RESEARCH ARTICLE

WILEY

Decoding point-light displays and fully visible hand grasping actions within the action observation network

Settimio Ziccarelli [ORCID] | Antonino Errante | Leonardo Fogassi

Department of Medicine and Surgery, University of Parma, Parma, Italy

Correspondence
Leonardo Fogassi, Department of Medicine and Surgery, University of Parma, 43125 Parma, Italy.
Email: leonardo.fogassi@unipr.it

Abstract

Action observation typically recruits visual areas and dorsal and ventral sectors of the parietal and premotor cortex. This network has been collectively termed as extended action observation network (eAON). Within this network, the elaboration of kinematic aspects of biological motion is crucial. Previous studies investigated these aspects by presenting subjects with point-light displays (PLDs) videos of whole-body movements, showing the recruitment of some of the eAON areas. However, studies focused on cortical activation during observation of PLDs grasping actions are lacking. In the present functional magnetic resonance imaging (fMRI) study, we assessed the activation of eAON in healthy participants during the observation of both PLDs and fully visible hand grasping actions, excluding confounding effects due to low-level visual features, motion, and context. Results showed that the observation of PLDs grasping stimuli elicited a bilateral activation of the eAON. Region of interest analyses performed on visual and sensorimotor areas showed no significant differences in signal intensity between PLDs and fully visible experimental conditions, indicating that both conditions evoked a similar motor resonance mechanism. Multivoxel pattern analysis (MVPA) revealed significant decoding of PLDs and fully visible grasping observation conditions in occipital, parietal, and premotor areas belonging to eAON. Data show that kinematic features conveyed by PLDs stimuli are sufficient to elicit a complete action representation, suggesting that these features can be disentangled within the eAON from the features usually characterizing fully visible actions. PLDs stimuli could be useful in assessing which areas are recruited, when only kinematic cues are available, for action recognition, imitation, and motor learning.

KEYWORDS
action observation, biological motion, fMRI, kinematic information, mirror neuron system, MVPA, point-light displays

1 | INTRODUCTION

During the observation of an action performed by another individual, a mechanism provided by the mirror neuron system (MNS) allows the observer to automatically understand another's action by matching its visual description onto her/his motor representation of the same action (Rizzolatti et al., 2014). Mirror neurons have been originally found in monkey ventral premotor cortex (F5) and then in the inferior parietal lobule (PFG) (Di Pellegrino et al., 1992; Fogassi et al., 2005); subsequently, neurons with mirror properties have been found in other cortical areas, such as, for example, the anterior intraparietal area (AIP) (Lanzilotto et al., 2019; Maeda et al., 2015), dorsal premotor cortex (PMd) (Papadourakis & Raos, 2019; Tkach et al., 2007) and pre-supplementary motor cortex (pre-SMA) (Lanzilotto et al., 2016; Yoshida et al., 2011).

By using electrophysiological and neuroimaging techniques, a comparable parieto-premotor MNS has been described also in humans, homologous to that of monkeys, that includes the ventral premotor cortex (PMv) plus the inferior frontal gyrus (IFG) and the inferior parietal lobule (IPL) (Molenberghs et al., 2012). Further studies demonstrated that mere action observation recruits an extended network of cortical and subcortical areas, collectively called "extended action observation network" (eAON), including PMd, a sector of intraparietal sulcus (IPS), superior parietal lobule (SPL), primary somatosensory cortex (SI), occipitotemporal areas (middle temporal area [MT]), posterior superior temporal sulcus (pSTS), and the lateral part of cerebellum (Abdelgabar et al., 2019; Errante & Fogassi, 2020; Filimon et al., 2007; Gazzola & Keysers, 2009; see also Caspers et al., 2010; Hardwick et al., 2018; Molenberghs et al., 2012).

Within the eAON, there are parietal and frontal areas crucial for generating in the observer an internal motor resonance with the observed action (Rizzolatti et al., 2014). Some of these areas have also been proposed to take part in the elaboration of additional features of the observed actions (Kemmerer, 2021) including kinematics (Filimon et al., 2007; Koul et al., 2018), type of grip used to achieve the final action goal (Errante et al., 2021; Grafton & Hamilton, 2007), and contextual information (Amoruso et al., 2016; Iacoboni et al., 2005).

The decoding of the kinematic aspects of observed actions relies on the elaboration of biological motion by high-order visual areas located in the inferotemporal cortex, as it has been clearly demonstrated in both monkeys and humans (Caspers et al., 2010; Perrett et al., 1989). The outcome of this elaboration is then provided to the eAON. Interestingly, within this latter network, the PMd sector has been shown to play a role in decoding observed actions complying with the "two-thirds power law" (Casile et al., 2010), which describes the relation between the speed and the trajectory of biological movements (Lacquaniti et al., 1983). According to this law, the movement velocity depends on the trajectory curvature, namely velocity is lower in more curved parts than in less curved parts of the trajectory. One method that allows to study the kinematic properties of biological motion, in absence of pictorial content, is that of point-light displays (PLDs) (Blake & Shiffrar, 2007; Johansson, 1973; Pavlova, 2012; Thornton, 2006). Using this technique, consisting in small lights attached to the main joints of one's body on a dark background, so that only the lights are visible, it is possible to present visually impoverished stimuli of several human behaviors. Behavioral data showed that motion information conveyed by PLDs is enough to distinguish biological from nonbiological actions (Hiris, 2007; Johansson, 1973; Lapenta et al., 2017) and also to recognize features such as, for example, the gender or the emotional state of the observed agent, or the effort exerted when lifting a weight (Chouchourelou et al., 2006; Kozlowski & Cutting, 1977; Shim et al., 2004). Furthermore, the results of previous electrophysiological and neuroimaging studies employing whole-body PLDs stimuli suggested that these latter can convey sufficient information about movements of the human body to activate sensory-motor processes within some of the areas of the eAON, similar to those typically involved during fully visible action observation (Bonda et al., 1996; Grossman & Blake, 2002; Peelen et al., 2006; Peuskens et al., 2005; Saygin et al., 2004; Ulloa & Pineda, 2007; van Kemenade et al., 2012). Interestingly, not only human adults are able to perceive and extract information by PLDs but also human infants (Pavlova et al., 2001; Simion et al., 2008) and even monkeys (Jastorff et al., 2012).

This technique has also found important applications in the clinical field. For example, recent evidence on people with autism spectrum disorder shows that deficits are particularly evident when biological motion is needed to infer intention, action goal or emotion from an observed action (Federici et al., 2020). The authors also propose that an analysis of the distinct levels of biological motion by means of PLDs technique, aimed at assessing brain processing of specific spatio-temporal components of the action, may be useful for deepening the understanding of this syndrome, laying the foundation for future clinical investigations in early infancy.

Most of neuroimaging evidence obtained during observation of PLDs concerns only whole-body movements, while very few studies focused on upper limb actions. One of them, by Lestou et al. (2008) investigated with functional magnetic resonance imaging (fMRI) the brain activation obtained during observation of upper limb PLDs actions (such as knocking, lifting, waving, and throwing) combined with their mental simulation. The study showed the activation of several areas including IPL, SPL, PMv, as compared to mere observation. However, the results of pure observation are not explicitly reported. Another study, performed by Quadrelli et al. (2019) with electroencephalographic technique, showed that in 9-month infants, it is possible to elicit an attenuation of alpha band activity during the observation of the reaching phase of a silhouette of a grasping action, suggesting that this type of stimulus is able to elicit a motor resonance mechanism. Altogether, these studies suggest that the motor system can be activated by observation of impoverished version of upper limb actions.

The present fMRI study aimed at investigating the activation of the eAON in healthy human participants during mere observation of hand grasping actions presented in PLDs version versus fully visible version, in order to reveal, at the level of cortical activation, the importance of the decoding of kinematic features for understanding others' actions. The use of several control conditions allowed us to assess more precisely the role of eAON areas in encoding grasping action

features in both PLDs and fully visible stimuli, excluding confounding effects due to low-level visual features, motion, and contextual information. To address the specific contribution of different eAON areas in the encoding of PLDs and fully visible actions, we used a combined approach based on univariate analysis and multivariate pattern analysis (MVPA), this latter allowing to extract finer grain information. Using this latter approach, we were able to investigate whether information encoded in eAON areas during the observation of PLDs and fully visible stimuli could be disentangled with machine-learning approaches.

2 | MATERIALS AND METHODS

2.1 | Participants

Twenty-three healthy human volunteers (12 female; mean age 25.5 years; range 21–32 years) with no history of neurological or orthopedic disorders, and of drug or alcohol abuse, participated in the study. All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Informed consent was obtained in accordance with Helsinki declaration. The study was approved by the local ethics committee (Comitato Etico Area Vasta Emilia Nord – AVEN; code NEUROIMAGE_UNIPR).

2.2 | Stimuli

Experimental stimuli consisted of video events showing grasping actions either fully visible (fully visible set) or as point-light displays (point-light displays set). All videos lasted 2 s.

2.2.1 | Fully visible set

In this set, videoclips showed a fully visible (FV) human right hand grasping an object (FV_Grasp) (Figure 1a). Videos were recorded from a lateral perspective (90° angle) in a well-lit environment on a neutral background by means of a digital HD camera (©GoPro, Inc., USA) with a frame rate of 25/s and a resolution of 1920 × 1080 p. In order to use a wide set of stimuli, we recorded four grasping actions each performed with a different grip (whole hand, five-fingers grip, three-fingers grip, and precision grip) congruent with the size of the object to be grasped, for a total number of 16 grasping videos. The first static frame of each grasp stimulus was used as a control condition (FV_Static).

To control for possible effects due to velocity and direction of motion of the observed stimuli, a box-like stimulus moving in the same direction at linear velocity, with colors and size/shape similar to those of a human arm, was created using an image editing software (©Affinity Photo v 1.6.7, Serif Europe Ltd) and animated in Final Cut Pro X (v 10.5.1, Apple Inc.) (FV_Box). In order to create FV_Box videos, first we carried out 2D kinematic analysis (©Tracker v 5.1.5, 2020,

Douglas Brown) to calculate the velocity profile of the wrist in each grasping video (Figure 1c). Then, we used these velocity data to animate the box, maintaining the same direction and mean velocity of each FV_Grasp video. Thus, the experimental grasping videos and the box control condition were matched for movement direction and mean velocity.

A further control condition was the scrambled version of FV_Grasp videos (FV_Scrambled). This was realized using an ad hoc script capable of dividing each frame of the experimental stimuli in squares of 10 × 10 pixels and randomize the position of each square in each frame, so that the basic visual features (e.g., contrast, luminance, and color) were the same of the original video, but the contents of the latter were no longer recognizable (Figure 1a).

2.3 | Point-light displays set

In this set, videoclips showed PLDs stimuli created starting from FV stimuli in order to reduce at minimum the pictorial aspects of the stimulus, keeping the same kinematic features (Figure 1b). In order to accurately match joints trajectory, grasping stimuli were realized by tracking the movement of the hand joints for each frame, using the tracker feature included in Motion software (v 5.5.1, Apple Inc.) and by placing, on each joint, a white point of 9 px diameter (PLD_Grasp). The first static frame of each PLD grasp stimulus was used as a control condition (PLD_Static) (Figure 1b). The PLDs box-like stimuli were created by overlapping a series of white points (ø 9 px) to the edges of the FV box, forming a silhouette of the box shape (PLD_Box).

The same ad hoc script used for FV stimuli was used to randomize the position of each square in each frame of the PLDs grasping stimuli so that the point configuration was not recognizable in spite of the same amount of visual basic information (PLD_Scrambled).

