ELSEVIER

NeuroImage

Editor-in-Chief
Karl J. Friston
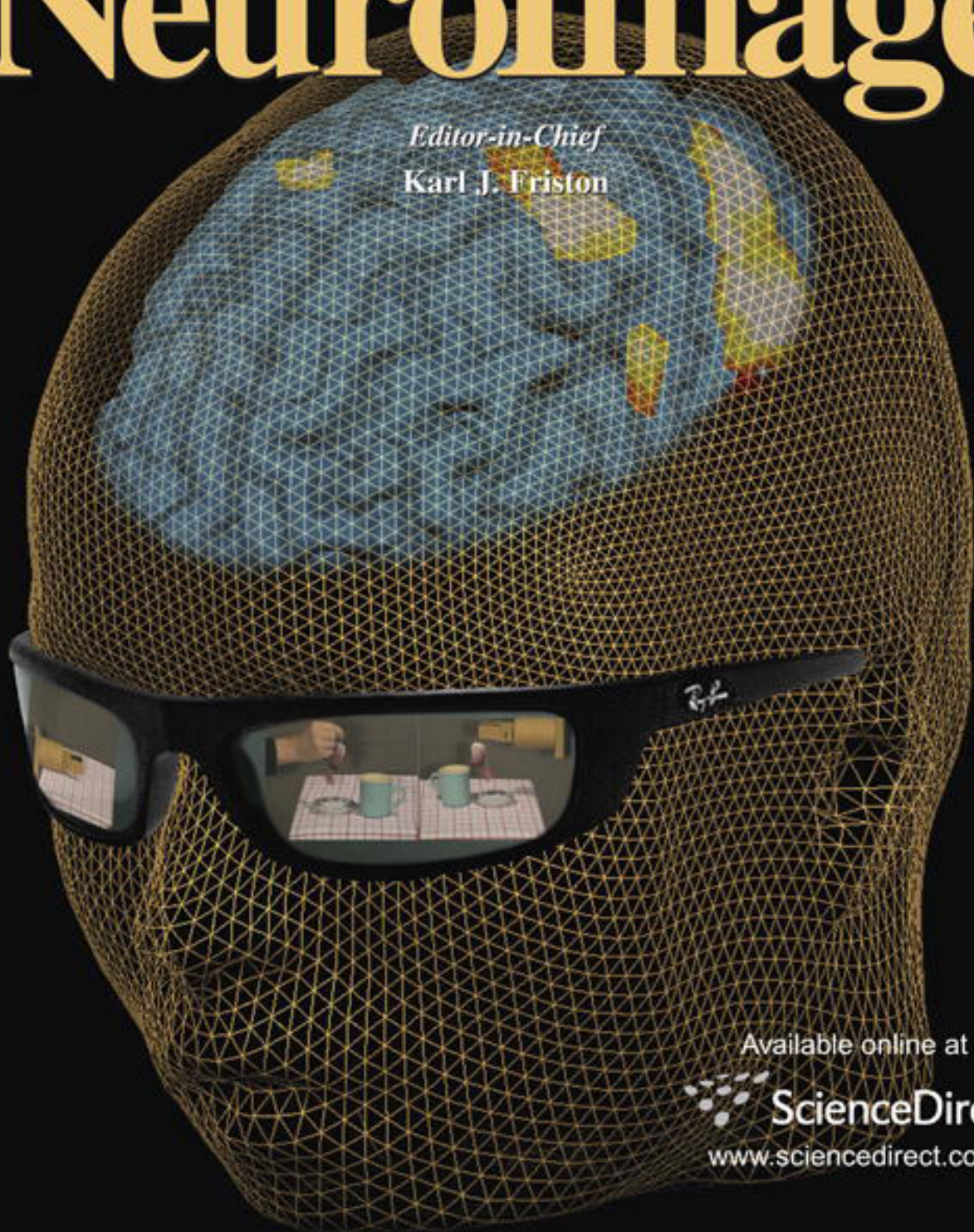
Available online at

ScienceDirect
www.sciencedirect.com

The anthropomorphic brain: The mirror neuron system responds to human and robotic actions

V. Gazzola,[a] G. Rizzolatti,[b] B. Wicker,[c] and C. Keysers[a,*]

[a]BCN Neuroimaging Center, University Medical Center Groningen, University of Groningen, A. Deusinglaan 2, 9713AW Groningen, The Netherlands
[b]Dipartimento di Neuroscienze, Universita di Parma, Via Volturno, 43100 Parma, Italy
[c]Institut de Neurosciences Cognitives de la Méditerranée, CNRS, 31, Chemin Joseph Aiguier 13009 Marseille, France

