# An action-observation network in the common marmoset identified by ultra-high field fMRI

Alessandro Zanini[1,*], Audrey Dureux[1], Janahan Selvanayagam[2], Stefan Everling[1,2]

Author affiliations:

[1]Centre for Functional and Metabolic Mapping, Robarts Research Institute, University of Western Ontario, London, Ontario, Canada

[2]Department of Physiology and Pharmacology, University of Western Ontario, London, Ontario, Canada

*Corresponding author:

Alessandro Zanini, Robarts Research Institute, University of Western Ontario, London, Ontario, Canada. **Email**: azanini4@uwo.ca

**Abstract**

The observation of others' actions activates a network of temporal, parietal and premotor/prefrontal areas in macaque monkeys and humans. This action-observation network (AON) has been shown to play important roles in understanding the actions of others, learning by imitation, and social cognition in both species. It is unclear whether a similar network exists in New World primates, which separated from Old Word Primates ~ 35 million years ago. Here we used ultra-high field fMRI at 9.4T in awake common marmosets (*Callithrix jacchus*) while they watched videos depicting the upper-limb of conspecifics performing goal-directed (grasping food) or non-goal-directed actions. We found that the observation of goal-directed actions, compared to non-goal directed ones, activated a temporo-parieto-frontal network, including areas 6 and 45 in premotor and prefrontal cortices, areas PGa-IPa, FST and the TE complex in occipito-temporal region and areas V6A, MIP, LIP and PG in the occipito-parietal cortex. These results show remarkable overlap with the AON observed in humans and macaques. These results demonstrate the existence of an evolutionarily conserved AON that likely predates the separation of Old and New World primates.

*Keywords*

Marmoset, action observation network, fMRI, grasping, mirror neurons

## 1. Introduction

Understanding the intentions and behavior of other individuals is crucial for our daily social interactions. When we observe other individuals carrying out a certain action, we are often able to draw conclusions about their intentions and/or the purpose of their actions. Functional imaging studies have shown that action observation activates a large-scale network of brain areas known as the action-observation network (AON, 1–3). In Old-World primates, the core of the AON includes three main nodes: the occipito-temporal region corresponding to the superior temporal sulcus (STS), the inferior parietal lobule (IPL or PFG/AIP) and a prefrontal/premotor cluster (human's ventral premotor cortex and inferior frontal gyrus, macaque's F5 and area 45) (1, 4–13, and see 14, 15, and 16 for reviews).

The AON has been described as a hierarchical system (but with both feedforward and feedback connections, 17) which contains areas activated by progressively greater complexity (18). Visual neurons in the STS are sensitive to action observation processes and integrate multiple low-features, such as shape and motion (5, 7, 19–22). This represents the visual input of the AON which is subsequently transmitted to the rest of the network through projections to visual and visuomotor neurons in the IPL (7, 9, 23–25). The parietal and premotor nodes of the AON represent observed actions according to their goal (12, 9, 26, 27, 6, 7, 28, 14, 25, 5, 29), enabling understanding of others' intentions and facilitating the planning of an adequate behavioural response.

The three nodes described above represent the "classic" AON core, but more recent studies suggest that this network may be more extensive (12), including clusters in dorsal premotor and motor cortices, somatosensory areas (SI and SII), ventrolateral prefrontal cortex and in the superior parietal lobule (IPS), in both humans (1, 15, 30–32) and macaques (33–38).

Support for a putative AON in New World monkeys has come from an electrophysiological study by Suzuki and colleagues, who reported a small number of neurons in the ventrolateral frontal region and in the STS of the common marmoset (*Callithrix jacchus*) that were activated by action observation (39). Marmosets live in family groups and shares some interesting similarities with humans, including prosocial behavior, imitation, and cooperative breeding. Their small size combined with recent wireless and datalogger recording techniques make it possible to record high density neural activity during natural social behaviour in this species. Their lissencephalic (smooth) cortex offers the opportunity for laminar

electrophysiological recordings and optical imaging in many cortical areas. These features could make marmosets an excellent primate model for the study of the neural basis of action observation.

To identify a putative AON in marmosets, we employed whole-brain ultra-high-field fMRI at 9.4T in awake marmosets. Similar to previous macaque studies (6, 7), we investigated the activations induced by the presentation of videos representing the upper limb of a marmoset performing a goal-directed (reaching and grasping for food) or non-goal-directed arm and hand movement (reaching and grasping to an empty space).

## 2. Materials and Methods

All the experimental procedures here described were performed in accordance with the guidelines of the Canadian Council on Animal Care policy on the care and use of experimental animals and an animal use protocol #2021-111 approved by the Animal Care Committee of the University of Western Ontario.

### *2.1. Subjects and experimental setup*

Seven common marmosets (*Callithrix jacchus*; three females, average age: 35.8±8.5 months, ranging from 30 to 54 months, average weight: 398 ± 46.7 g, ranging from 328 to 462 g) took part in an awake fMRI study after the surgical implantation of an MRI-compatible head restraint chamber (n=4, see 40 for details) or of an MRI-compatible head post (n=3). In order to reduce stress and anxiety due to scanning and to head-fixation, all marmosets underwent a 3-week training following the acclimatization procedure prior to the experiments described in Gilbert et al. (2021). All animals scored 1 or 2 on the behavioural assessment scale described by Silva et al. (2011) after the training, reporting little or no signs of agitation and a calm behaviour.

Before MRI acquisition, the marmoset was placed in the sphinx position using the same MRI-compatible restraint system described in Schaeffer et al. (2019). After entering the animal holder, the monkey was firstly restrained using a neck and a tail plates. Taking advantage of either the MRI-compatible chamber or the head post, the head of the marmoset was then fixed to a five-channel (chamber) or eight-channel (head post) receive coil. A lubricating gel (MUKO SM1321N, Canadian Custom Packaging Company, Toronto, Ontario, Canada) was applied on the top of the chamber and on its edges in order to reduce magnetic susceptibility and artifacts.

An MRI-compatible camera (model 12M-i, 60 Hz sampling rate, MRC Systems GmbH, Heidelberg, Germany) mounted on the front of the animal holder allowed the monitoring of the animal's condition and awake state by a veterinarian technician. Once in the scanner, marmosets faced a translucent plexiglass screen attached to the front of the scanner bore (~119 cm from animal's eyes) on which the visual stimuli projected by a SONY VPL-FE40 projector were reflected via a wall-mounted mirror. Maximum visual angle from the centre to the side of the screen was 6.5°. Visual stimuli were presented via PowerPoint in synchronization with MRI TTL pulses triggered by a Raspberry Pi (model 3B+, Raspberry Pi Foundation, Cambridge, UK) running a custom python script.

