results. *Cereb Cortex*. 16:268–279.
- Ferri S, Kolster H, Jastorff J, Orban GA. 2013. The overlap of the EBA and the MT/V5 cluster. *Neuroimage*. 66:412–425.
- Ferri S, Peeters R, Nelissen K, Vanduffel W, Rizzolatti G, Orban GA. 2015. A human homologue of monkey F5c. *Neuroimage*. 111:251–266.
- Ferri S, Rizzolatti G, Orban GA. 2015. The organization of the posterior parietal cortex devoted to upper limb actions: An fMRI study. *Hum Brain Mapp*. 36:3845–3866.
- Frank SM, Baumann O, Mattingley JB, Greenlee MW. 2014. Vestibular and visual responses in human posterior insular cortex. *J Neurophysiol*. 112:2481–2491.
- Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. 1997. Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage*. 6:218–229.
- Friston KJ, Frith CD, Frackowiak RS, Turner R. 1995. Characterizing dynamic brain responses with fMRI: a multivariate approach. *Neuroimage*. 2:166–172.
- Gallivan J, Johnsrude I, Flanagan J. 2016. Planning ahead: object-directed sequential actions decoded from human frontoparietal and occipitotemporal networks. *Cereb Cortex*. 26(2):708–730.
- Gallivan JP, McLean DA, Flanagan JR, Culham JC. 2013. Where one hand meets the other: limb-specific and action-dependent movement plans decoded from preparatory signals in single human frontoparietal brain areas. *J Neurosci*. 33:1991–2008.
- Gazzola V, Rizzolatti G, Wicker B, Keysers C. 2007. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage*. 35:1674–1684.
- Gentile G, Petkova VI, Ehrsson HH. 2011. Integration of visual and tactile signals from the hand in the human brain: an fMRI study. *J Neurophysiol*. 105:910–922.
- Georgieva S, Peeters R, Kolster H, Todd JT, Orban GA. 2009. The processing of three-dimensional shape from disparity in the human brain. *J Neurosci*. 29:727–742.

- Geyer S. 2004. The microstructural border between the motor and the cognitive domain in the human cerebral cortex. *Adv Anat Embryol Cell Biol.* 174:1–89.
- Geyer S, Ledberg A, Schleicher A, Kinomura S, Schormann T, Bürgel U, Klingberg T, Larsson J, Zilles K, Roland PE. 1996. Two different areas within the primary motor cortex of man. *Nature.* 382:805–807.
- Gitelman DR, Penny WD, Ashburner J, Friston KJ. 2003. Modeling regional and psychophysiological interactions in fMRI: The importance of hemodynamic deconvolution. *Neuroimage.* 19:200–207.
- Glasser MF, Preuss MT, Nair G, Rilling J, Zhang X, Li L, Van Essen DC. 2012. Improved cortical myelin maps in humans, chimpanzees, and macaques allow identification of putative areal homologies. New Orleans, LA: Society for Neuroscience.
- Grafton ST, Hamilton AF. 2007. Evidence for a distributed hierarchy of action representation in the brain. *Hum Mov Sci.* 26:590–616.
- Graziano MS, Cooke DF. 2006. Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia.* 44:845–859.
- Green DM, Swets JA. 1966. *Signal Detection Theory and Psychophysics.* New York: Wiley.
- Guipponi O, Clery J, Odouard S, Wardak C, Ben Hamed S. 2015. Whole brain mapping of visual and tactile convergence in the macaque monkey. *Neuroimage.* 117:93–102.
- Guldin WO, Grüsser OJ. 1998. Is there a vestibular cortex? *Trends Neurosci.* 21:(6)254–259.
- Harvey B, Fracasso A, Petridou N, Dumoulin S. 2015. Overlapping topographic representations of numerosity and object size in human parietal cortex. *J Vis.* 15:1283.
- Harvey BM, Klein BP, Petridou N, Dumoulin SO. 2013. Topographic representation of numerosity in the human parietal cortex. *Science.* 341:1123–1126.
- Heesy CP, Kamilar JM, Willms J. 2011. Retinogeniculostriate pathway components scale with orbit convergence only in primates and not in other mammals. *Brain Behav Evol.* 77:105–115.
- Hirschmuller H. 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Trans Pattern Anal Mach Intell.* 30:328–341.
- Holmes A, Friston K. 1998. Generalisability, random effects and population inference. *Neuroimage.* 7:S754.
- Holt DJ, Cassidy BS, Yue X, Rauch SL, Boeke EA, Nasr S, Tootell RB, Coombs G 3rd. 2014. Neural correlates of personal space intrusion. *J Neurosci.* 34:4123–4134.
- Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G. 1999. Cortical mechanisms of human imitation. *Science.* 286:2526–2528.
- Indovina I, Maffei V, Bosco G, Zago M, Macaluso E, Lacquaniti F. 2005. Representation of visual gravitational motion in the human vestibular cortex. *Science.* 308:416–419.