2.4 | fMRI task

Before fMRI scanning, participants were briefly informed about the scanning procedure in order to help them to familiarize with the experimental environment and setting. During MR scanning, they laid supine in the bore of the scanner in a dimly lit environment.

The experiment was performed in a single imaging acquisition session divided in six functional runs each lasting 4 min and 56 s (148 volumes) during which participants had to observe the video stimuli presented by means of a digital goggles system (Resonance Technology, Northridge, CA) (60 Hz refresh rate) with a resolution of 800 horizontal pixels × 600 vertical pixels and horizontal eye field of 30°. To dampen scanner noise, sound-attenuating (30 dB) headphones were employed. During the whole imaging session, a white cross was presented in the center of the screen, and participants were instructed to fixate it. Each run was acquired using a block paradigm. Each block lasted 20 s and comprised 10 videos of
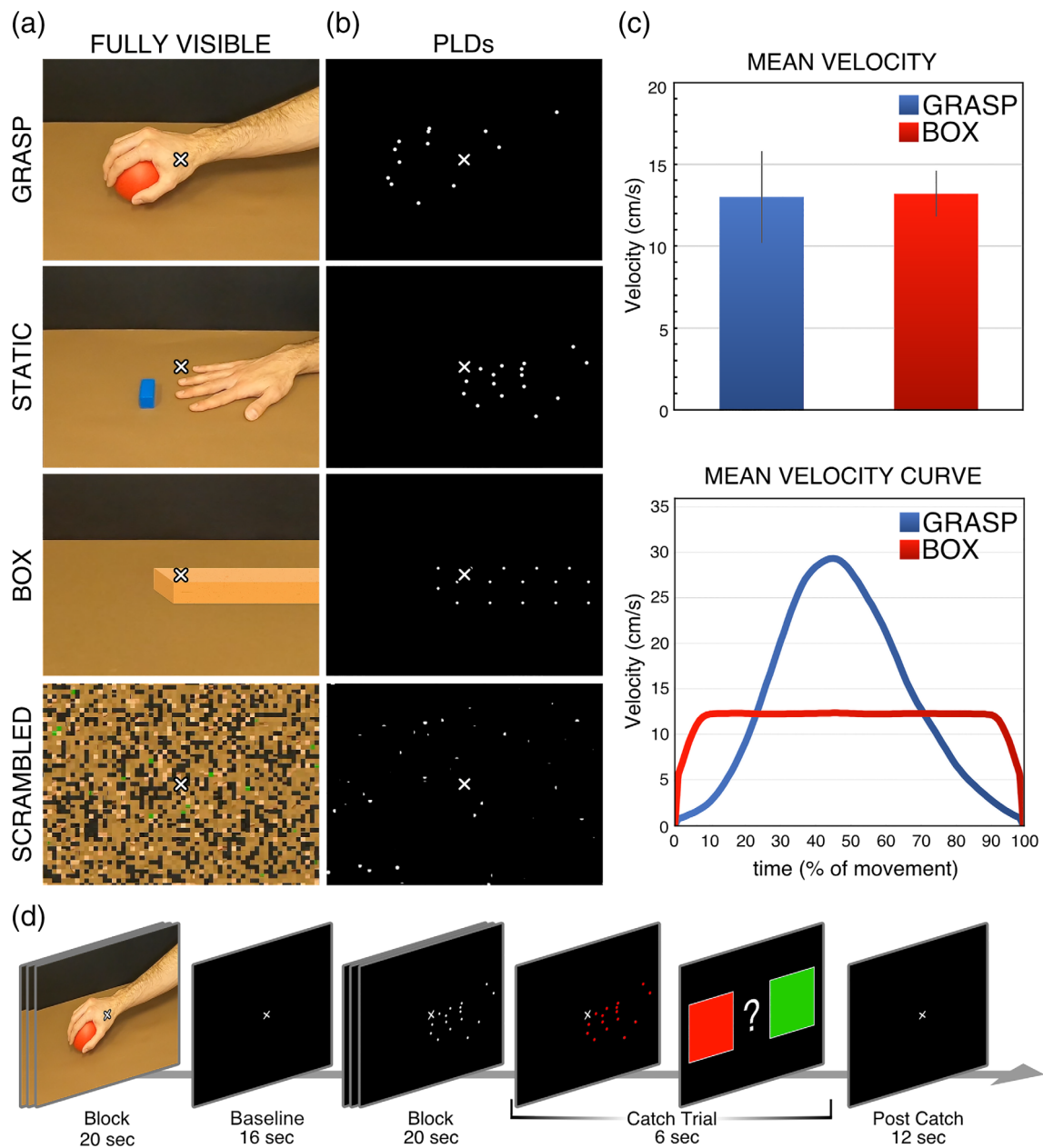
FIGURE 1 Stimuli and paradigm. Illustrations of fully visible (a) and point-light displays (b) sets including experimental and control stimuli. (c) Top: Mean wrist velocity of all grasping actions and matched mean box velocity. Bottom: Velocity profiles of the wrist and the box, the latter moving linearly. (d) Action observation paradigm, presented in six functional runs, formed by independent blocks of 20 s, each consisting of 10 randomly presented videos of the same condition, alternated with a baseline period of 16 s

the same experimental or control condition. During a single run, a total of eight blocks were presented, one for each condition. A total of 48 blocks were presented during the whole imaging session, for a total of 480 trials (60 trial per condition). Blocks were interleaved by a fixation no-videoclip event lasting 16 s used as baseline during which only the fixation cross at the center of a black background was visible.

In order to monitor participants' attention to the visual stimuli, catch trials were presented, in 25% of the blocks, equally distributed among all conditions. During the catch trial, participants observed a video stimulus (2 s duration), whose color was altered by applying a color correction filter (blue, red, and green) (Final Cut Pro X 10.5.1, Apple Inc.), after which they had to indicate the main color of the stimulus, using a response pad, by selecting one of the two options presented on the screen (4 s time window). On all valid given answers, participants were accurate 99% of the times. In order to remove potential signal artifacts due to the hand movement, a 12 s signal denoising period (post-catch), in which participant had to remain still, followed the attention task (Figure 1d).

2.5 | fMRI data acquisition

Both anatomical T1-weighted and functional T2*-weighted MR images were acquired with a 3 T General Electric scanner (MR750 Discovery) equipped with an 8-channel receiver head-coil. Functional volumes were acquired with the following parameters: 40 axial slices of functional images covering the whole brain acquired in an interleaved bottom-up order using a gradient-echo echo-planar imaging (EPI) pulse sequence, slice thickness = 3.0 mm, interslice gap = 0.5 mm, $64 \times 64 \times 37$ matrix with a spatial resolution of $3.5 \times 3.5 \times 3.5$ mm, TR = 2000 ms, TE = 30 ms, FOV = 205×205 mm^2, flip angle = 90°, in plane resolution = 3.2×3.2 mm^2. A morphological 3D T1-weighted (Bravo_Mik) volume was acquired as anatomical reference. Its acquisition parameters were 192 slices, 512×512 matrix, spatial resolution $0.9 \times 0.5 \times 0.5$ mm, TR = 9700 ms, TE 4 ms, FOV = 252×252 mm, flip angle 90°.

2.6 | fMRI data analysis

2.6.1 | Data preprocessing and analysis

Processing was carried out using SPM12 (Wellcome Department of Imaging Neuroscience, University College, London, UK; http://www.fil.ion.ucl.ac.uk/spm) on MATLAB R2017a (The Mathworks, Inc.). The first four volumes of each run were discarded to allow T1 equilibration so that magnetization could reach a steady state. For each participant, all volumes were preprocessed using the same standard pipeline. Images were spatially realigned to the first volume of the first functional run and un-warped to correct for between scan motion, and slice timing corrected considering slice acquisition order. Spatial transformation parameters derived from the segmentation and spatial normalization of the anatomical T1-weighted images to the Montreal Neurological Institute (MNI) space were then applied to the realigned EPIs and re-sampled in $2 \times 2 \times 2$ mm^3 voxels using a fourth degree B-spline interpolation in space. Lastly, all functional T2*-weighted volumes were spatially smoothed with an 8-mm full-width half-maximum isotropic Gaussian kernel (FWHM).

Data were analyzed using a random-effects model (Friston et al., 1999), implemented in a two-level procedure. In the first-level analysis, single-subject fMRI time series were modeled using the general linear model. The design-matrix included the onsets and the durations of all experimental and control condition as well as the response of catch trial conditions for each of the six functional runs. Each predictor (*FV_Grasp, FV_Static, FV_Box, FV_Scrambled, PLD_Grasp, PLD_Static, PLD_Box, PLD_Scrambled*), including 10 consecutive videos in a block, was modeled as a single epoch lasting 20 s. The catch trials were modeled as consecutive blocks lasting 18 s (*Catch_Trial*; 2 s color altered stimulus, 4 s explicit response, plus 12 s post-catch signal denoise period). Rest periods between blocks were considered as implicit baseline. Contrasts between each experimental/control condition versus implicit baseline were calculated. Specific effects were

tested using *t* statistical parametric maps, with degrees of freedom corrected for nonsphericity at each voxel. In the second level group-analysis, corresponding *t*-contrast images of the first-level conditions, except for *Catch-Trial*, were entered into a flexible ANOVA with sphericity-correction for repeated measures (Friston et al., 2002). In particular, since we did not want to test for all possible main effects and interactions, we modeled two factors specifying our scans and conditions, corresponding to: (a) subjects, modeling participants variability; (b) conditions, modeling task effects.

Within this model, in order to exclude the confounding effects due to the context, direction, mean velocity of the movement and low-level visual processing, the activation maps resulting from the contrast between experimental condition *FV_Grasp* and its corresponding control conditions (*FV_Static, FV_Box, FV_Scrambled; collectively termed as FV_Ctrls*) were calculated. In addition, *PLD_Grasp* condition was contrasted with its corresponding PLDs controls (*PLD_Static, PLD_Box, PLD_Scrambled; collectively termed as PLD_Ctrls*). The rationale for this type of analysis is that experimental stimuli were embedded with information about (a) the context in which the action was performed, (b) kinematic aspects, for example, velocity and movement direction, and (c) low-level visual characteristics. All these features are concurrent and collectively contribute to the encoding of the grasping action. Thus, the subtraction of these characteristics altogether allowed us to better assess which areas mostly contribute to the encoding of grasping actions, after excluding confounding effects.

Significant brain activations shared between FV and PLD contrasts, were assessed by means of a conjunction analysis (*FV_Grasp* vs. *FV_Ctrls* $\wedge$ *PLD_Grasp* vs. *PLD_Ctrls*) revealing cortical areas involved in both experimental conditions. In order to highlight general significant cortical activation for FV or PLDs contrasts, a global analysis was conducted (*FV_Grasp* vs. *FV_Ctrls* and *PLD_Grasp* vs. *PLD_Ctrls*) (see Section 2.6.3). Statistical inference was drawn at a voxel level, corrected with family-wise error (FWE) with a threshold of *p* < .01.

2.6.2 | Lateralization index analysis

In order to assess activation pattern distribution between the two hemispheres, we performed a lateralization analysis using LI-toolbox (Wilke & Lidzba, 2007). This index is computed, first, by calculating the number of voxels surviving a statistical threshold in each hemisphere and then applying the formula $(\Sigma_{left} - \Sigma_{right})/(\Sigma_{left} + \Sigma_{right})$, yielding a lateralization index value ranging from −1 to 1. Positive values correspond to left lateralization, while negative values correspond to right lateralization.