*2.2. Task and Stimuli*

Similarly to Nelissen and colleagues' studies (6, 7), the visual stimuli used in this experiment consisted of movies showing reaching and grasping actions or movements of the upper limb. All videos were recorded inside the marmoset facility of the University of Western Ontario and included two different motor acts performed by the upper limb of a marmoset (none of the "actors" were included in the experimental sample nor were they housed in the same room as the fMRI experimental animals). Video editing was performed in Adobe Premiere Pro (Adobe, San Jose, California, USA). In the first experimental condition, hereafter called "Grasping Hand" (S1 movie), the movie represented the hand and forearm of a marmoset reaching towards and grasping a small piece of food (marshmallow, ~2 cm of diameter). In this condition, a complete goal-directed action is therefore presented in which the food is reached, grabbed and finally transported outside the scene. In the second experimental condition, hereafter called "Empty Hand" (S2 movie), the movement of the upper limb is very similar, but in this case the action is purposeless, as no object is present in the scene. To ensure the greatest similarity between the movements of the two conditions, Empty Hand videos were obtained using grasping movements towards a marshmallow, placed in the same position used in the Grasping Hand condition, removed just before the grasping phase. The marshmallow was then removed from the video in post-production. All videos lasted for 3 seconds; due to the speed of the recorded grasping movement, these clips were slowed down to fill this time interval (S1 and S2 movies are examples of slowed-down actions, slowing range 20-45%). Twelve different videos were selected for each condition. To check the effect of the direction of action, all videos

were mirrored, presenting rightward and leftward sequences within each run for each monkey. In addition, to check possible effects due simply to the perception of moving stimuli, all videos were scrambled (see S3 and S4 movies for examples of grasping and empty hand scrambled conditions respectively) by random rotation of the phase information performed with a custom MATLAB script (MATLAB, The MathWorks, Natick, MA). The use of a constant seed made it possible to preserve the aspects of motion and luminance in scrambled videos, applying the same rotation matrix to each frame of each movie. Each run included 8 stimulation blocks (two blocks for each of the four experimental conditions: grasping hand, scrambled grasping hand, empty hand and scrambled empty hand), whose order was fully randomized across runs, sessions and animals. Each block was composed of a sequence of four 3 seconds videos belonging to the same experimental condition, for a duration of 12 seconds, and was interleaved 18 seconds of baseline, in order to allow the washing-out of the BOLD signal between experimental blocks (total duration of a run: 258 seconds). During the baseline, a black filled circular shape (1.5 cm of diameter, 0.72°) was presented at the center of the screen, but the animals were not required (nor trained) to fix: the only purpose of the black dot was to reduce the nystagmus sometimes present because of the ultra-high magnetic field. A diagram of the structure of a run can be seen in Figure 1. The compliance of the animal during each run was checked and noted online by the investigator; runs in which the animal closed its eyes for two or more stimulation blocks (regardless of the experimental condition) were discarded from further analyses. Animals did not receive any reward during the sessions.

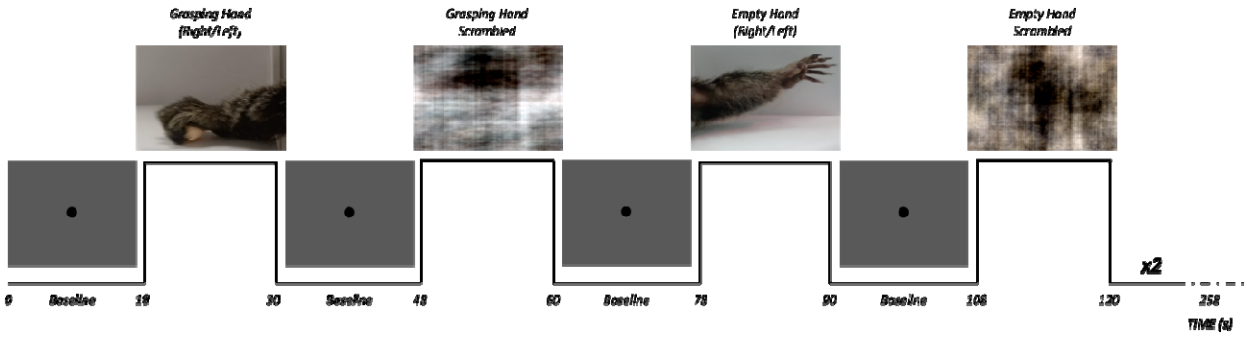


**Figure 1** Experimental design for task-based fMRI. In each run, 9 baseline blocks (18 seconds long) were alternated with experimental blocks (12 seconds long) of four different conditions: Grasping Hand, Empty Hand, Grasping Hand Scrambled and Empty Hand Scrambled. Each run lasted 258 seconds globally.

*2.3. MRI Acquisition*

All scanning sessions were carried out in a 9.4T, 31 cm horizontal bore magnet (Varian/Agilent) and a Bruker BioSpec Avance III console with the software package Paravision-7 (Bruker BioSpin Corp), a custom-built high-performance 15-cm diameter gradient coil with 400-mT/m maximum gradient strength (44), and a five- (43) or eight-channel receive coil (Gilbert et al., submitted) at the Centre for Functional and Metabolic Mapping at the University of Western Ontario. An in-house built quadrature birdcage coil (12-cm diameter) functioned as transmit coil.

For functional imaging, gradient-echo-based, single-shot echo-planar images covering the whole brain were acquired over multiple daily sessions (TR=1500 ms; TE=15 ms; flip angle=40°; FOV=64×48 mm; matrix size=96×128; voxel size=0.5 mm isotropic; number of slices=42 [axial]; bandwidth=400 kHz; GRAPPA acceleration factor (left-right=2). To correct for spatial distortion, a second set of echo-planar images with the opposite phase-encoding direction (right-left) was collected. To perform anatomical registration, a T2-weighted structural image was acquired for each animal with the following parameters: TR=7000 ms; TE=52 ms; FOV=51.20 × 51.20 mm; voxel size=0.133×0.133×0.5 mm; number of slices=45 (axial); bandwidth=50 kHz, GRAPPA acceleration factor=2.

Each session lasted approximately 50-60 minutes, including preparation of the animal and of the sequences. We acquired 10 functional runs that fulfilled our inclusion criteria for each marmoset, which required between 4-7 sessions, depending on monkey's compliance.

*2.4. Image Preprocessing*

The data were preprocessed using a combination of AFNI (45) and FSL's functions (46). The raw functional images were first converted to the NIfTI format using dcm2niix (47) and then reoriented (FSL's fslswapdim and fslorient) to correct the sphinx posture. The reoriented functional images were preprocessed through procedures for eliminating any outliers (detected via AFNI's 3dToutcount), despiking (AFNI's 3dDespike) and time shifting (AFNI's 3dTshift). The images thus obtained were registered to the base volume (extracted at the half of each run and therefore corresponding to the 86th volume) using AFNI's 3dvolreg function. Then, all volumes were smoothed (AFNI's 3dmerge, FWHM Gaussian kernel of 2mm) and bandpass-filtered (AFNI's 3dBandpass, lowest frequency of 0.1 and highest frequency of 0.01). An

average functional image was calculated for each run of each animal and then registered (FSL's FLIRT function) to the respective T2-weighted image. The transformation matrix thus obtained was subsequently used to carry out the 4D time series data. T2-weighted images were manually skull-stripped (removing the olfactory bulb) and the mask obtained was applied to the functional images.

Finally, anatomical and functional images were registered to the NIH marmoset brain atlas (48) via the nonlinear registration operated by ANTs' (Advanced Normalization Tools, 49) ApplyTransforms function.

2.5. *Contrasts analysis*

The scan timing was convolved to the BOLD response (AFNI's 3dDeconvolve) specifying the 'BLOCK' convolution and extracting a regressor for each condition (Grasping Hand, Empty Hand and the two Scrambled conditions) for each run to be used in the subsequent regression analysis. All the conditions were entered in the model, along with polynomial detrending regressors (n=5). This regression generated four T-value maps, corresponding to the four experimental conditions, per animal per run, that were registered to the NIH marmoset brain atlas (48). These maps were thus compared through paired t tests (AFNI's 3dttest++).