- Jastorff J, Abdollahi R, Fasano F, Orban GA. 2016. Seeing biological actions in 3D: an fMRI study. *Hum Brain Mapp.* 37:203–219.
- Jastorff J, Begliomini C, Fabbri-Destro M, Rizzolatti G, Orban GA. 2010. Coding observed motor acts: different organizational principles in the parietal and premotor cortex of humans. *J Neurophysiol.* 104:128–140.
- Kahane P, Hoffmann D, Minotti L, Berthoz A. 2003. Reappraisal of the human vestibular cortex by cortical electrical stimulation study. *Ann Neurol.* 54:615–624.
- Kimmig H, Greenlee MW, Gondan M, Schira M, Kassubek J, Mergner T. 2001. Relationship between saccadic eye movements and cortical activity as measured by fMRI: quantitative and qualitative aspects. *Exp Brain Res.* 141(2):184–194.
- Kolster H, Peeters R, Orban GA. 2010. The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *J Neurosci.* 30:9801–9820.
- Konen CS, Kastner S. 2008. Two hierarchically organized neural systems for object information in human visual cortex. *Nat Neurosci.* 11:224–231.
- Konen CS, Mruzec RE, Montoya JL, Kastner S. 2013. Functional organization of human posterior parietal cortex: grasping- and reaching-related activations relative to topographically organized cortex. *J Neurophysiol.* 109:2897–2908.
- Koyama M, Hasegawa I, Osada T, Adachi Y, Nakahara K, Miyashita Y. 2004. Functional magnetic resonance imaging of macaque monkeys performing visually guided saccade tasks: comparison of cortical eye fields with humans. *Neuron.* 41:795–807.
- Lambot MA, Depasse F, Noel JC, Vanderhaeghen P. 2005. Mapping labels in the human developing visual system and the evolution of binocular vision. *J Neurosci.* 25:7232–7237.
- Lemon RN. 2008. Descending pathways in motor control. *Annu Rev Neurosci.* 31:195–218.
- Likova LT, Tyler CW. 2007. Stereomotion processing in the human occipital cortex. *Neuroimage.* 38:293–305.
- Lobel E, Kleine JF, Leroy-Willig A, Van de Moortele PF, Le Bihan D, Grusser OJ, Berthoz A. 1999. Cortical areas activated by bilateral galvanic vestibular stimulation. *Ann N Y Acad Sci.* 871:313–323.
- Lopez C, Blanke O, Mast FW. 2012. The human vestibular cortex revealed by coordinate-based activation likelihood estimation meta-analysis. *Neuroscience.* 212:159–179.
- Macmillan NA, Creelman CD. 2005. *Detection Theory a user guide.* Mahwah, New Jersey, London: Lawrence Erlbaum Associates.
- Maffei V, Indovina I, Macaluso E, Ivanenko YP, Orban GA, Lacquaniti F. 2015. Visual gravity cues in the interpretation of biological movements: neural correlates in humans. *Neuroimage.* 104:221–230.
- Maffei V, Macaluso E, Indovina I, Orban G, Lacquaniti F. 2010. Processing of targets in smooth or apparent motion along the vertical in the human brain: an fMRI study. *J Neurophysiol.* 103:360–370.
- Miller J. 1996. The sampling distribution of d' . *Percept Psychophys.* 58:65–72.
- Mukamel R, Ekstrom AD, Kaplan J, Iacoboni M, Fried I. 2010. Single-neuron responses in humans during execution and observation of actions. *Curr Biol.* 20:750–756.
- Nichols T, Brett M, Andersson J, Wager T, Poline JB. 2005. Valid conjunction inference with the minimum statistic. *Neuroimage.* 25:653–660.
- Norman KA, Polyn SM, Detre GJ, Haxby JV. 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cogn Sci.* 10:424–430.
- Orban GA. 2011. The extraction of 3D shape in the visual system of human and nonhuman primates. *Annu Rev Neurosci.* 34:361–388.
- Orban GA, Claeys K, Nelissen K, Smans R, Sunaert S, Todd JT, Wardak C, Durand J-B, Vanduffel W. 2006. Mapping the parietal cortex of human and non-human primates. *Neuropsychologia.* 44:2647–2667.
- Parker AJ. 2007. Binocular depth perception and the cerebral cortex. *Nat Rev Neurosci.* 8:379–391.
- Pauwels K, Van Hulle MM. 2009. Optic flow from unstable sequences through local velocity constancy maximization. *Image and Vision Comput.* 27:579–587.
- Person AL, Perkel DJ. 2007. Pallidal neuron activity increases during sensory relay through thalamus in a songbird circuit essential for learning. *J Neurosci.* 27:8687–8698.
- Pilgramm S, Haas de B, Helm F, Zentgraf K, Stark R, Munzert J, Krüger B. 2016. Motor imagery of hand actions: Decoding the

- content of motor imagery from brain activity in frontal and parietal motor areas. *Hum Brain Mapp.* 37:81–93.