Contrast images of PLDs and FV grasping conditions from each subject were entered as input in the toolbox and one default value threshold ($t = 3$) was applied to all images. An interhemispheric exclusive mask of ±5 mm was applied, masking out the midline. Finally, the lateralization index was computed following the previously described formula.

2.6.3 | Region of interest selection and analysis

To examine potential differences between blood oxygen level dependent (BOLD) activations in FV and PLDs conditions, a region of interest (ROI) analysis was performed. ROIs were selected using an anatomical approach starting from the group level results of the global analysis. Brain areas revealed by global analysis were determined by means of cytoarchitectonic probabilistic maps of the human brain using SPM-Anatomy toolbox v. 2.1 (Eickhoff et al., 2005). In order to determine which anatomical areas were included in the activation map, global analysis pattern was overlayed onto cytoarchitectonic probabilistic maps and the cytoarchitectonic area corresponding to the maxima of each cortical cluster, plus adjacent areas with a high assigned probability, were selected. This allowed the selection of 19 cortical ROIs, located in left hemisphere (see Figures 4–6, for a visualization of ROI positions).

ROIs masks were created using Anatomy toolbox v. 3.0 (Eickhoff et al., 2005), Automated Anatomical Labeling atlas (AAL) included in the WFU-PickAtlas Toolbox (https://www.nitrc.org/projects/wfu_pickatlas; Maldjian et al., 2003]), the Human Motor Area Template (HMAT; http://lrnlab.org; Mayka et al., 2006) and Brainnetome Atlas (https://www.nitrc.org/projects/bn_atlas; Fan et al., 2016). To preserve only the voxels within the activation pattern, we used a masking procedure provided by MRIcron software (https://www.nitrc.org/projects/mricron). In addition, a spherical ROI built in the white matter (CTRL_WM; $r = 4.5$ mm, $x = -20$, $y = 42$, $z = 2$) was used as control and created by means of MarsBaR software for SPM (http://masbar.sourceforge.net/).

The average BOLD signal change across all significant voxels was extracted separately in each ROI using the SPM Rex Toolbox (http://web.mit.edu/swg/rex). BOLD signal change was compared between all conditions by means of repeated measures ANOVAs. Significant differences were assessed with post hoc comparisons computation by using paired-sample t-test with *Bonferroni* correction for multiple comparisons.

2.6.4 | Multivoxel pattern analysis

To detect all fine information included in the fMRI data patterns, MVPA was conducted on the un-smoothed normalized T2* functional brain images using Pattern Recognition for Neuroimaging Toolbox (PRoNTo v.2.1; Schrouff et al., 2013), a MATLAB (The MathWorks Inc.) based toolbox.

The intensity value of each voxel (*feature*) is represented as a series of points in a multidimensional space (*feature space*). Then a classifier algorithm is trained and employed to find the optimal separating boundary between the *features*, associated with a categorical label corresponding to the experimental conditions, in the feature space. After training and testing the classification model, which consists of applying a trained model to the tested set of data, the classifier returns a predicted label for different brain patterns. The performance of the classifier in discriminating between

different conditions is assessed on new data not previously used to train it.

In order to perform MVPA analyses, the experimental design elements, that is, *labels*, *onsets*, *duration* and *number* of each block and *interscan interval* were specified. The un-smoothed normalized T2* functional brain images belonging to the experimental conditions (*FV_Grasp* and *PLD_Grasp*) for each subject were selected (1380 volumes total) and then a first-level mask, including only voxels containing relevant features and discarding those with nonrelevant information, that is, voxels outside the brain, was applied to the data. Successively, a linear kernel included in PRoNTo toolbox, was used to compute a similarity matrix. The kernel function, by calculating the dot product of each *feature* in pairs, returns a value characterizing the similarity between each pair, creating a kernel matrix of the feature space. A first degree polynomial detrending was applied to the data since fMRI data represent continuous temporal series. As a second-level masks, the same ROIs used in the univariate analysis (See Section 2.6.3) were entered, in order to focus on specific sets of *features*.

A binary classification model was computed using a support vector machine algorithm (SVM) which, using the similarity matrix previously determined, computes a hyperplane that splits the feature space, maximizing the margin that separates values belonging to the two experimental conditions, finally extracting the weight vectors running perpendicularly to the hyperplane. For each binary classification model, *FV_Grasp* functional images were assigned to Class 1 and *PLD_Grasp* to Class 2.

The performance of the classifier and its ability to generalize the results of its computations on an independent nontrained dataset was assessed by means of a leave one subject out cross validation scheme. Specifically, the entire dataset was separated into two sets: one used for training and the other for test. The number of folds in which data were partitioned was equal to the number of subjects. The set of data used for training was equal to the number of subjects minus one. The learned function was then used to predict the labels on the remaining unused subject's data. Further operations were applied to the data, including sample averaging within subjects, mean centering the features using training data and dividing the data vectors by their Euclidean norm. In order to estimate the model p value, 1000 permutations were run so that the model was retrained by the specified number of times. Model accuracy and the area under curve were computed to assess the model performance.

3 | RESULTS

3.1 | Univariate analysis

Figure 2 shows cortical activation maps, FWE corrected at a voxel level with a significance threshold of $p < .01$, overlaid on an MNI template. Contrast between *FV_Grasp* versus baseline shows a large activation pattern in the left cerebral hemisphere, including occipito-temporal (pMTG, pFG) and occipito-parietal (SPOC) clusters, superior

and inferior parietal areas (SPL, IPS, IPL), both dorsal and ventral sector of premotor cortex (PMd, PMV), pars triangularis of the IFG, SMA. In the right cerebral hemisphere, clusters were roughly symmetrical but less extended than in the left hemisphere. No significant activation of IPL and SMA was found in the right hemisphere. Subcortical clusters were localized in the thalamus at the level of the Pulvinar (bilaterally), and in cerebellum, including bilateral lobules VI, Crus I, VIIb, and VIIIb (Figure 2a).

The contrast between *PLD_Grasp* versus baseline shows a pattern rather similar to that of *FV_Grasp* condition, although the activation in the former condition was more bilateral than in the latter one. Further differences consist in the absence of activation of left IFG, pars triangularis, the presence of a small cluster in the posterior sector of the left middle cingulate cortex (pMCC) and the bilateral activation of inferior parietal cortex. Subcortical activations were comparable to those of *FV_Grasp* (Figure 2b).

The activation map corresponding to the contrast between *FV_Grasp* versus all fully visible control conditions (*FV_Static*, *FV_Box*, and *FV_Scrambled*) revealed activated clusters in areas pMTG, pFG, SPOC, SPL, PMd, bilaterally, although more extended in the left hemisphere. Some clusters were fully lateralized to the left hemisphere, such as those in PMv, IPL, and pMCC. Subcortical clusters were localized in left pulvinar, cerebellar lobules VI and Crus I bilaterally, and right lobule VIIb (Figure 2a; Figure 2c).

Cortical activation map derived from the contrast between *PLD_Grasp* versus PLDs control conditions (*PLD_Static*, *PLD_Box*, *PLD_Scrambled*) revealed bilateral clusters in areas pMTG, pFG, SPOC, SPL, IPS, IPL, PMd, and PMv. A significant cluster was also present in the left pMCC. Additional subcortical structures included the lateral sectors of cerebellar lobules VI, Crus I and VIIb bilaterally (Figure 2d). For statistical details about MNI coordinates of activation peaks see supplementary table 1.
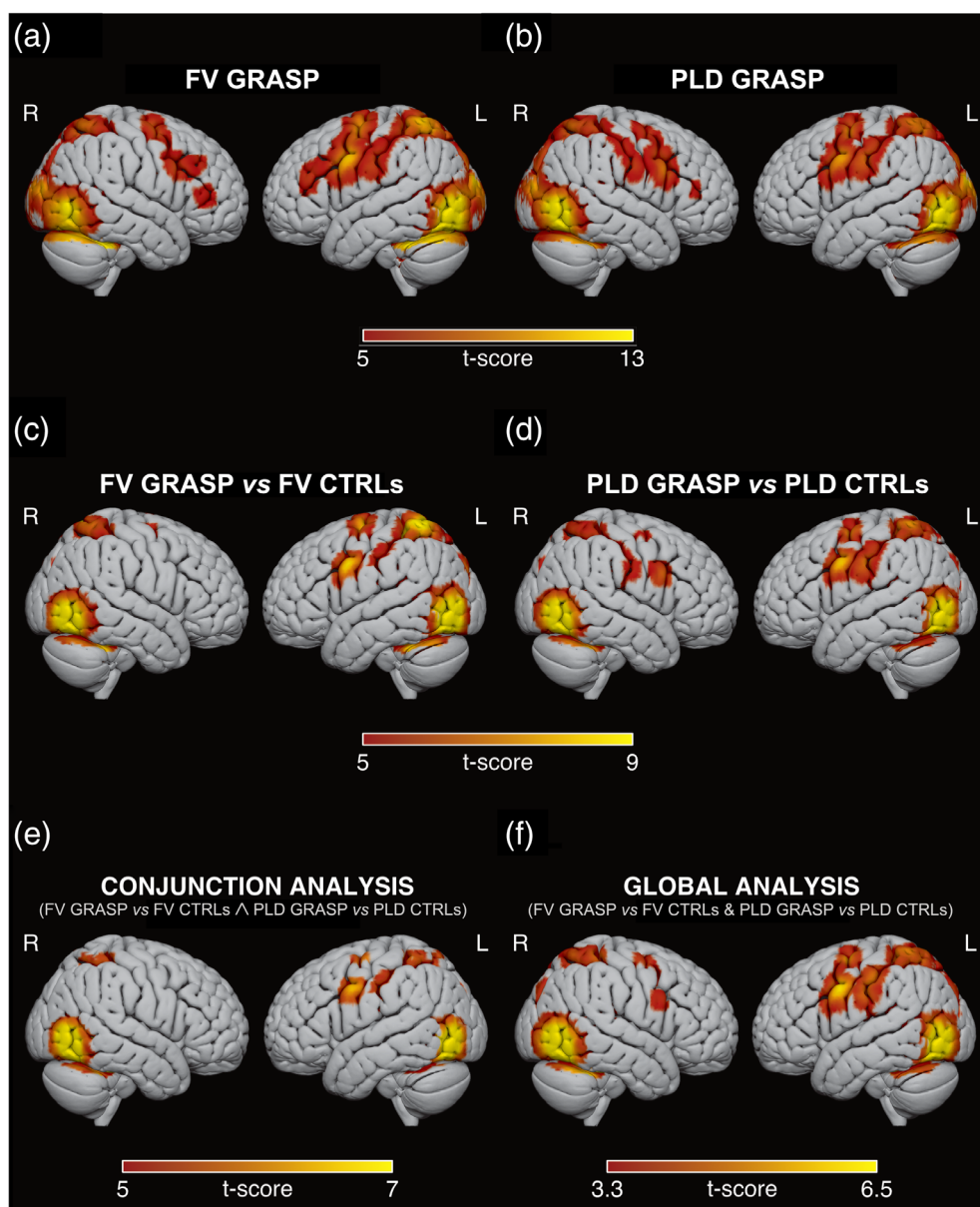

FIGURE 2 Cortical activations projected onto a 3D MNI152 Brain template (Surfice; https://www.nitrc.org/projects/surfice/). (a) Contrast between the observation of *FV_Grasp* and baseline. (b) Contrast between the observation of *PLD_Grasp* and baseline. (c) Contrast between *FV_Grasp* and all corresponding controls. (d) Contrast between *PLD_Grasp* and its respective control conditions. Conjunction (e) and global (f) analysis between *FV_Grasp* and *PLD_Grasp* versus their corresponding controls

Figure 2e shows the conjunction analysis between *FV_Grasp* and *PLD_Grasp*, versus their corresponding controls. This analysis revealed shared bilateral clusters in pMTG, pFG, and SPL, and left-lateralized clusters in SPOC, IPL, IPS, PMd, and PMv. Subcortical clusters were localized in cerebellar lobules VI and VIIb, bilaterally, while the cluster in Crus I was more evident on the right.