At the group level, the T-maps of each animal for each condition and for each run (n=10) were converted into Z-maps. To investigate the neural responses induced by the perception of an "intact" biological movement versus a scrambled movement pattern, the Z score maps (n=140) of the Grasping and Empty Hand conditions of each animal were compared via t-test to those of the Scrambled Grasping Hand and Scrambled Empty Hand conditions. For protection against false positives, the results of this t-tests have been corrected with a minimum cluster-size resulting from 10000 Monte Carlo simulations (AFNI's 3dttest++ with Clustsim option, $p<0.001$ and $\alpha=0.05$, nearest-neighbor clustering method 2, two-sided).

Finally, the presence of areas selective for goal-directed actions was investigated by comparing the Z value maps (n=70) of the Grasping Hand and Empty Hand conditions. As described above, we compared these maps using a paired t-test, which result was protected from false positives thanks to the same cluster-size correction (AFNI's 3dttest++ with Clustsim option, $p<0.001$ and $\alpha=0.05$, nearest-neighbor clustering method 2, two-sided).

*2.6. Eye Tracking*

Marmosets have very large pupils which make video eye tracking challenging in the MR environment. Thus, the signal from the MRI-compatible camera used to assess the alert state of the animals in the scanner was not of sufficient quality to guarantee a reliable gaze analysis. To investigate potential differences in eye movements between our two conditions of interest (i.e., Grasping Hand and Empty Hand) we conducted an eye-tracking experiment outside of the scanner, in a sound attenuating chamber (Crist Instruments Co., Hagerstown, MD, USA) free of MRI-induced noise. 10 common marmosets (5 females, average age 35.5±6.78 months, average weight 417.4±57.7 grams) took part in this experiment. Among them, 7 performed the eye-tracking task before (n=4) or after (n=3) the fMRI sessions. Depending on the type of implant (chamber or head-post), the animals were head restrained in the sphinx position in the MRI chair, following the same procedure described above, or upright, in a custom chair (40), inside the sound attenuating chamber. Grasping (24 videoclips for right- and leftward grasping) and Empty Hand videos (24 videoclips for right- and leftward grasping) were presented with NIMH Monkeylogic (50) at 57 cm from the eye of the animal with the same size (in visual angle) described above (see *2.2. Task and Stimuli*). At the beginning of each session, the eye position was calibrated by rewarding the monkey for 300 to 600ms fixations on a 1-degree dot presented at the display centre and at 6 degrees in each of the cardinal directions using NIMH Monkeylogic (50).

All stimuli were presented on a CRT monitor (ViewSonic Optiquest Q115, 76 Hz non-interlaced, 1600 x 1280 resolution). Eye position was digitally recorded at 1 kHz via video tracking of the left pupil (EyeLink 1000, SR Research, Ottawa, ON, Canada). Eye movement analysis was conducted using an in-house written Python code. Eye velocity (visual degrees/second) was calculated through smoothing and numerical differentiation, and saccades were defined as eye movements with a radial eye velocity greater than 30 deg/s and amplitude greater than 0.35 visual degrees. To replicate the experimental conditions of the fMRI sessions, no reward was delivered during the eye-tracking. Differences in saccade frequency (number of saccades/second) and median amplitude were investigated through Wilcoxon signed-rank tests for two-paired samples.

## 3. Results

In this fMRI study we used a block-design in which each run contained 8 video sequences belonging to four different experimental conditions, alternating with baseline blocks. The videos belonging to the Grasping Hand condition showed the upper limb of a marmoset performing a complete reaching-to-grasp movement of a small piece of food (i.e., marshmallow). In the Empty Hand condition, a similar movement was shown, but in this case the piece of food was absent, rendering the action non-goal-directed. The other two experimental conditions (Scrambled Grasping Hand and Scrambled Empty Hand) were the phase-scrambled version of these two classes of videos (see Methods section for details). 7 common marmosets were included in the study, using ten runs per animal in the analyses (70 runs in total). In the eye-tracking experiment controlling for differences in oculomotor patterns, the same Grasping and Empty Hand videos were presented in a sound-attenuating chamber to 10 common marmosets.

*3.1. An occipito-temporal cluster to process vision of body-parts in motion*

We first performed an analysis to identify brain activations induced by videos displaying intact movements (i.e, a movement of an arm and hand), without considering the goal of the action, (Grasping and Empty Hand conditions), versus phase-scrambled videos of these movements (Scrambled Grasping and Empty Hand conditions). The results of this group-level analysis are shown in Figure 2 (paired t-test, results reported at $p<0.001$ and corrected for cluster-size threshold via Monte Carlo simulations, with $\alpha=0.05$). We found that the observation of the intact movement videos elicited stronger activations both in the occipital (in high-level visual areas, such as V3, V4 and V4T, in particular in the left hemisphere) and temporal cortices. In the latter case, we observed two large bilateral activation clusters, including area FST, the PGa-IPa conjunction, and a wide portion of the TE complex (TE1, TE2, TE3 and TEO) in the left hemisphere, as opposed to a less extensive contralateral activation (including almost all of the TE3 area). Significant activations at the level of $p <0.001$ but not extended enough to resist cluster-wise correction can be observed at the level of the bilateral prefrontal cortex, including areas 45 and 47, and the left premotor cortex (6Va and 6Vb).

These results show regions activated by videos of grasping movements with or without goal-directed actions, consistent with previously described body patches in the marmoset temporal cortex (51, 52).