In humans and monkeys the mirror neuron system transforms seen actions into our inner representation of these actions. Here we asked if this system responds also if we see an industrial robot perform similar actions. We localised the motor areas involved in the execution of hand actions, presented the same subjects blocks of movies of humans or robots perform a variety of actions. The mirror system was activated strongly by the sight of both human and robotic actions, with no significant differences between these two agents. Finally we observed that seeing a robot perform a single action repeatedly within a block failed to activate the mirror system. This latter finding suggests that previous studies may have failed to find mirror activations to robotic actions because of the repetitiveness of the presented actions. Our findings suggest that the mirror neuron system could contribute to the understanding of a wider range of actions than previously assumed, and that the goal of an action might be more important for mirror activations than the way in which the action is performed.
© 2007 Elsevier Inc. All rights reserved.

Introduction

Understanding the goals and intentions behind the actions of other individuals is essential for survival and for normal social functioning. The discovery of mirror neurons in the monkey has provided an opportunity to gain insights into the neural basis of this capacity. Mirror neurons in the premotor and parietal cortex of the monkey respond when the monkey executes a particular action, and when it sees or hears someone else perform similar actions (Fogassi et al., 2005; Gallese et al., 1996; Keysers et al., 2003; Kohler et al., 2002; Rizzolatti and Craighero, 2004). A number of laboratories using different methods have shown that a similar system appears to exist in humans (Buccino et al., 2001, 2004a,b; Decety et al., 1997; Fadiga et al., 1995, 2002; Gazzola et al., 2006; Grafton et al., 1996; Grezes et al., 2003; Hari et al., 1998; Iacoboni et al., 1999, 2001; Nishitani and Hari,

2000; Rizzolatti et al., 1996). By linking the actions of others to the observer's corresponding actions, the existence of mirror neurons suggests that our understanding of the actions of others derives from translating them into the vocabulary of our own actions (Gallese et al., 2004; Keysers and Gazzola, 2006; Rizzolatti and Craighero, 2004).

A question that has remained relatively uninvestigated is the visual requirements of the mirror neuron system (MNS). Does the MNS need a tight match between the observed movements and the observer's motor programs to be activated? Or is it enough that the goal that is seen matches goals that are represented within the premotor cortex of the observer? An opportunity to dissociate these two aspects of motor acts is to compare how the brain processes the sight of human motor acts against corresponding acts performed by an industrial robot. 'What' is achieved is the same (e.g. reaching an object) but 'how' it is achieved differs: industrial robots are designed to be simple and efficient. As a result their movements have kinematics that radically differ from those of humans. Measuring brain activity while subjects observe human and robotic hands perform similar actions can shed light into the visual requirements of the MNS: if the activity in the MNS is similar for the different agents, the MNS would primarily match the goal of the robotic action without requiring that the kinematics of the action match. If activity is much reduced in the robotic case, it would require a tight kinematics match.

Although this question sounds simple, results so far have been very contradictory. Tai et al. (2004) scanned volunteers using Positron Emission Tomography while having a human or a robot perform actions in the scanner room in front of the subjects. Stimuli were (a) the experimenter grasping a cylinder 3 times in a row, with the inevitable variability of human grasping and (b) an artificial arm grasping the object three times according to a preprogrammed motion, that was identical in all trials. The authors found that the MNS was activated in (a) but not in (b). From these data, it is impossible to establish whether the differences between (a) and (b) were due to differences between human and robotic kinematics, as argued by the authors, or rather to a difference in action *variability*. This variability was present in the human case, where movements obviously varied from one trial to another, but absent in the robotic case, where movements were always identical. An fMRI study performed in

monkeys suggests that at least the most anterior sector of the ventral premotor cortex (area F5a) does respond to robotic movements, while the most posterior sector (F5c) does not (Nelissen et al., 2005). This finding receives additional support from evidence that the mu-rithm is reduced both when participants view human and robotic actions (Oberman et al., 2007). Psychophysical investigations of the consequences of seeing robotic vs. human actions have lead to similarly ambiguous results: some authors find that observing robotic actions does influence an observer's own actions (e.g. Press et al., 2005) while others find that it does not (e.g. Kilner et al., 2003, see Press et al., 2005 for a review).