The cortical activation map emerging from global analysis reveals clusters significantly active in either *FV_Grasp*, *PLD_Grasp*, or in both conditions, each versus its respective controls (Figure 2f). The pattern obtained with this analysis was used to define the areas in which to perform a subsequent ROI analysis.

3.2 | Lateralization analysis

Lateralization index computation showed a significant difference ($p < .001$) between PLDs (mean = 0.04, s.d. = 0.15) and FV (mean = 0.21, s.d. = 0.16) grasping conditions, indicating a more bilateral pattern for observation of PLDs grasping (Figure 3).

3.3 | ROI analysis

Comparisons between BOLD signal change in the different conditions were also carried out by means of a ROI analysis, in cortical areas chosen following an anatomical approach (see Section 2.6.3). For each ROI, BOLD signal change was compared between conditions by means of repeated measure ANOVA and significant differences were assessed with post hoc comparisons by using paired-sample *t*-test with Bonferroni correction.

The analysis carried out revealed a significant effect for all considered ROIs except for the CTRL_WM ROI ($F_{(1,7)} = 0.75$, $p < .63$, $\eta2 = 0.03$). We clustered the remaining cortical ROIs using a functional criterion. Significant effects between conditions were found in the following:

(a) *Visual* ROIs: V4_hOc4la ($F_{(1,7)} = 82.72$, $p < .001$, $\eta2 = 0.79$), V5_hOc5 ($F_{(1,7)} = 82.12$, $p < .001$, $\eta2 = 0.79$), pMTG ($F_{(1,7)} = 32.90$, $p < .001$, $\eta2 = 0.60$), FG2 ($F_{(1,7)} = 41.64$, $p < .001$, $\eta2 = 0.65$), FG4 ($F_{(1,7)} = 35.16$, $p < 0.001$, $\eta2 = 0.62$), SPOC ($F_{(1,7)} = 23.45$, $p < .001$, $\eta2 = 0.52$) (Figure 4).

(b) *Parieto-premotor* ROIs: SPL_5L ($F_{(1,7)} = 32.41$, $p < .001$, $\eta2 = 0.60$), SPL_7A ($F_{(1,7)} = 42.51$, $p < .001$, $\eta2 = 0.66$), SPL_7PC ($F_{(1,7)} = 38.82$, $p < .001$, $\eta2 = 0.64$), aIPS_IP1 ($F_{(1,7)} = 16.65$, $p < .001$, $\eta2 = 0.43$), aIPS_IP3 ($F_{(1,7)} = 25.60$, $p < .001$, $\eta2 = 0.54$], IPL_PFt ($F_{(1,7)} = 16.93$, $p < .001$, $\eta2 = 0.43$), SMG ($F_{(1,7)} = 22.68$, $p < .001$, $\eta2 = 0.51$), PMd ($F_{(1,7)} = 23.58$, $p < .001$, $\eta2 = 0.52$), PMv ($F_{(1,7)} = 19.97$, $p < .001$, $\eta2 = 0.48$) (Figure 5).

(c) *Somatomotor* ROIs: PSC_2 ($F_{(1,7)} = 21.85$, $p < .001$, $\eta2 = 0.50$) PSC_3b ($F_{(1,7)} = 13.60$, $p < .001$, $\eta2 = 0$), pMCC ($F_{(1,7)} = 17.48$, $p < .001$, $\eta2 = 0.44$) (Figure 6).

Post hoc tests showed no significant difference between experimental conditions (*FV_Grasp* and *PLD_Grasp*) in all considered ROIs. Instead, a significant difference between FV experimental condition and all its corresponding control conditions (*FV_Grasp* vs. *FV_Static*, *FV_Box*, *FV_Scrambled*) was observed in: V4_hOc4la ($p < .001$), V5_hOc5 ($p < .001$), pMTG ($p < .001$), SPOC ($p < .001$), FG2 ($p < .001$), FG4 ($p < .05$), SPL_5L ($p < .001$), SPL_7A ($p < .001$), SPL_7PC ($p < .001$), aIPS_IP3 ($p < .01$), IPL_PFt ($p < .01$), SMG ($p < .001$), PSC_2 ($p < .001$), PSC_3a ($p < .001$), PSC_3b ($p < .001$), PMv ($p < .01$), PMd ($p < .001$).
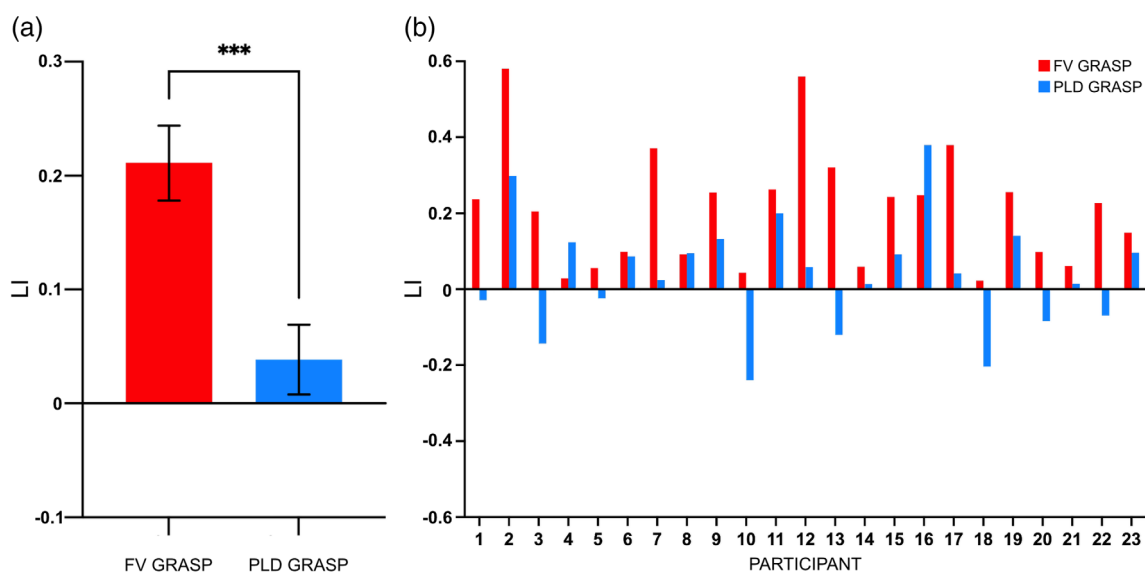

FIGURE 3 Lateralization index (LI) analysis results at group (a) and single subject (b) level. Positive values correspond to left lateralization, while negative values correspond to right lateralization. In (a), the data show a significant difference ($p < .001$) in LI between PLDs (mean = 0.04, s.d. = 0.15) and FV (mean = 0.21, s.d. = 0.16) grasping conditions. Error bars indicate the standard error of the mean
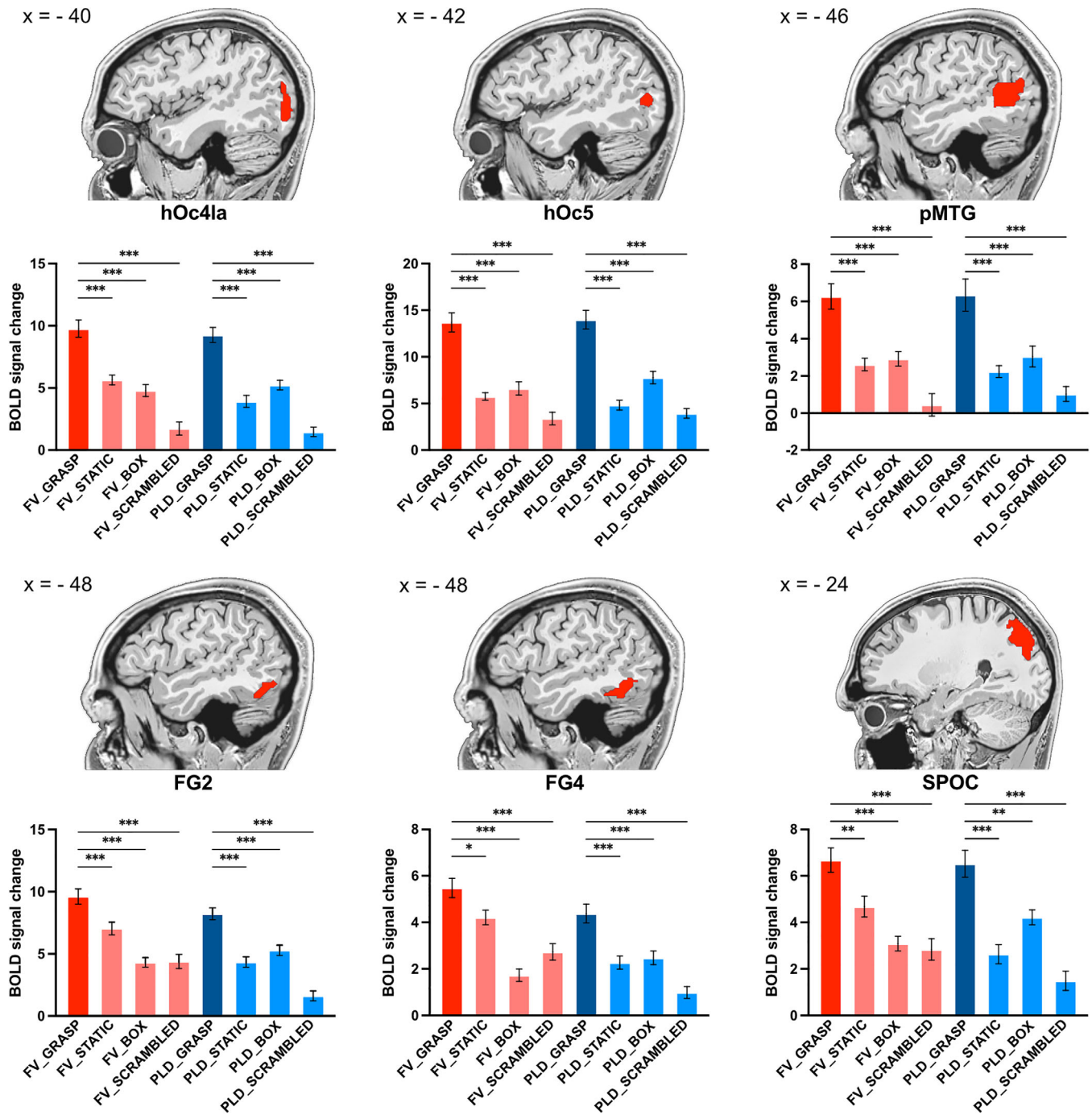
FIGURE 4 Results of the region of interest (ROI) analyses on visual areas. Histograms show the BOLD signal change in each ROI. The red colored bars refer to fully visible (FV) conditions, blue to point-light displays (PLDs) ones. Vertical lines in the histograms indicate standard error of the mean. Above each histogram, the corresponding ROI is represented on a sagittal slice of a human left hemisphere. Asterisks indicate significant effects corrected for multiple comparisons (*$p < .05$, **$p < .01$, ***$p < .001$; Bonferroni corr)

Considering PLDs conditions, the analysis revealed a significant difference between *PLD_Grasp* versus its corresponding controls (*PLD_Static, PLD_Box,* and *PLD_Scrambled*) in the following ROIs: V4_hOc4la ($p < .001$), V5_hOc5 ($p < .001$), pMTG ($p < .001$), SPOC ($p < .01$), FG2 ($p < .001$), FG4 ($p < .001$), SPL_5L ($p < .001$), SPL_7A ($p < .01$), SPL_7PC ($p < .001$), aIPS_IP3 ($p < .01$), IPL_PFt ($p < .01$),

SMG ($p < .001$), PSC_2 ($p < .001$), PSC_3a ($p < .01$), PSC_3b ($p < .001$), PMv ($p < .001$), PMd ($p < .001$).