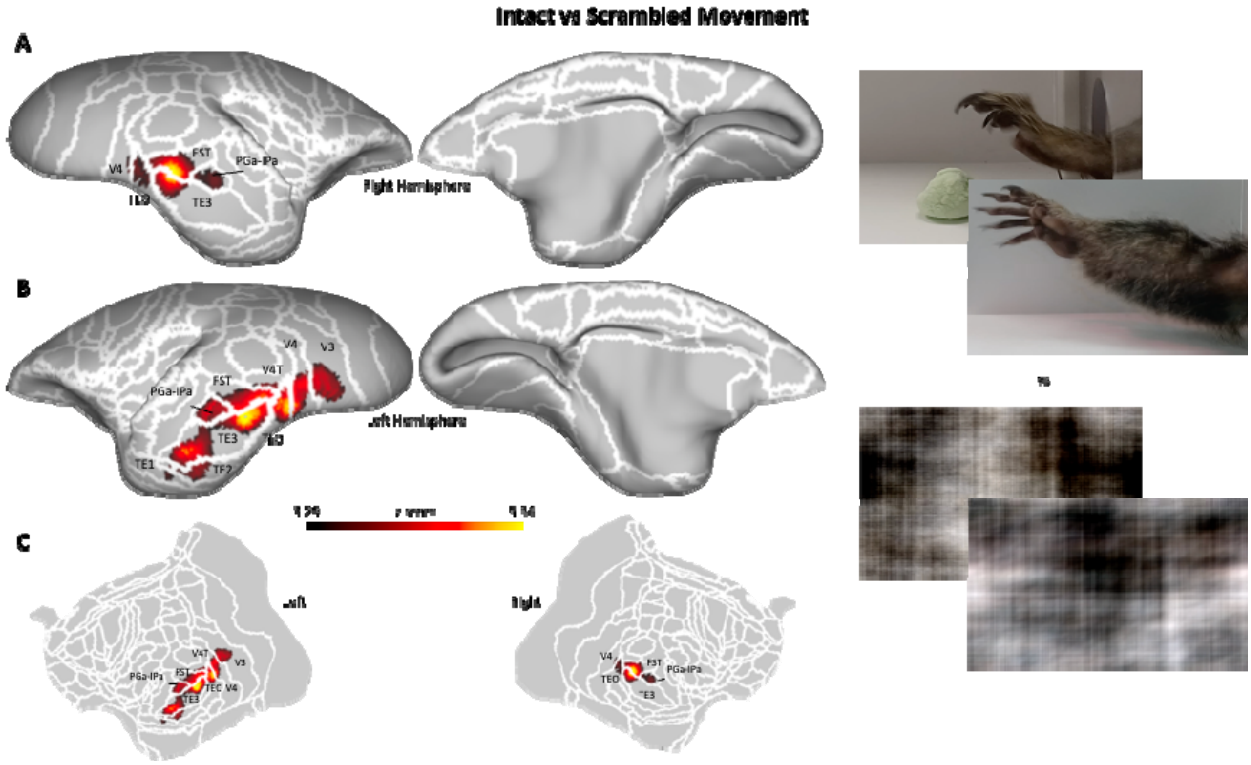


**Figure 2**. Functional maps of the contrasts between the "Intact" versus the "Scrambled" Motion conditions. Maps are displayed on the left (A) and right (B) fiducial brain surfaces (lateral view on the left, medial view on the right) and on flat maps (C, left hemisphere on the left, right hemisphere on the right). White lines delineate the cerebral areas included in the Paxinos parcellation of the NIH marmoset brain atlas (Liu et al., 2018). The responses here reported have intensity higher than z = 3.29 (corresponding to p<0.001, AFNI's 3dttest++) and survived the cluster-size correction (10000 Monte-Carlo simulations, $\alpha$=0.05).

*3.2. An action-observation network for goal-directed actions*

Whereas the occipito-temporal node of the AON is mainly involved in processing several low-level features of the action observed, the parietal and premotor/prefrontal nodes process actions in terms of their goal (5, 7). To test whether areas in the marmoset brain show a similar pattern, we compared the activations elicited by the Grasping condition with the Empty Hand condition (paired t-test, results reported at p<0.001 with cluster-size threshold determined by Monte Carlo simulations for $\alpha$=0.05). The results show a large network of fronto-temporo-parietal regions that is strongly activated by goal-directed actions, compared to a similar but non-goal directed motor act (Figure 3). The premotor/prefrontal cluster includes rostral and caudal dorsal premotor cortices (6DR and 6DC) and the ventral premotor cortex (6Va), bilaterally. These clusters extend to a small portion of dlPFC (8C in both hemispheres, plus right 8aV) and

to the vlPFC (bilateral area 45, plus wide portions of the right area 47). In the right hemisphere, we observed that this cluster also extends to area 13 in the OFC, while in the left hemisphere there is a region more active for goal-directed actions in the more rostral part of the motor cortex (4ab).

A similar bilateral pattern of activation is observed at the occipito-parietal junction, in which stronger responses to goal-directed actions are present in higher-level visual areas (V3A, V6 and V6A) and in parietal areas MIP and LIP, with a greater extension of these latter activations in the left hemisphere. Here, we also find an extension of the cluster to parietal area PG, at the border to LIP.

Finally, the third and largest bilateral activation cluster is found in the occipito-temporal cortex, where stronger responses to an observed goal-directed action are located in both visual (V2, V3, V4, V4T and V5) and in more temporal regions including the most dorsal part of TEO, the FST, the PGa-IPa and the TE3.

In addition to these bilateral activations, we also observed a unilateral activation in the right hemisphere, in the posterior cingulate region, corresponding to medial areas 29, 30 and PGM.

Subcortically, we found bilateral activations in the hippocampal formation and in the thalamus. In the right hemisphere, a first cluster included the anteromedial and anteroventral thalamic nuclei, extending into the laterodorsal and ventrolateral dorsal thalamic nucleus. In the left hemisphere, we found activations in the laterodorsal and in ventrolateral thalamic nuclei. Both this clusters extended posteriorly, including wide portions of the bilateral pulvinar

These results demonstrate the existence of a widespread fronto-temporo-parietal network activated by the observation of goal-directed reaching-to-grasp movements, similar to the AON already described in both humans and macaque monkeys (6, 7, 5, 12, 53, 23, and see 54, 1, and 14 for review).

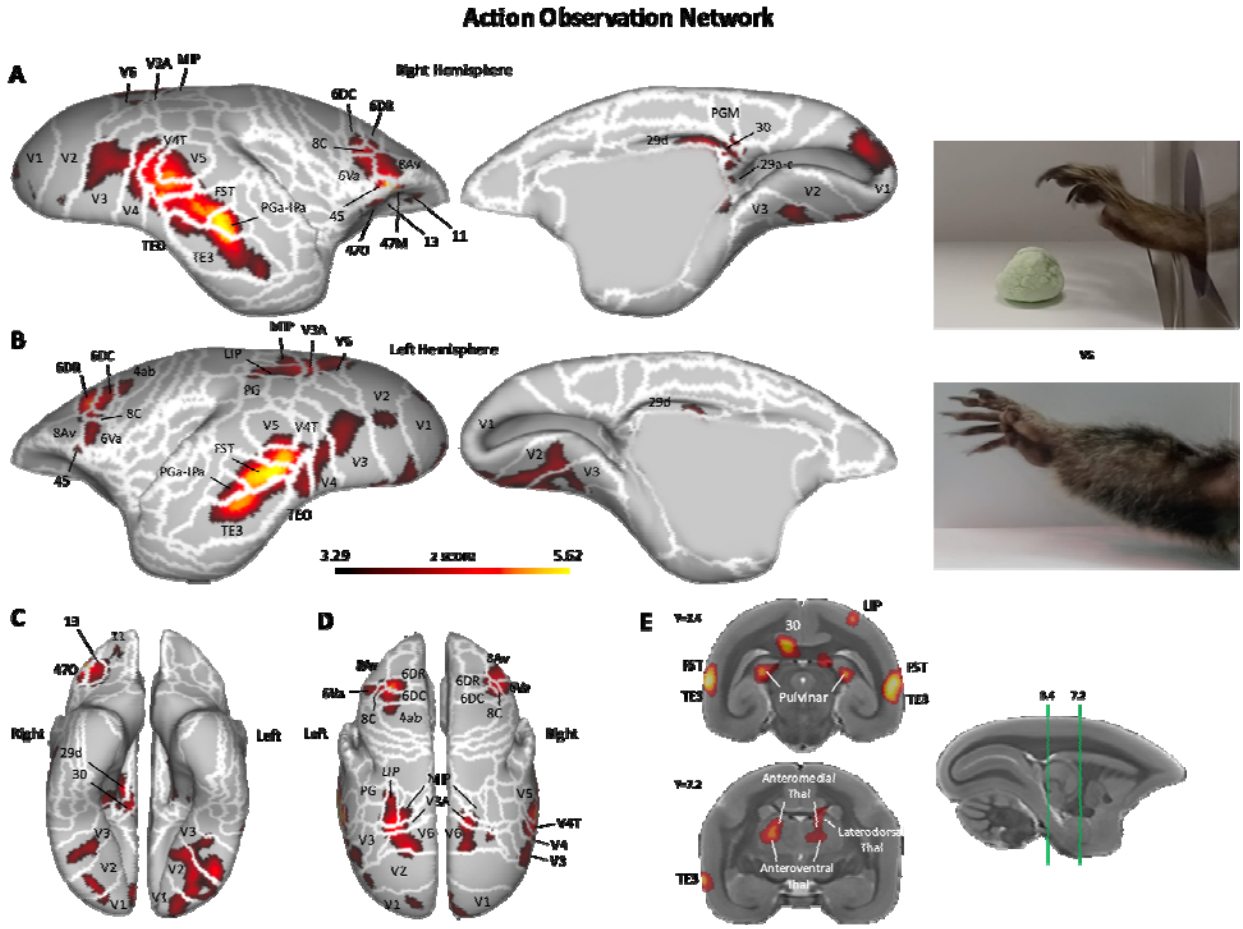


**Figure 3**. Functional maps of the contrasts between Grasping Hand and Empty Hand conditions. Maps are displayed on the left [A] and right [B] fiducial brain surfaces (lateral view on the left, medial view on the right, ventral view in [C], dorsal view in [D]). Particular of the subcortical thalamic activations in [E]. White lines delineate the cerebral areas included in the Paxinos parcellation of the NIH marmoset brain atlas (Liu et al., 2018). The responses here reported have intensity higher than z = 3.29 (corresponding to p<0.001, AFNI's 3dttest++) and survived the cluster-size correction (10000 Monte-Carlo simulations, $\alpha$=0.05).