Evidence for the fact that the goal of motor acts is represented in the MNS is provided by a number of other investigations. In a recent study, we show that subjects born without hands and arms do activate their MNS as strongly as typically developed adults while watching *hand* actions of other individuals. This finding suggests that the MNS can bridge differences in embodiment, at least for actions with goals that are familiar to the observer (Gazzola et al., submitted for publication). Sensitivity for goals is also suggested by studies showing that the MNS is responsive to the intention with which observed motor acts are carried out (Fogassi et al., 2005; Iacoboni et al., 2005), and studies showing that at least in the parietal cortex, two actions with the same goal but different kinematics, presented one after another, cause stronger habituation than two actions with the same kinematics but different goals (Hamilton and Grafton, 2006).

On the other hand, TMS studies show that the detailed temporal characteristics of an observed human movement are reflected in the motor system of the observer (Borroni et al., 2005), showing that the MNS is very sensitive to how a movement is performed, suggesting that differences in movement kinematics between humans and robots could lead to significant differences in the brain activity. Thus, there is evidence that the MNS is sensitive to both the *goal* of an action and *how* that goal is achieved, but the question of whether the MNS could represent the goal of an action even if it is performed with kinematics that do not match human kinematics at all remains poorly investigated.

Here, we therefore tested, using fMRI, 16 subjects using a 2 agent × 4 stimuli factorial design. Subjects watched either a human or a robot perform various actions. All visual stimuli were video-clips lasting between 2.5 and 4 s. We investigated the impact of the context of the action by having both the human and the robotic agent perform actions either on simple objects (colored wooden blocks) placed on a table that suggested no further purpose or on more complex objects (such as a cocktail glass or a bowl of soup) that suggest a further purpose. Two control conditions were also included. In the first, the agent's hand entered the screen, to rest on the table. In the second, the hand of the agent was immobile on the table, with the complex objects placed in front of the hand in a pose that did not suggest interactions between the hand and the object. All conditions are recapitulated in Table 1 and frames from the movies are shown in Fig. 1.

On separate days, the same participants were requested to perform actions similar to the complex actions seen in the movies within the scanner to map the motor areas responsible for performing the observed actions. This data allowed us to chart the "mirror" voxels, i.e. voxels involved both during the vision and execution of hand actions. Additionally, participants were scanned while viewing a variety of static pictures to quantify brain responses to the various visual components of the stimuli.

Materials and methods

Subjects and general procedures

16 healthy volunteers (14 right and 2 left handed; 9 female and 7 male; mean age 31 years ranging 25–45 years) with normal or corrected to normal vision were tested on three different days. During the first scanning day subjects viewed still pictures of an industrial robot, a human being or objects; during the second day, they viewed movies with the same robot and human agent performing actions, and during the third day, they executed actions in the scanner. All subjects were informed about the content of the study on a day-by-day basis, remaining naïve to the content of the days to follow. They did not know that they would see movies of