In pMCC, no significant difference was found between both *FV_Grasp* versus *FV_Box* ($p = 1$; n.s.) and *PLD_Grasp* versus *PLD_Box* ($p = .09$; n.s.), in aIPS_IP1 between *FV_Grasp* versus *FV_Static* ($p = .06$; n.s.) and *PLD_Grasp* versus *PLD_Box* ($p = .07$; n.s.). Note that
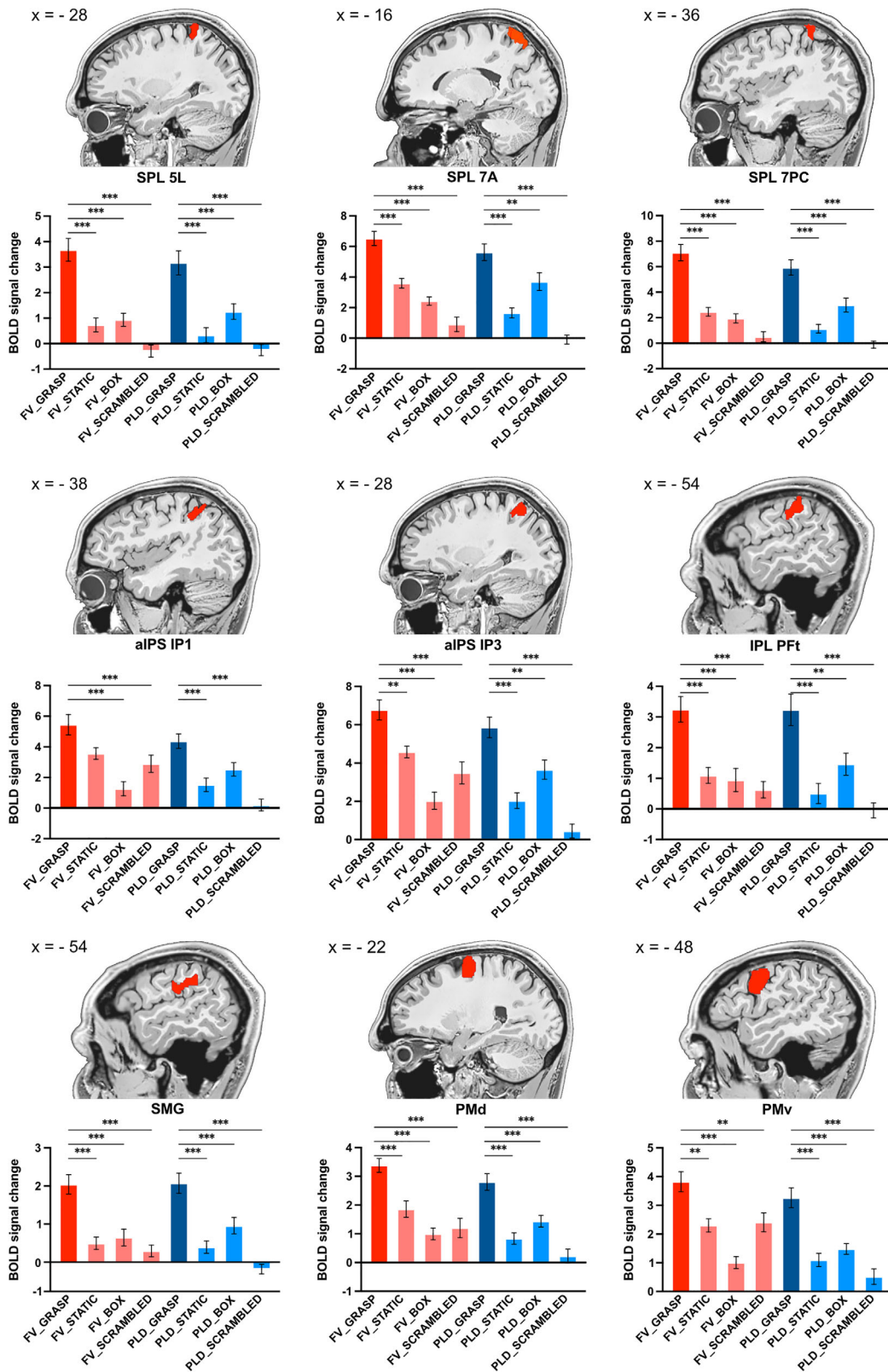
FIGURE 5 Results of the region of interest (ROI) analyses on parieto-premotor areas. Histograms show the BOLD signal change in each ROI. The red colored bars refer to fully visible (FV) conditions, blue to point-light displays (PLDs) ones. Vertical lines in the histograms indicate standard error of the mean. Above each histogram, the corresponding ROI is represented on a sagittal slice of a human left hemisphere. Asterisks indicate significant effects corrected for multiple comparisons (*$p < .05$, **$p < .01$, ***$p < .001$; Bonferroni corr).
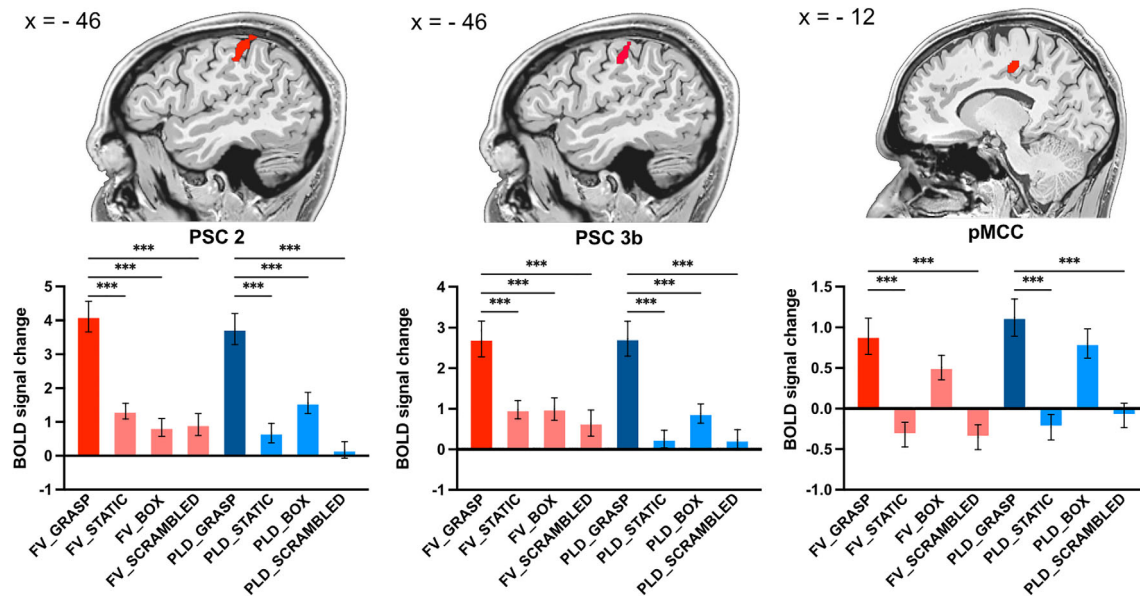
FIGURE 6 Results of the region of interest (ROI) analyses on somatomotor areas. Histograms show the BOLD signal change in each ROI. The red colored bars refer to fully visible (FV) conditions, blue to point-light displays (PLDs) ones. Vertical lines in the histograms indicate standard error of the mean. Above each histogram, the corresponding ROI is represented on a sagittal slice of a human left hemisphere. Asterisks indicate significant effects corrected for multiple comparisons (*p < .05, **p < .01, ***p < .001; Bonferroni corr).

in this latter ROIs p values were close to significance threshold, although not reaching it.

3.4 | Multivariate analysis

The analysis, performed by computing a binary classification model using an SVM algorithm and performing the permutation testing for 1000 times on the un-smoothed normalized T2* functional brain images belonging to *FV_Grasp* and *PLD_Grasp* experimental conditions for each subject, revealed significant decoding accuracy in the following ROIs: FG4 (model accuracy = 69.57%, p < .05), SPOC (model accuracy = 69.57%, p < .05), SPL_7A (model accuracy = 78.26%, p = .001), SPL_7PC (model accuracy = 78.26%, p < .01), PMd (model accuracy = 67.39%, p < .05) and PMv (model accuracy = 73.91%, p = .01) (Figure 7). No significant above threshold model accuracy was found in the remaining ROIs (for details see Table 1).

In order to evaluate the contribution of the voxels to the decision function in each ROI, we computed the weight maps of the classification models for both *FV_Grasp* and *PLD_Grasp* conditions (Figure 8). A voxel's weight parameter reflects the contribution of that voxel to the discrimination process. Since all voxels with a value different from zero contribute to the function value, we represented the intensity of each weight with a color grading: colder colors for the weights with an intensity <0; warmer ones for intensities >0. Weights with a positive value tend to move the classification boundaries toward class 1 (*FV_Grasp*), on the contrary those with a negative one, toward class 2 (*PLD_Grasp*).
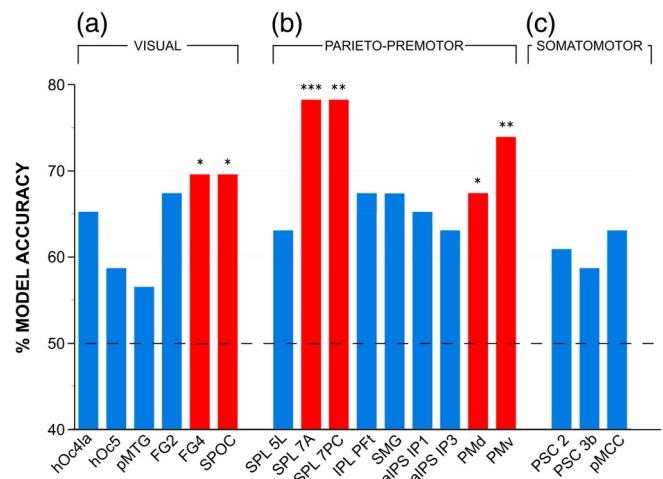

FIGURE 7 Results of the Multivoxel pattern analysis (MVPA). Histograms show the percentage of model accuracy in each region of interest (ROI) clustered in visual, parieto-premotor and somatomotor areas. The dotted line represents the chance level. Red colored bars show the ROIs with a significant (*p < .05, **p < .01, ***p < .001) model accuracy assessed by means of a permutation testing (n° permutation = 1000)

4 | DISCUSSION

In the present fMRI study, healthy participants observed hand grasping actions performed by a fully visible human hand or PLDs representation of it. The results show that (a) kinematic information conveyed by observation of PLDs hand grasping action elicits activation of the eAON; (b) the activation pattern is more bilateral during observation

	ROI	Accuracy %	p value	AUC	Class 1%	Class 2%
Visual	hOc4la	65.22	.06	0.70	73.91	56.52
	hOc5	58.70	.17	0.63	65.22	52.17
	pMTG	56.52	.28	0.63	69.57	43.48
	FG2	67.39	.08	0.72	69.57	65.22
	FG4	**69.57**	**.04**	**0.81**	**69.57**	**69.57**
	SPOC	**69.57**	**.03**	**0.75**	**73.91**	**65.22**
Parieto-premotor	SPL 5A	63.04	.12	0.66	60.87	65.22
	SPL 7A	**78.26**	**.001**	**0.78**	**78.26**	**78.26**
	SPL 7PC	**78.26**	**.003**	**0.78**	**69.57**	**86.96**
	IPL PFt	67.39	.06	0.75	73.91	60.87
	SMG	67.39	.08	0.72	78.26	56.52
	aIPS IP1	65.22	.08	0.70	65.22	65.22
	aIPS IP3	63.04	.12	0.70	73.91	52.17
	PMd	**67.39**	**.04**	**0.80**	**78.26**	**56.52**
	PMv	**73.91**	**.01**	**0.75**	**86.96**	**60.87**
Somatomotor	PSC 2	60.87	.18	0.69	65.22	56.52
	PSC 3b	58.70	.24	0.67	65.22	52.17
	pMCC	63.04	.14	0.60	78.26	47.83
	CTRL WM	54.35	.44	0.64	60.87	47.83

TABLE 1 MVPA detailed results of the binary support vector machine classification models

Note: Model accuracy and the area under curve (AUC) were computed to assess model performance. Significant values are indicated in boldface.