*3.3. Similar eye movement patterns for body-parts in motion*

Figure 4 reports the results of the analyses conducted on the frequency and amplitude of saccades during the viewing of the Grasping and Empty hand videos outside of the scanner. Two separate Wilcoxon signed rank tests were conducted on these two variables to investigate potential differences in oculomotor patterns during the two main experimental conditions. As shown in Figure 4, the eye-tracking experiment did not reveal any significant difference in the frequency of the saccades (W=26, p=0.92) nor in saccade amplitude (W=23, p =0.70) between the Grasping and Empty hand videos. Overall, these results suggest that the cortical activations

observed in the fMRI experiment between videos depicting either goal-directed or non-goal-directed actions cannot be attributed to differences in saccadic eye movements.

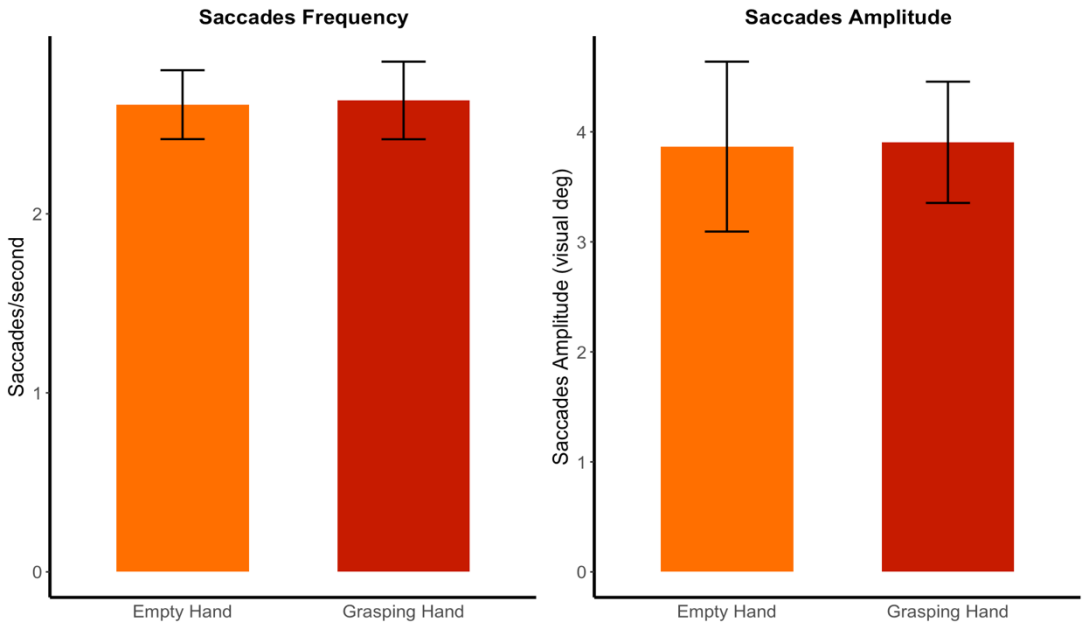


**Figure 4**. Frequency (left panel, number of eye movements per second with a radial eye velocity greater than 30 deg/s and amplitude greater than 0.35 visual degrees) and amplitude (right panel, in visual degrees) of saccades performed during the presentation of videos depicting goal-directed (Grasping Hand condition, in red) and non-goal-directed actions (Empty Hand condition, in orange). No differences between conditions emerged using Wilcoxon signed rank test (in both cases, $p>0.6$).

## 4. Discussion

Observing and understanding others' actions is crucial for social animals like primates. Numerous studies have shown that a fronto-temporo-parietal network mediates this function in Old World macaques and humans (1, 5, 7, 14, 15, 55). This so-called action observation network (AON) responds to goal-directed actions (6, 7), and its main function seems to be to attribute a purpose to the actions of others in order to plan an effective response, providing the foundation for social interactions (29, 56). A strong conceptual overlapping exists between the definition of AON and that of the mirror neuron system, a brain network activated both during the observation and the execution of goal-directed actions (57–59). While the AON is solely activated by action observation, the mirror neurons system is activated by both observation and execution of actions. Here, we investigated the existence of an AON in the common marmoset, a small New World

primate that is becoming popular as powerful additional nonhuman primate model (60). For this purpose, we presented awake marmosets with videos depicting a finalized action (grasping food) or a similar but non-goal-directed movement (grasping air) in an fMRI block-designed task, alternated by videos of their phase-scrambled versions.

The comparison between videos depicting an action ("intact") and their phase-scrambled versions elicited strong activations in the occipital and temporal lobes, consistent with what we observed in previous studies (61, 62) and overlapping with the location of body and face patches (51, 52). These activations, proceeding in a caudo-rostral direction, include portions of the visual areas V4/V4T, the fundus of the superior temporal sulcus (FST), the PGa-IPa and a wide part of the TE complex, in particular area TE3 and the most dorsal part of TEO. Therefore, these patches seem to respond specifically to the presentation of a part of the body (the hand or the entire arm) in motion. A similar pattern of occipito-temporal activations has been reported during the presentation of videos depicting faces in marmosets (51, 52, 62), and low-level features of moving body-parts in macaques and human (11, 22, 53, 63–69). Within the AON, this node therefore seems to be responsible for the detection and processing of a biological movement, a fundamental step in the recognition of an external agent (53). Only after this recognition is it possible to attribute an intention and a goal to the actions of the external agent.

The main finding of this study comes from the comparison between goal-directed and non-goal-directed actions. As already observed by the first electrophysiological (57–59) and functional imaging studies in macaques (6, 7, 65, 70) and humans (8, 71, 72), neurons in the AON show greater activity during the observation of goal- versus non-goal-directed movements. It has been hypothesized that this feature could provide the basis of the ability to recognize others' actions and understand their behavior and intentions (53, 73).

While the existence of a AON has been established in both human and non-human Old World primates (1, 14, 53, 54, 74), only a small number of neurons responding to goal-directed actions have been identified by a single neuron recording study in the ventrolateral prefrontal region of New World marmosets (39). Thus, our study represents the first support for an extended AON in the common marmoset. This network shows strong correspondences with the AON in macaques and humans. The first (mirror) neurons in macaques responding to the observation of others' actions have been observed in areas F5 and 45 (6, 7, 57–59), and regions with similar properties have also been found with fMRI in homologous brain areas in humans (5,

17, 30, 71, 72, 75–79). Our results show that large parts of the bilateral marmoset area 6V (likely corresponding to macaque's area F5 and human area PMv) and of the ventrolateral prefrontal cortex (area 45 bilaterally, plus right areas 47 and 13) are part of this AON. The homologies between the AON of the three species continue with activations in the dorsal area 6 (6DR and 6DC, bilaterally), resembling what has been observed by electrophysiological (33, 36, 38) and functional imaging studies (30–32) in macaques and humans.