- Platonov A, Orban GA. 2015. Psychometric curves describe action discrimination in humans. *J Vis.* 15:561.
- Raos V, Umiltá M-A, Gallese V, Fogassi L. 2004. Functional properties of grasping-related neurons in the dorsal premotor area F2 of the macaque monkey. *J Neurophysiol.* 92:1990–2002.
- Regan D, Gray R. 2001. Hitting what one wants to hit and missing what one wants to miss. *Vision Res.* 41:3321–3329.
- Rizzolatti G, Cattaneo L, Fabbri-Destro M, Rozzi S. 2014. Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol Rev.* 94:655–706.
- Rizzolatti G, Luppino G. 2001. The cortical motor system. *Neuron.* 31:889–901.
- Rogalsky C, Raphel K, Tomkovicz V, O'Grady L, Damasio H, Bellugi U, Hickok G. 2013. Neural Basis of Action Understanding: Evidence from Sign Language Aphasia. *Aphasiology.* 27:1147–1158.
- Schaafsma SJ, 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.* 76:4056–4068.
- Scheperjans F, Hermann K, Eickhoff SB, Amunts K, Schleicher A, Zilles K. 2008. Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. *Cereb Cortex.* 18:846–867.
- Schubotz RI, Anwander A, Knosche TR, von Cramon DY, Tittgemeyer M. 2010. Anatomical and functional parcellation of the human lateral premotor cortex. *Neuroimage.* 50:396–408.
- Serences JT, Saproo S, Scolari M, Ho T, Muftuler LT. 2009. Estimating the influence of attention on population codes in human visual cortex using voxel-based tuning functions. *Neuroimage.* 44:223–231.
- Shmuelof L, Zohary E. 2006. A mirror representation of others' actions in the human anterior parietal cortex. *J Neurosci.* 26:9736–9742.
- Solimini. 2013. Are there side effects to watching 3d movies? a prospective crossover observational study on visually induced motion sickness. *PLoS ONE.* 8(2):e56160.
- Solis MM, Doupe AJ. 1997. Anterior forebrain neurons develop selectivity by an intermediate stage of birdsong learning. *J Neurosci.* 17:6447–6462.
- Sunaert S, Van Hecke P, Marchal G, Orban GA. 1999. Motion-responsive regions of the human brain. *Exp Brain Res.* 127:355–370.
- Theunissen F, Doupe AJ. 1998. Temporal and spectral sensitivity of complex auditory neurons in the nucleus HVC of male zebra finches. *J Neurosci.* 18:3786–3802.
- Tomassini V, Jbabdi S, Klein JC, Behrens TEJ, Pozzilli C, Matthews PM, Rushworth MFS, Johansen-Berg H. 2007. Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *J Neurosci.* 27:10259–10269.
- Tsao DY, Vanduffel W, Sasaki Y, Fize D, Knutsen TA, Mandeville JB, Wald LL, Dale AM, Rosen BR, Van Essen DC, et al. 2003. Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron.* 39:555–568.
- Vanduffel W, Zhu Q, Orban GA. 2014. Monkey Cortex through fMRI Glasses. *Neuron.* 83:533–550.
- Van Essen DC. 2005. A population-average, landmark- and surface-based (PALS) atlas of human cerebral cortex. *Neuroimage.* 28:635–662.
- Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH. 2001. An integrated software suite for surface-based analyses of cerebral cortex. *J Am Med Inform Assoc.* 8:443–459.
- Vannuscorps G, Caramazza A. 2015. Motor simulation does not underlie action perception: evidence from upper limb dysmelia. *J Vis.* 15:559.
- Wardle S, Ritchie J, Seymour K, Carlson T. 2015. What information is 'decoded' from stimulus orientation with fMRI and MVPA? *J Vis.* 15:993.
- Watt SJ, Bradshaw MF. 2003. The visual control of reaching and grasping: binocular disparity and motion parallax. *J Exp Psychol Hum Percept Perform.* 29:404–415.
- Wheaton KJ, Thompson JC, Syngeniotes A, Abbott DF, Puce A. 2004. Viewing the motion of human body parts activates different regions of premotor, temporal, and parietal cortex. *Neuroimage.* 22:277–288.
- Yang SN, Schlieski T, Selmins B, Cooper SC, Doherty RA, Coriveau PJ, Sheedy JE. 2012. Stereoscopic viewing and reported perceived immersion and symptoms. *Optom Vis Sci.* 89:1068–1080.
- Zhang T, Heuer HW, Britten KH. 2004. Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. *Neuron.* 42:993–1001.
- zu Eulenburg P, Caspers S, Roski C, Eickhoff SB. 2012. Meta-analytical definition and functional connectivity of the human vestibular cortex. *Neuroimage.* 60:162–169.