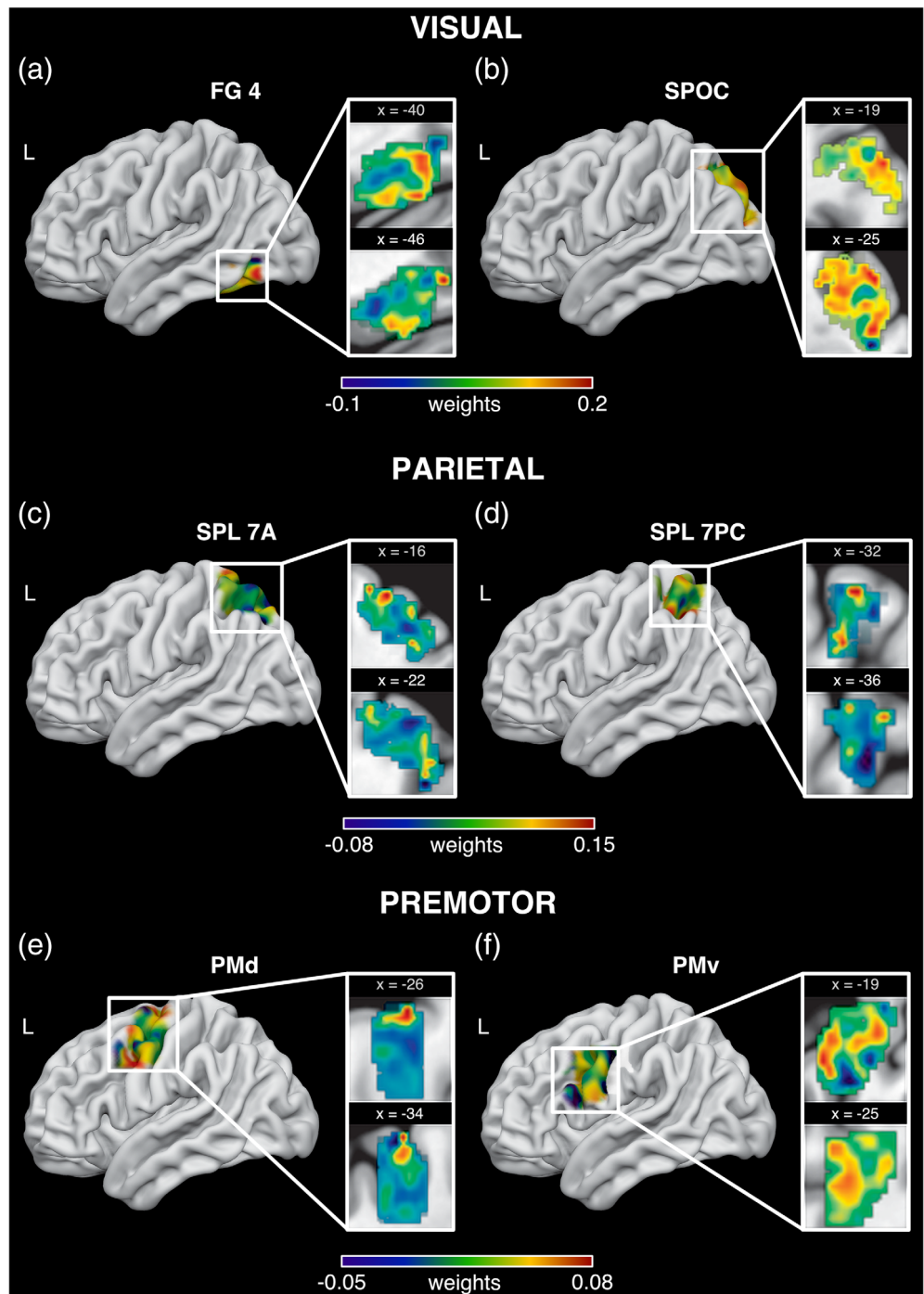Abbreviation: MVPA, multivoxel pattern analysis.

of PLDs stimuli than during observation of fully visible grasping, that is lateralized to the left hemisphere; (c) activation, assessed within multiple ROIs, is comparable between the experimental conditions; and (d) visual, parietal, and premotor cortex discriminate between the two versions of grasping actions with significant decoding accuracy.

4.1 | Brain activations during observation of PLDs and fully visible grasping actions

A large body of studies demonstrated that observation of hand grasping actions recruits a bilateral cortical network of occipito-temporal, parietal, and premotor areas belonging to the eAON (Caspers et al., 2010; Hardwick et al., 2018). It is well established that areas of this network code the goal of an observed action. However, activation within eAON could also involve the processing of low-level visual characteristics, motion aspects (e.g., linear velocity and motion direction), as well as the elaboration of the target object and the context in which the action is performed. To exclude all these effects, leaving only those features that still allow to decode action goal, we introduced control conditions, thus subtracting potential confounds from both PLDs and fully visible experimental conditions. The contrast between PLDs grasping condition and all its controls revealed a bilateral activation pattern including occipito-temporal, posterior parietal, and premotor areas known to be involved in the processing of observed actions, corresponding to the eAON. This demonstrates

that the information conveyed by PLDs hand grasping stimuli is enough to elicit activation in eAON, suggesting the involvement of a motor resonance mechanism similar to that elicited by fully visible actions. Thus, the observation of a visually impoverished grasping performed with a specific effector, in our case a right hand, is sufficient to activate areas coding action goal even when low-level visual characteristics (contrast, luminance), motion direction and velocity, as well as static hand-shape pattern, are excluded. Therefore, the remaining biological kinematic information is still able to elicit this mechanism. Neuroimaging studies using observation of PLDs versions of whole-body complex movements revealed an activation of some areas within the eAON (Beauchamp et al., 2003; Grossman & Blake, 2002; Peelen et al., 2006; Peuskens et al., 2005; Saygin et al., 2004; Vaina et al., 2001). Beauchamp et al. (2003), comparing observation of fully visible whole-body human actions with PLDs ones, showed that inferior temporal cortex and fusiform gyrus were more strongly activated by fully visible videos than by PLDs. Studies focused on the perception of whole-body PLDs, comparing biological with nonbiological motion, consistently showed activations in temporal and occipital areas (Grossman & Blake, 2002; Peelen et al., 2006; Peuskens et al., 2005). On the other hand, there are studies showing the recruitment of parietal areas during observation of whole-body biological PLDs stimuli, in particular of IPS and SPL (Grèzes et al., 2003; Vaina et al., 2001). The present study, that is focussed on the observation of PLDs hand grasping actions, is in line with the above findings, while adding new information, since it also

shows premotor cortex activations and a more extended parietal activation pattern, which includes IPL, part of human parieto-premotor MNS, usually recruited in both observation and execution of hand grasping actions (Caspers et al., 2010; Hardwick et al., 2018).

The activation of occipito-temporal areas, that in the present study includes V4, MT/V5, pMTG and pFG, very likely reflects the elaboration of biological motion, also in line with previous studies (Chang et al., 2018; Grossman et al., 2000; Grossman & Blake, 2002; Peelen et al., 2006; Pelphrey et al., 2005; Servos et al., 2002). Indeed,

the random motion of PLDs in the scrambled control condition and the motion of the PLDs box control stimuli, although sharing several low-level visual characteristics and motion features with the experimental stimuli, elicited a weaker activation in these areas. This suggests their tuning to biological features of motion rather than to motion in general. The activation of IPL and PMv is in line with the results of a large body of studies on action observation of fully visible stimuli (Caspers et al., 2010; Hardwick et al., 2018), thus suggesting the involvement of a common motor resonance mechanism in both PLDs and fully visible grasping stimuli. These regions can be involved

in coding action goal and specific aspects of performed acts, for example, grip type and action outcome (Binkofski et al., 1999; Errante et al., 2021; Grafton & Hamilton, 2007). The activation also involves areas within the so-called "dorsal circuit" such as PMd, SPL, and SPOC (Cavina-Pratesi et al., 2010; Filimon et al., 2007; Gallivan et al., 2009; Gazzola & Keysers, 2009), usually considered as involved in the observation, as well as in the execution of reaching motor acts. However, more recent human and monkey studies reported that the dorsal parieto-premotor circuit plays an important role in processing grasping/manipulation components (Errante & Fogassi, 2019; Nelissen et al., 2017). Noteworthy, in the present study, also the ROI analysis reveals that activation of areas within the dorsal circuit was higher during observation of grasping actions as compared to the box control condition, although in this case the stimulus was moving with the same direction and mean velocity, thus reaching the same portion of space of the PLDs hand. Therefore, the involvement of dorsal areas could reflect the processing of other features of observed grasping acts, such as hand/finger posture for grip configuration and the desired end state of an action, very likely as during the observation of fully visible grasping (Errante et al., 2021; Majdandzic et al., 2009). A similar activation pattern, even though more left-lateralized, was observed when contrasting fully visible grasping experimental condition and its controls. The eAON activation elicited by the observation of fully visible grasping stimuli is in line with previous literature on action observation (Caspers et al., 2010; Hardwick et al., 2018).

Interestingly, conjunction analysis between the two experimental conditions after subtraction of the respective control conditions indicates that a specific set of shared areas including left SPL and IPL as well as PMd and PMv are similarly activated. This suggests that although the processing of PLDs actions relies only on the available biological kinematic features; nonetheless, this latter information is sufficient to elicit in the observer a full action representation.

4.2 | Differential eAON contribution in the processing of observed PLDs and fully visible grasping actions

The more left-lateralized activation obtained during observation of fully visible grasping actions can be explained by a motor resonance mechanism that allows the observer to understand the action goal, likely grounded on a *praxic* knowledge that, in right-handed individuals, is usually left-lateralized (Biagi et al., 2010; Binkofski et al., 1999; De Renzi, 1982). The activation of a more bilateral pattern during observation of PLDs grasping actions suggests that very likely this type of processing is based on the elaboration of movement kinematic features, by the recruitment of the parieto-premotor grasping network of both hemispheres, in order to extract information about the final action goal.