Outside the premotor and prefrontal cortex, one of the main AON regions is the STS in both humans (5, 21, 32, 80) and macaques (7, 19, 67). Here, we found a posterior region in the fundus of the temporal sulcus (FST) in the New World marmoset that was more active for the observation of goal-directed versus non-goal-directed movements, consistent with the findings of a macaque fMRI study that used a similar experimental paradigm (7). In marmosets, this cluster extends also towards more lateral and inferior temporal regions, involving the more dorsal part of area TE3 and the PGa-IPa area, confirming previous observations in macaques (65).

At the parietal level, the observation of goal-directed actions induces a bilateral cluster of activation including high-level visual areas (V3A, V6, V6A), the intraparietal areas MIP and LIP and the PG area. Although not part of the classic AON core, responses in these visual areas have previously been identified in macaques (24, 70) and could be involved in the processing of visuospatial information required for the coding of reaching actions. Moreover, a circuit recruited in the planning and execution of arm reaching movements includes area V6A, area MIP and dorsal premotor area F2 (corresponding to area 6DC in marmosets) (81) in macaques. These regions are also activated bilaterally in our study and therefore could be a part of an extended AON. Further studies using different types of goal-directed actions and/or different effectors will be required to test this hypothesis.

Surprisingly, our study did not find activations at the level of the inferior parietal lobule and in area AIP, one of the hubs of the AON often reported in human (12, 14, 82–85) and macaque studies (9, 14, 25–27, 86). The rostral part of the IPL would indeed be the junction between the visual input coming from the STS and directed to the prefrontal and premotor cortices, structures not directly connected to each other in the macaque monkey (23, 87–89). Conversely, recent studies demonstrated a direct bidirectional connection between the ventrolateral prefrontal cortex, including area 6V, and area FST in common marmosets (39, 90). This could explain the reduced parietal activation for goal-directed actions in our study. A

second possible explanation concern the content of the videos that we used: while humans and macaques are more accustomed to the use of the upper limbs for reaching, grasping and manipulating objects, marmosets often use their mouth if left free to choose. It is therefore possible that the use of videos depicting grasping-with-mouth actions may prove even more effective in mapping brain regions that respond to action observation in this primate species.

Although the frontal and parietal regions reported in our study are homologous to those found in the AON in Old World primates, some of them are also known to play a role in eye movements in marmosets (91–94). However, the results of our eye-tracking study show that the frequency and amplitudes of saccades did not differ between the two main experimental conditions. Hence, it is highly unlikely that these activations are due to differences in eye movements between conditions. Moreover, a previous microstimulation study from our lab also showed facial and forelimb responses in this frontal region (94), and single neuron recordings have identified neurons involved in the generation and processing of vocalizations in these same frontal areas (45, 8Av, 6V, 6D) in marmosets (95). Further electrophysiological recordings will be necessary to test whether separate neural populations perform these different functions or whether the same neurons have multiple response patterns (e.g. saccade, vocal-motor, or grasping responses).

In summary, our study provides the first functional mapping of an extensive AON in the common marmoset. This network exhibits striking similarities with the well-characterized AON in humans and macaques, suggesting that the common ancestor of New and Old World primates possessed a system for understanding the actions of others. The identification of an AON in marmosets provides the foundation for targeted recording and stimulation studies that will advance our mechanistic understanding of primate social cognition.

**Acknowledgements**

Support was provided by the Canadian Institutes of Health Research (FRN 148365), the Canada First Research Excellence Fund to BrainsCAN, and a Discovery grant by the Natural Sciences and Engineering Research Council of Canada. We wish to thank Cheryl Vander Tuin, Whitney Froese, Hannah Pettypiece, and Miranda Bellyou for animal preparation and care, Dr. Alex Li for scanning assistance, Dr. Kyle Gilbert and Peter Zeman for coil designs and Dr. Kevin Johnston for video footage assistance and his careful reading of this manuscript.

**Data and code availability statement**

Raw data and code supporting this study are available on OSF at https://osf.io/m6d58/?view_only=41b350b0049a4698b5e017fff02973a7.

**Competing interests**

Authors do not report any conflicts of interest.

## References


1. S. Caspers, K. Zilles, A. R. Laird, S. B. Eickhoff, ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage* **50**, 1148–1167 (2010).

2. E. S. Cross, A. F. de C. Hamilton, D. J. M. Kraemer, W. M. Kelley, S. T. Grafton, Dissociable substrates for body motion and physical experience in the human action observation network. *Eur. J. Neurosci.* **30**, 1383–1392 (2009).

3. V. Gazzola, G. Rizzolatti, B. Wicker, C. Keysers, The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *NeuroImage* **35**, 1674–1684 (2007).

4. G. Buccino, *et al.*, Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.*, 6 (2001).

5. B. A. Urgen, S. Pehlivan, A. P. Saygin, Distinct representations in occipito-temporal, parietal, and premotor cortex during action perception revealed by fMRI and computational modeling. *Neuropsychologia* **127**, 35–47 (2019).

6. K. Nelissen, G. Luppino, W. Vanduffel, G. Rizzolatti, G. A. Orban, Observing Others: Multiple Action Representation in the Frontal Lobe. *Science* **310**, 332–336 (2005).

7. K. Nelissen, *et al.*, Action Observation Circuits in the Macaque Monkey Cortex. *J. Neurosci.* **31**, 3743–3756 (2011).

8. J. Jastorff, C. Begliomini, M. Fabbri-Destro, G. Rizzolatti, G. A. Orban, Coding Observed Motor Acts: Different Organizational Principles in the Parietal and Premotor Cortex of Humans. *J. Neurophysiol.* **104**, 128–140 (2010).

9. L. Fogassi, *et al.*, Parietal Lobe: From Action Organization to Intention Understanding. *Science* **308**, 662–667 (2005).

10. L. Bonini, *et al.*, Ventral Premotor and Inferior Parietal Cortices Make Distinct Contribution to Action Organization and Intention Understanding. *Cereb. Cortex* **20**, 1372–1385 (2010).

11. J. Sliwa, W. A. Freiwald, A dedicated network for social interaction processing in the primate brain. *Science* **356**, 745–749 (2017).

12. P. A. Fiave, S. Sharma, J. Jastorff, K. Nelissen, Investigating common coding of observed and executed actions in the monkey brain using cross-modal multi-variate fMRI classification. *NeuroImage* **178**, 306–317 (2018).

13. S. H. Johnson-Frey, *et al.*, Actions or Hand-Object Interactions? Human Inferior Frontal Cortex and Action Observation. *Neuron* **39**, 1053–1058 (2003).

14. G. Rizzolatti, L. Craighero, The Mirror-Neuron System. *Annu. Rev. Neurosci.* **27**, 169–192 (2004).

15. E. E. Condy, *et al.*, Characterizing the Action-Observation Network Through Functional Near-Infrared Spectroscopy: A Review. *Front. Hum. Neurosci.* **15**, 627983 (2021).

16. E. Kilroy, S. A. Cermak, L. Aziz-Zadeh, A Review of Functional and Structural Neurobiology of the Action Observation Network in Autism Spectrum Disorder and Developmental Coordination Disorder. *Brain Sci.* **9**, 75 (2019).

17. B. A. Urgen, A. P. Saygin, Predictive processing account of action perception: Evidence from effective connectivity in the action observation network. *Cortex* **128**, 132–142 (2020).

18. J. M. Kilner, K. J. Friston, C. D. Frith, Predictive coding: an account of the mirror neuron system. *Cogn. Process.* **8**, 159–166 (2007).

19. J. Jastorff, I. D. Popivanov, R. Vogels, W. Vanduffel, G. A. Orban, Integration of shape and motion cues in biological motion processing in the monkey STS. *NeuroImage* **60**, 911–921 (2012).

20. J. Jastorff, R. O. Abdollahi, F. Fasano, G. A. Orban, Seeing biological actions in 3D: An fMRI study. *Hum. Brain Mapp.* **37**, 203–219 (2016).

21. J. Jastorff, G. A. Orban, Human Functional Magnetic Resonance Imaging Reveals Separation and Integration of Shape and Motion Cues in Biological Motion Processing. *J. Neurosci.* **29**, 7315–7329 (2009).

22. D. I. Perrett, *et al.*, Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: A preliminary report. *Behav. Brain Res.* **16**, 153–170 (1985).

23. S. Bruni, *et al.*, Cortical and subcortical connections of parietal and premotor nodes of the monkey hand mirror neuron network. *Brain Struct. Funct.* **223**, 1713–1729 (2018).

24. M. N. Evangeliou, V. Raos, C. Galletti, H. E. Savaki, Functional Imaging of the Parietal Cortex during Action Execution and Observation. *Cereb. Cortex* **19**, 624–639 (2009).

25. S. Rozzi, P. F. Ferrari, L. Bonini, G. Rizzolatti, L. Fogassi, Functional organization of inferior parietal lobule convexity in the macaque monkey: electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. *Eur. J. Neurosci.* **28**, 1569–1588 (2008).

26. L. Fogassi, G. Luppino, Motor functions of the parietal lobe. *Curr. Opin. Neurobiol.* **15**, 626–631 (2005).

27. M. Lanzilotto, *et al.*, Anterior Intraparietal Area: A Hub in the Observed Manipulative Action Network. *Cereb. Cortex* **29**, 1816–1833 (2019).

28. K. Ogawa, T. Inui, Neural representation of observed actions in the parietal and premotor cortex. *NeuroImage* **56**, 728–735 (2011).

29. L. Bonini, C. Rotunno, E. Arcuri, V. Gallese, Mirror neurons 30 years later: implications and applications. *Trends Cogn. Sci.* **26**, 767–781 (2022).

30. F. Filimon, J. D. Nelson, D. J. Hagler, M. I. Sereno, Human cortical representations for reaching: mirror neurons for execution, observation, and imagery. *NeuroImage* **37**, 1315–1328 (2007).

31. V. Gazzola, C. Keysers, The Observation and Execution of Actions Share Motor and Somatosensory Voxels in all Tested Subjects: Single-Subject Analyses of Unsmoothed fMRI Data. *Cereb. Cortex* **19**, 1239–1255 (2009).

32. J. Grèzes, J. L. Armony, J. Rowe, R. E. Passingham, Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. *NeuroImage* **18**, 928–937 (2003).

33. P. Cisek, J. F. Kalaska, Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature* **431**, 993–996 (2004).

34. S. Hihara, M. Taoka, M. Tanaka, A. Iriki, Visual Responsiveness of Neurons in the Secondary Somatosensory Area and its Surrounding Parietal Operculum Regions in Awake Macaque Monkeys. *Cereb. Cortex* **25**, 4535–4550 (2015).

35. A. Kraskov, N. Dancause, M. M. Quallo, S. Shepherd, R. N. Lemon, Corticospinal Neurons in Macaque Ventral Premotor Cortex with Mirror Properties: A Potential Mechanism for Action Suppression? *Neuron* **64**, 922–930 (2009).

36. V. Papadourakis, V. Raos, Neurons in the Macaque Dorsal Premotor Cortex Respond to Execution and Observation of Actions. *Cereb. Cortex* **29**, 4223–4237 (2019).

37. L. Simone, M. Bimbi, F. Rodà, L. Fogassi, S. Rozzi, Action observation activates neurons of the monkey ventrolateral prefrontal cortex. *Sci. Rep.* **7**, 44378 (2017).

38. D. Tkach, J. Reimer, N. G. Hatsopoulos, Congruent Activity during Action and Action Observation in Motor Cortex. *J. Neurosci.* **27**, 13241–13250 (2007).

39. W. Suzuki, *et al.*, Mirror Neurons in a New World Monkey, Common Marmoset. *Front. Neurosci.* **9** (2015).

40. K. D. Johnston, K. Barker, L. Schaeffer, D. Schaeffer, S. Everling, Methods for chair restraint and training of the common marmoset on oculomotor tasks. *J. Neurophysiol.* **119**, 1636–1646 (2018).

41. K. M. Gilbert, *et al.*, Simultaneous functional MRI of two awake marmosets. *Nat. Commun.* **12**, 6608 (2021).

42. A. C. Silva, *et al.*, "Longitudinal functional magnetic resonance imaging in animal models" in *Magnetic Resonance Neuroimaging: Methods and Protocols*, M. Modo, J. W. M. Bulte, Eds. (Humana Press, 2011), pp. 281–302.

43. D. J. Schaeffer, *et al.*, Integrated radiofrequency array and animal holder design for minimizing head motion during awake marmoset functional magnetic resonance imaging. *NeuroImage* **193**, 126–138 (2019).

44. W. B. Handler, *et al.*, Design and construction of a gradient coil for high resolution marmoset imaging. *Biomed. Phys. Eng. Express* **6**, 045022 (2020).

45. R. W. Cox, AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).

46. S. M. Smith, *et al.*, Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219 (2004).

47. X. Li, P. S. Morgan, J. Ashburner, J. Smith, C. Rorden, The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J. Neurosci. Methods* **264**, 47–56 (2016).

48. C. Liu, *et al.*, A digital 3D atlas of the marmoset brain based on multi-modal MRI. *NeuroImage* **169**, 106–116 (2018).

49. B. B. Avants, *et al.*, A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage* **54**, 2033–2044 (2011).

50. J. Hwang, A. R. Mitz, E. A. Murray, NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. *J. Neurosci. Methods* **323**, 13–21 (2019).

51. Y. Hori, *et al.*, Interspecies activation correlations reveal functional correspondences between marmoset and human brain areas. *Proc. Natl. Acad. Sci.* **118**, e2110980118 (2021).

52. C.-C. Hung, *et al.*, Functional Mapping of Face-Selective Regions in the Extrastriate Visual Cortex of the Marmoset. *J. Neurosci.* **35**, 1160–1172 (2015).

53. J. Schultz, C. D. Frith, Animacy and the prediction of behaviour. *Neurosci. Biobehav. Rev.* **140**, 104766 (2022).

54. A. Casile, Mirror neurons (and beyond) in the macaque brain: An overview of 20 years of research. *Neurosci. Lett.* **540**, 3–14 (2013).

55. A. Hafri, J. C. Trueswell, R. A. Epstein, Neural Representations of Observed Actions Generalize across Static and Dynamic Visual Input. *J. Neurosci.* **37**, 3056–3071 (2017).

56. L. Bonini, P. F. Ferrari, Evolution of mirror systems: a simple mechanism for complex cognitive functions. *Ann. N. Y. Acad. Sci.* **1225**, 166–175 (2011).