Although ROI analysis reveals that in both PLDs and fully visible actions signal intensity was comparable in all considered areas, it is reasonable to suppose that spatial distribution of the activation pattern may differ between the two conditions. This was tested by means of MVPA, the results of which show different features patterns in FG, SPOC, SPL, PMd, and PMv. The classification model accuracy was statistically significant, showing that spatially distributed information can correctly disentangle, in the considered ROIs, the two classes of experimental conditions.

Differences in FG pattern may be attributable to a dissimilarity in the appearance and visual complexity of the two types of hand grasping stimuli, as well as to the presence of the object in fully visible grasping condition (Weiner & Zilles, 2016). In fact, this is a high-order extrastriate visual area, that is known to be also recruited during observation of actions performed with a visible upper limb (Hardwick et al., 2018), as well as by object-directed hand movements (Grosbras et al., 2012). Thus, visual appearance of the fully visible grasping stimuli as well as its complex visual characteristics, such as shape, color, and texture may be key factors in discriminating the two activation patterns.

Visual dorsal stream area SPOC has been reported to be recruited during both execution (Cavina-Pratesi et al., 2010; Gallivan et al., 2009) and observation (Filimon et al., 2007) of arm reaching actions, as well as of objects within reach (Gallivan et al., 2009). This region includes human V6 and V6A, the latter being involved in the visual analysis of the transport phase of reaching-grasping actions (Pitzalis et al., 2015). Although in our study, SPOC is recruited in a comparable manner in terms of BOLD intensity during observation of both PLDs and fully visible grasping actions, a possible interpretation is that differences in pattern distribution are mainly due to the prevalence, in fully visible condition, of information about arm movement and the presence of the object within reach.

The MVPA results also indicate a high-level accuracy (~80%) in decoding between PLDs and fully visible actions in SPL and in PMd. This is not surprising because, as previously described, both the dorsal and ventral parietal and premotor areas are involved in the processing of reaching but also of some features of the observed grasping, such as specific grip configuration (Errante et al., 2021; Errante & Fogassi, 2019). Here, however, the decoding accuracy in SPL and PMd could not be explained only by differences in reaching movement features or grip configuration because both PLDs and fully visible actions were matched for these characteristics. Thus, a further possible interpretation is that dorsal areas differentially encode proprioceptive information associated to fully visible hand–object interaction (Casile et al., 2010; Errante & Fogassi, 2019). In addition, it is also plausible that the more complete vision of the arm in fully visible actions elicits a more specific representation of this effector in the dorsal areas, according with their somatotopic organization.

Finally, pattern differences in PMv could be in principle attributed to the processing of the action goal and/or kinematic features of the movement. However, these variables were matched between the two main conditions. Therefore, the possible role of PMv in decoding between PLDs and fully visible grasping actions could be related to object presence only in fully visible stimuli (Grèzes et al., 2003).

5 | CONCLUSIONS

Activation of eAON evoked by PLDs stimuli, in particular, in parietal and premotor areas, demonstrates that motion features are sufficient to determine goal encoding without any confounding effect relative to the observation of contextual information. In addition, the use of machine learning methods allowed us to assess which areas of the eAON play a key role in disentangling between PLDs and fully visible stimuli and, together with data from literature, whether they encode specific features of the observed grasping action. Based on the present data, in the future, it could be interesting to investigate whether kinematic information provided by PLDs stimuli is exploited during motor learning tasks to improve some aspects of action execution such as precise hand/finger configurations.

From a clinical perspective, the present results could be useful to improve the observation-based methods for rehabilitation in patients with motor disorders (Buccino, 2014; Franceschini et al., 2012; Pelosin et al., 2010; Sgandurra et al., 2013). The implementation of PLDs stimuli in the clinical rehabilitation setting could bring improvements for a personalized therapy focused not only on the imitation of the action performed by another individual in terms of goal achievement but also in the imitation of the kinematics of the observed action, achieving a finer use of the hand.

The application of deep learning models and neural networks to identify features of biological motion has seen a rapid growth, leading to the creation of specific tools for such purposes (Insafutdinov et al., 2017; Nath et al., 2019; Toshev & Szegedy, 2014). Such machine-learning approaches have been used with whole body PLDs stimuli of several human actions, which were entered as input in complex pattern classification algorithms (Tanisaro et al., 2017) and in convolutional neural networks (Peng et al., 2021). The use of deep learning-based classification models to improve hand actions recognition, by extracting the kinematic features from PLDs actions of both healthy people and patients, could be useful in the field of human–robot interaction.

CONFLICT OF INTEREST
The authors declare no competing financial interests.

DATA AVAILABILITY STATEMENT
The data of the present study can be made available, on request, from the corresponding author.

ORCID
Settimio Ziccarelli https://orcid.org/0000-0002-1021-6986

REFERENCES
Abdelgabar, A. R., Suttrup, J., Broersen, R., Bhandari, R., Picard, S., Keysers, C., De Zeeuw, C. I., & Gazzola, V. (2019). Action perception recruits the cerebellum and is impaired in patients with spinocerebellar ataxia. *Brain*, 142(12), 3791–3805. https://doi.org/10.1093/brain/awz337

Amoruso, L., Finisguerra, A., & Urgesi, C. (2016). Tracking the time course of top-down contextual effects on motor responses during action comprehension. *The Journal of Neuroscience*, 36(46), 11590–11600. https://doi.org/10.1523/JNEUROSCI.4340-15.2016

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). FMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, 15(7), 991–1001. https://doi.org/10.1162/089892903770007380

Biagi, L., Cioni, G., Fogassi, L., Guzzetta, A., & Tosetti, M. (2010). Anterior intraparietal cortex codes complexity of observed hand movements. *Brain Research Bulletin*, 81(4–5), 434–440. https://doi.org/10.1016/j.brainresbull.2009.12.002

Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., & Freund, H. (1999). A fronto-parietal circuit for object manipulation in man: Evidence from an fMRI-study. *The European Journal of Neuroscience*, 11(9), 3276–3286. https://doi.org/10.1046/j.1460-9568.1999.00753.x

Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology*, 58, 47–73. https://doi.org/10.1146/annurev.psych.57.102904.190152

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience*, 16(11), 3737–3744. https://doi.org/10.1523/JNEUROSCI.16-11-03737.1996

Buccino, G. (2014). Action observation treatment: A novel tool in neuro-rehabilitation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 369(1644), 20130185. https://doi.org/10.1098/rstb.2013.0185

Casile, A., Dayan, E., Caggiano, V., Hendler, T., Flash, T., & Giese, M. A. (2010). Neuronal encoding of human kinematic invariants during action observation. *Cerebral Cortex*, 20(7), 1647–1655. https://doi.org/10.1093/cercor/bhp229

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage*, 50(3), 1148–1167. https://doi.org/10.1016/j.neuroimage.2009.12.112

Cavina-Pratesi, C., Monaco, S., Fattori, P., Galletti, C., McAdam, T. D., Quinlan, D. J., Goodale, M. A., & Culham, J. C. (2010). Functional magnetic resonance imaging reveals the neural substrates of arm transport and grip formation in reach-to-grasp actions in humans. *The Journal of Neuroscience*, 30(31), 10306–10323. https://doi.org/10.1523/JNEUROSCI.2023-10.2010

Chang, D. H. F., Ban, H., Ikegaya, Y., Fujita, I., & Troje, N. F. (2018). Cortical and subcortical responses to biological motion. *NeuroImage*, 174, 87–96. https://doi.org/10.1016/j.neuroimage.2018.03.013

Chouchourelou, A., Matsuka, T., Harber, K., & Shiffrar, M. (2006). The visual analysis of emotional actions. *Social Neuroscience*, 1(1), 63–74. https://doi.org/10.1080/17470910600630599

De Renzi, E. (1982). *Disorders of space exploration and cognition*. John Wiley & Sons, Inc.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91(1), 176–180. https://doi.org/10.1007/Bf00230027

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25(4), 1325–1335. https://doi.org/10.1016/j.neuroimage.2004.12.034

Errante, A., & Fogassi, L. (2019). Parieto-frontal mechanisms underlying observation of complex hand-object manipulation. *Scientific Reports*, 9(1), 348. https://doi.org/10.1038/s41598-018-36640-5

Errante, A., & Fogassi, L. (2020). Activation of cerebellum and basal ganglia during the observation and execution of manipulative actions. *Scientific Reports*, 10(1), 12008. https://doi.org/10.1038/s41598-020-68928-w

Errante, A., Ziccarelli, S., Mingolla, G. P., & Fogassi, L. (2021). Decoding grip type and action goal during the observation of reaching-grasping actions: A multivariate fMRI study. *NeuroImage*, 243, 118511. https://doi.org/10.1016/j.neuroimage.2021.118511

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T. (2016). The human Brainnetome atlas: A new brain atlas based on connectional architecture. *Cerebral Cortex*, 26(8), 3508–3526. https://doi.org/10.1093/cercor/bhw157

Federici, A., Parma, V., Vicovaro, M., Radassao, L., Casartelli, L., & Ronconi, L. (2020). Anomalous perception of biological motion in autism: A conceptual review and meta-analysis. *Scientific Reports*, 10(1), 4576. https://doi.org/10.1038/s41598-020-61252-3

Filimon, F., Nelson, J. D., Hagler, D. J., & Sereno, M. I. (2007). Human cortical representations for reaching: Mirror neurons for execution, observation, and imagery. *NeuroImage*, 37(4), 1315–1328. https://doi.org/10.1016/j.neuroimage.2007.06.008

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, 308(5722), 662–667. https://doi.org/10.1126/science.1106138

Franceschini, M., Ceravolo, M. G., Agosti, M., Cavallini, P., Bonassi, S., Dall'Armi, V., Massucci, M., Schifini, F., & Sale, P. (2012). Clinical relevance of action observation in upper-limb stroke rehabilitation: A possible role in recovery of functional dexterity. A randomized clinical trial. *Neurorehabilitation and Neural Repair*, 26(5), 456–462. https://doi.org/10.1177/1545968311427406

Friston, K. J., Glaser, D. E., Henson, R. N., Kiebel, S., Phillips, C., & Ashburner, J. (2002). Classical and Bayesian inference in neuroimaging: Applications. *NeuroImage*, 16(2), 484–512. https://doi.org/10.1006/nimg.2002.1091

Friston, K. J., Holmes, A. P., Price, C. J., Buchel, C., & Worsley, K. J. (1999). Multisubject fMRI studies and conjunction analyses. *NeuroImage*, 10(4), 385–396. https://doi.org/10.1006/nimg.1999.0484

Gallivan, J. P., Cavina-Pratesi, C., & Culham, J. C. (2009). Is that within reach? fMRI reveals that the human superior parieto-occipital cortex encodes objects reachable by the hand. *The Journal of Neuroscience*, 29(14), 4381–4391. https://doi.org/10.1523/JNEUROSCI.0377-09.2009

Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: Single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, 19(6), 1239–1255. https://doi.org/10.1093/cercor/bhn181

Grafton, S. T., & Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4), 590–616. https://doi.org/10.1016/j.humov.2007.05.009

Grèzes, J., Armony, J. L., Rowe, J., & Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurones in the human brain: An fMRI study. *NeuroImage*, 18(4), 928–937. https://doi.org/10.1016/s1053-8119(03)00042-9

Grosbras, M. H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human Brain Mapping*, 33(2), 431–454. https://doi.org/10.1002/hbm.21222