57. G. di Pellegrino, L. Fadiga, L. Fogassi, V. Gallese, G. Rizzolatti, Understanding motor events: a neurophysiological study. *Exp. Brain Res.* **91**, 176–180 (1992).

58. V. Gallese, L. Fadiga, L. Fogassi, G. Rizzolatti, Action recognition in the premotor cortex. *Brain* **119**, 593–609 (1996).

59. G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogassi, Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.*, 131–141 (1996).

60. H. Okano, Current Status of and Perspectives on the Application of Marmosets in Neurobiology. *Annu. Rev. Neurosci.* **44**, 27–48 (2021).

61. J. C. Cléry, Y. Hori, D. J. Schaeffer, R. S. Menon, S. Everling, Neural network of social interaction observation in marmosets. *eLife* **10**, e65012 (2021).

62. D. J. Schaeffer, *et al.*, Face selective patches in marmoset frontal cortex. *Nat. Commun.* **11**, 4856 (2020).

63. M. A. Giese, G. Rizzolatti, Neural and Computational Mechanisms of Action Processing: Interaction between Visual and Motor Representations. *Neuron* **88**, 167–180 (2015).

64. T. Jellema, C. I. Baker, B. Wicker, D. I. Perrett, Neural Representation for the Perception of the Intentionality of Actions. *Brain Cogn.* **44**, 280–302 (2000).

65. M. Kilintari, V. Raos, H. E. Savaki, Involvement of the Superior Temporal Cortex in Action Execution and Action Observation. *J. Neurosci.* **34**, 8999–9011 (2014).

66. P. Kuravi, V. Caggiano, M. Giese, R. Vogels, Repetition suppression for visual actions in the macaque superior temporal sulcus. *J. Neurophysiol.* **115**, 1324–1337 (2016).

67. D. I. Perrett, *et al.*, Frameworks of analysis for the neural representation of animate objects and actions. *J. Exp. Biol.* **146**, 87–113 (1989).

68. J. Vangeneugden, *et al.*, Distinct Mechanisms for Coding of Visual Actions in Macaque Temporal Cortex. *J. Neurosci.* **31**, 385–401 (2011).

69. J. Vangeneugden, M. V. Peelen, D. Tadin, L. Battelli, Distinct Neural Mechanisms for Body Form and Body Motion Discriminations. *J. Neurosci.* **34**, 574–585 (2014).

70. M. Kilintari, V. Raos, H. E. Savaki, Grasping in the Dark Activates Early Visual Cortices. *Cereb. Cortex* **21**, 949–963 (2011).

71. P. O. Jacquet, A. Avenanti, Perturbing the Action Observation Network During Perception and Categorization of Actions' Goals and Grips: State-Dependency and Virtual Lesion TMS Effects. *Cereb. Cortex* **25**, 598–608 (2015).

72. M. Stosic, M. Brass, N. Van Hoeck, N. Ma, F. Van Overwalle, Brain activation related to the perception of minimal agency cues: The role of the mirror system. *NeuroImage* **86**, 364–369 (2014).

73. M. A. Umiltà, *et al.*, I Know What You Are Doing: A Neurophysiological Study. *Neuron* **31**, 155–165 (2001).

74. P. Molenberghs, R. Cunnington, J. B. Mattingley, Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neurosci. Biobehav. Rev.* **36**, 341–349 (2012).

75. L. Aziz-Zadeh, S. M. Wilson, G. Rizzolatti, M. Iacoboni, Congruent Embodied Representations for Visually Presented Actions and Linguistic Phrases Describing Actions. *Curr. Biol.* **16**, 1818–1823 (2006).

76. G. Buccino, F. Binkofski, L. Riggio, The mirror neuron system and action recognition. *Brain Lang.* **89**, 370–376 (2004).

77. S. Ferri, *et al.*, A human homologue of monkey F5c. *NeuroImage* **111**, 251–266 (2015).

78. J. M. Kilner, A. Neal, N. Weiskopf, K. J. Friston, C. D. Frith, Evidence of Mirror Neurons in Human Inferior Frontal Gyrus. *J. Neurosci.* **29**, 10153–10159 (2009).

79. L. Kislinger, Photographs of Actions: What Makes Them Special Cues to Social Perception. *Brain Sci.* **11**, 1382 (2021).

80. R. Saxe, D.-K. Xiao, G. Kovacs, D. I. Perrett, N. Kanwisher, A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* **42**, 1435–1446 (2004).

81. M. Matelli, G. Luppino, Parietofrontal circuits for action and space perception in the macaque monkey. *NeuroImage* **14**, 27–32 (2001).

82. D. Arnstein, F. Cui, C. Keysers, N. M. Maurits, V. Gazzola, μ-Suppression during Action Observation and Execution Correlates with BOLD in Dorsal Premotor, Inferior Parietal, and SI Cortices. *J. Neurosci.* **31**, 14243–14249 (2011).

83. T. T.-J. Chong, R. Cunnington, M. A. Williams, N. Kanwisher, J. B. Mattingley, fMRI Adaptation Reveals Mirror Neurons in Human Inferior Parietal Cortex. *Curr. Biol.* **18**, 1576–1580 (2008).

84. M. Iacoboni, *et al.*, Cortical Mechanisms of Human Imitation. *Science* **286**, 2526–2528 (1999).

85. G. Rizzolatti, L. Fogassi, V. Gallese, Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* **2**, 661–670 (2001).

86. V. Gallese, L. Fogassi, L. Fadiga, G. Rizzolatti, "Action representation and the inferior parietal lobule" in *Attention and Performance XIX*, W. Prinz&B. Hommel, (Oxford University Press, 2002), pp. 247–266.

87. M. Matelli, R. Camarda, M. Glickstein, G. Rizzolatti, Afferent and efferent projections of the inferior area 6 in the macaque monkey. *J. Comp. Neurol.* **251**, 281–298 (1986).

88. M. Petrides, D. N. Pandya, Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *J. Comp. Neurol.* **228**, 105–116 (1984).

89. B. Seltzer, D. N. Pandya, Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *J. Comp. Neurol.* **281**, 97–113 (1989).

90. K. J. Burman, *et al.*, Cortical and thalamic projections to cytoarchitectural areas 6Va and 8C of the marmoset monkey: Connectionally distinct subdivisions of the lateral premotor cortex. *J. Comp. Neurol.* **523**, 1222–1247 (2015).

91. M. Ghahremani, K. D. Johnston, L. Ma, L. K. Hayrynen, S. Everling, Electrical microstimulation evokes saccades in posterior parietal cortex of common marmosets. *J. Neurophysiol.* **122**, 1765–1776 (2019).

92. K. D. Johnston, L. Ma, L. Schaeffer, S. Everling, Alpha Oscillations Modulate Preparatory Activity in Marmoset Area 8Ad. *J. Neurosci.* **39**, 1855–1866 (2019).

93. D. J. Schaeffer, *et al.*, Task-based fMRI of a free-viewing visuo-saccadic network in the marmoset monkey. *NeuroImage* **202**, 116147 (2019).

94. J. Selvanayagam, K. D. Johnston, D. J. Schaeffer, L. K. Hayrynen, S. Everling, Functional Localization of the Frontal Eye Fields in the Common Marmoset Using Microstimulation. *J. Neurosci.* **39**, 9197–9206 (2019).

95. C. T. Miller, A. W. Thomas, S. U. Nummela, L. A. de la Mothe, Responses of primate frontal cortex neurons during natural vocal communication. *J. Neurophysiol.* **114**, 1158–1171 (2015).