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience*, 12(5), 711–720. https://doi.org/10.1162/089892900562417

Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, 35(6), 1167–1175. https://doi.org/10.1016/s0896-6273(02)00897-8

Hardwick, R. M., Caspers, S., Eickhoff, S. B., & Swinnen, S. P. (2018). Neural correlates of action: Comparing meta-analyses of imagery, observation, and execution. *Neuroscience and Biobehavioral Reviews*, 94, 31–44. https://doi.org/10.1016/j.neubiorev.2018.08.003

Hiris, E. (2007). Detection of biological and nonbiological motion. *Journal of Vision*, 7(12), 1–16. https://doi.org/10.1167/7.12.4

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. C., & Rizzolatti, G. (2005). Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, 3(3), e79. https://doi.org/10.1371/journal.pbio.0030079

Insafutdinov, E., Andriluka, M., Pishchulin, L., Tang, S., Levinkov, E., Andres, B., & Schiele, B. (2017). Arttrack: Articulated multi-person tracking in the wild. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6457-6465). https://doi.org/10.48550/arXiv.1612.01465

Jastorff, J., Popivanov, I. D., Vogels, R., Vanduffel, W., & Orban, G. A. (2012). Integration of shape and motion cues in biological motion processing in the monkey STS. *NeuroImage*, 60(2), 911–921. https://doi.org/10.1016/j.neuroimage.2011.12.087

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201–211. https://doi.org/10.3758/BF03212378

Kemmerer, D. (2021). What modulates the Mirror neuron system during action observation? Multiple factors involving the action, the actor, the observer, the relationship between actor and observer, and the context. *Progress in Neurobiology*, 205, 102128. https://doi.org/10.1016/j.pneurobio.2021.102128

Koul, A., Cavallo, A., Cauda, F., Costa, T., Diano, M., Pontil, M., & Becchio, C. (2018). Action observation areas represent intentions from subtle kinematic features. *Cerebral Cortex*, 28(7), 2647–2654. https://doi.org/10.1093/cercor/bhy098

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21(6), 575–580. https://doi.org/10.3758/BF03198740

Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica*, 54(1–3), 115–130. https://doi.org/10.1016/0001-6918(83)90027-6

Lanzilotto, M., Ferroni, C. G., Livi, A., Gerbella, M., Maranesi, M., Borra, E., Passarelli, L., Gamberini, M., Fogassi, L., Bonini, L., & Orban, G. A. (2019). Anterior intraparietal area: A hub in the observed manipulative action network. *Cerebral Cortex*, 29(4), 1816–1833. https://doi.org/10.1093/cercor/bhz011

Lanzilotto, M., Livi, A., Maranesi, M., Gerbella, M., Barz, F., Ruther, P., Fogassi, L., Rizzolatti, G., & Bonini, L. (2016). Extending the cortical grasping network: Pre-supplementary motor neuron activity during vision and grasping of objects. *Cerebral Cortex*, 26(12), 4435–4449. https://doi.org/10.1093/cercor/bhw315

Lapenta, O. M., Xavier, A. P., Correa, S. C., & Boggio, P. S. (2017). Human biological and nonbiological point-light movements: Creation and validation of the dataset. *Behavior Research Methods*, 49(6), 2083–2092. https://doi.org/10.3758/s13428-016-0843-9

Lestou, V., Pollick, F. E., & Kourtzi, Z. (2008). Neural substrates for action understanding at different description levels in the human brain. *Journal of Cognitive Neuroscience*, 20(2), 324–341. https://doi.org/10.1162/jocn.2008.20021

Maeda, K., Ishida, H., Nakajima, K., Inase, M., & Murata, A. (2015). Functional properties of parietal hand manipulation-related neurons and mirror neurons responding to vision of own hand action. *Journal of Cognitive Neuroscience*, 27(3), 560–572. https://doi.org/10.1162/jocn_a_00742

Majdandzic, J., Bekkering, H., van Schie, H. T., & Toni, I. (2009). Movement-specific repetition suppression in ventral and dorsal premotor cortex during action observation. *Cerebral Cortex*, 19(11), 2736–2745. https://doi.org/10.1093/cercor/bhp049

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19(3), 1233–1239. https://doi.org/10.1016/s1053-8119(03)00169-1

Mayka, M. A., Corcos, D. M., Leurgans, S. E., & Vaillancourt, D. E. (2006). Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: A meta-analysis. *NeuroImage*, 31(4), 1453–1474. https://doi.org/10.1016/j.neuroimage.2006.02.004

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies.

Neuroscience and Biobehavioral Reviews, 36(1), 341–349. https://doi. org/10.1016/j.neubiorev.2011.07.004

Nath, T., Mathis, A., Chen, A. C., Patel, A., Bethge, M., & Mathis, M. W. (2019). Using DeepLabCut for 3D markerless pose estimation across species and behaviors. *Nature Protocols*, 14(7), 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Nelissen, K., Fiave, P. A., & Vanduffel, W. (2017). Decoding grasping movements from the parieto-frontal reaching circuit in the nonhuman primate. *Cerebral Cortex*, 28(4), 1245–1259. https://doi.org/10.1093/cercor/bhx037

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Papadourakis, V., & Raos, V. (2019). Neurons in the macaque dorsal premotor cortex respond to execution and observation of actions. *Cerebral Cortex*, 29(10), 4223–4237. https://doi.org/10.1093/cercor/bhy304

Pavlova, M., Krageloh-Mann, I., Sokolov, A., & Birbaumer, N. (2001). Recognition of point-light biological motion displays by young children. *Perception*, 30(8), 925–933. https://doi.org/10.1068/p3157

Pavlova, M. A. (2012). Biological motion processing as a hallmark of social cognition. *Cerebral Cortex*, 22(5), 981–995. https://doi.org/10.1093/cercor/bhr156

Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron*, 49(6), 815–822. https://doi.org/10.1016/j.neuron.2006.02.004

Pelosin, E., Avanzino, L., Bove, M., Stramesi, P., Nieuwboer, A., & Abbruzzese, G. (2010). Action observation improves freezing of gait in patients with Parkinson's disease. *Neurorehabilitation and Neural Repair*, 24(8), 746–752. https://doi.org/10.1177/1545968310368685

Pelphrey, K. A., Morris, J. P., Michelich, C. R., Allison, T., & McCarthy, G. (2005). Functional anatomy of biological motion perception in posterior temporal cortex: An FMRI study of eye, mouth and hand movements. *Cerebral Cortex*, 15(12), 1866–1876. https://doi.org/10.1093/cercor/bhi064

Peng, Y., Lee, H., Shu, T., & Lu, H. (2021). Exploring biological motion perception in two-stream convolutional neural networks. *Vision Research*, 178, 28–40. https://doi.org/10.1016/j.visres.2020.09.005

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., & Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *The Journal of Experimental Biology*, 146, 87–113.

Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *The European Journal of Neuroscience*, 21(10), 2864–2875. https://doi.org/10.1111/j.1460-9568.2005.04106.x

Pitzalis, S., Fattori, P., & Galletti, C. (2015). The human cortical areas V6 and V6A. *Visual Neuroscience*, 32, E007. https://doi.org/10.1017/S0952523815000048

Quadrelli, E., Roberti, E., Turati, C., & Craighero, L. (2019). Observation of the point-light animation of a grasping hand activates sensorimotor cortex in nine-month-old infants. *Cortex*, 119, 373–385. https://doi.org/10.1016/j.cortex.2019.07.006

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. *Physiological Reviews*, 94(2), 655–706. https://doi.org/10.1152/physrev.00009.2013

Saygin, A. P., Wilson, S. M., Hagler, D. J., Jr., Bates, E., & Sereno, M. I. (2004). Point-light biological motion perception activates human premotor cortex. *The Journal of Neuroscience*, 24(27), 6181–6188. https://doi.org/10.1523/JNEUROSCI.0504-04.2004

Schrouff, J., Rosa, M. J., Rondina, J. M., Marquand, A. F., Chu, C., Ashburner, J., Phillips, C., Richiardi, J., & Mourao-Miranda, J. (2013). PRoNTo: Pattern recognition for neuroimaging toolbox. *Neuroinformatics*, 11(3), 319–337. https://doi.org/10.1007/s12021-013-9178-1

Servos, P., Osu, R., Santi, A., & Kawato, M. (2002). The neural substrates of biological motion perception: An fMRI study. *Cerebral Cortex*, 12(7), 772–782. https://doi.org/10.1093/cercor/12.7.772

Sgandurra, G., Ferrari, A., Cossu, G., Guzzetta, A., Fogassi, L., & Cioni, G. (2013). Randomized trial of observation and execution of upper extremity actions versus action alone in children with unilateral cerebral palsy. *Neurorehabilitation and Neural Repair*, 27(9), 808–815. https://doi.org/10.1177/1545968313497101

Shim, J., Carlton, L. G., & Kim, J. (2004). Estimation of lifted weight and produced effort through perception of point-light display. *Perception*, 33(3), 277–291. https://doi.org/10.1068/p3434

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences of the United States of America*, 105(2), 809–813. https://doi.org/10.1073/pnas.0707021105

Tanisaro, P., Lehman, C., Sütfeld, L., Pipa, G., & Heidemann, G. (2017). Classifying bio-inspired model of point-light human motion using echo state networks. In: A. Lintas, S. Rovetta, P. Verschure, A. Villa (Eds.), *Artificial Neural Networks and Machine Learning – ICANN 2017. ICANN 2017. Lecture Notes in Computer Science*, (Vol. 10613). Cham, Springer. https://doi.org/10.1007/978-3-319-68600-4_11

Thornton, I. M. (2006). Biological motion: Point-light walkers and beyond. In *Human body perception from the inside out: Advances in visual cognition* (pp. 271–303). Oxford University Press.

Tkach, D., Reimer, J., & Hatsopoulos, N. G. (2007). Congruent activity during action and action observation in motor cortex. *The Journal of Neuroscience*, 27(48), 13241–13250. https://doi.org/10.1523/JNEUROSCI.2895-07.2007

Toshev, A., & Szegedy, C. (2014). Deeppose: Human pose estimation via deep neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1653–1660. https://doi.org/10.1109/CVPR.2014.214

Ulloa, E. R., & Pineda, J. A. (2007). Recognition of point-light biological motion: Mu rhythms and mirror neuron activity. *Behavioural Brain Research*, 183(2), 188–194. https://doi.org/10.1016/j.bbr.2007.06.007

Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 98(20), 11656–11661. https://doi.org/10.1073/pnas.191374198

van Kemenade, B. M., Muggleton, N., Walsh, V., & Saygin, A. P. (2012). Effects of TMS over premotor and superior temporal cortices on biological motion perception. *Journal of Cognitive Neuroscience*, 24(4), 896–904. https://doi.org/10.1162/jocn_a_00194

Weiner, K. S., & Zilles, K. (2016). The anatomical and functional specialization of the fusiform gyrus. *Neuropsychologia*, 83, 48–62. https://doi.org/10.1016/j.neuropsychologia.2015.06.033

Wilke, M., & Lidzba, K. (2007). LI-tool: A new toolbox to assess lateralization in functional MR-data. *Journal of Neuroscience Methods*, 163(1), 128–136. https://doi.org/10.1016/j.jneumeth.2007.01.026

Yoshida, K., Saito, N., Iriki, A., & Isoda, M. (2011). Representation of others' action by neurons in monkey medial frontal cortex. *Current Biology*, 21(3), 249–253. https://doi.org/10.1016/j.cub.2011.01.004

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.