Table 1
Experimental condition for the experiment involving movies

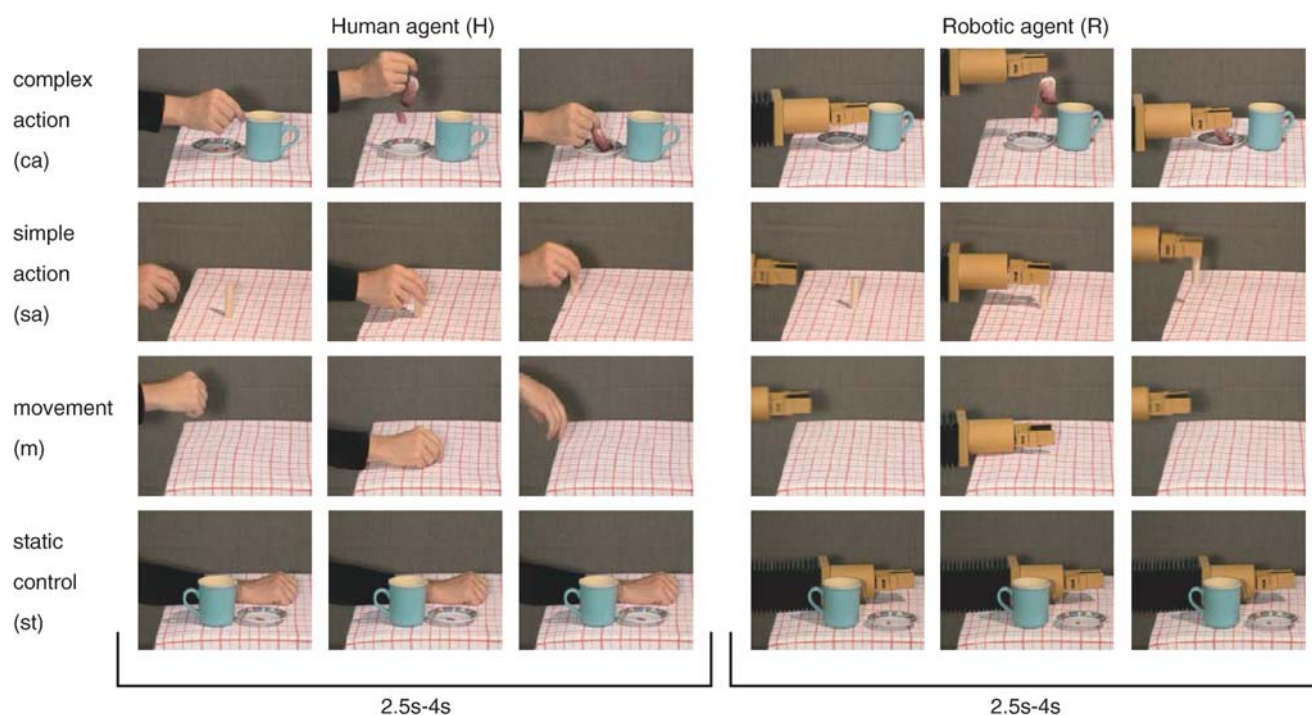| Agent	Condition name	Action description	Abbreviation
Human right arm+hand	Complex actions	–Placing a lid on a salt jar –Removing tea bag from mug and placing it on a saucer –Grasping an espresso cup –Grasping a cocktail glass by the stem –Grasping wine glass by the stem –Grasping spoon, and scooping soup out of the bowl with it	Hca
	Simple actions	6 different actions involving colored wooden blocks matching in duration and approximate complexity the complex actions	Hsa
	Static control	6 different images of the agent behind the objects used in complex actions	Hst
	Movement	The agent enters the screen and rests on the table in 6 different ways	Hm
Robot arm+claw	Complex actions	–Placing a lid on a salt jar –Removing tea bag from mug and placing it on a saucer –Grasping an espresso cup –Grasping a cocktail glass by the stem –Grasping wine glass by the stem –Grasping spoon, and scooping soup out of the bowl with it	Rca
	Simple actions	6 different actions involving colored wooden blocks matching in duration and approximate complexity the complex actions	Rsa
	Static control	6 different images of the agent behind the objects used in complex actions	Rst
	Movement	The agent enters the screen and rests on the table in 6 different ways	Rm

Fig. 1. Illustration of the movies used on day 2. Three snapshots from one of the 6 movie of each category are shown. The name of each condition is composed of the agent followed by the condition as detailed in Table 1. The duration of movies varied from 2.5 to 4 s.

actions while being scanned with the static images and they did not know that they would need to produce actions while viewing the movies of actions. This was done to avoid biasing the processing of the stimuli based on the other conditions.

Subjects signed an informed consent. All experiments were approved by the Medical Ethical Commission (METc) of the University Medical Center Groningen (NL).

Day 1: pictures

The static images used in the experiment are shown in Fig. S1, and included images of: (a) a human (right forearm and hand) or robotic (arm and claw) agent alone; (b) simple (6 wooden blocks with different shapes and colors) or complex objects (a coffee cup, a wine glass, a salt box, a cocktail glass, a tea cup and a bowl) alone; (c) the agents behind the objects. Since the basic setting was a table with a chequered red and white tablecloth, one of the two control conditions was the image of the table alone. The second control condition was the scrambled version (obtained by randomly moving square sections of the image to another location) of the human agent behind one of the complex object pasted together using photoshop with the robotic agent. Each stimulus existed in two forms: a left (agent on the left of the screen) and a right (agent on the left of the screen) version. Images were shown for 1 s in a fully randomized rapid event related design, with stimulus onset asynchrony ranging from 3 to 9 s. A fixation cross was shown between stimuli. Each category of stimuli was shown 36 times (18 right and 18 left examples) split in 6 functional runs.

Day 2: movies

This session was conducted in a 2 (agents) × 4 (stimuli) factorial design. The agents were either the industrial robot or the human

being used in the pictures of day 1. The construction of the robot was such that it could move its arm up and down, and for- and backwards. Its wrist could turn around the main axis of his arm. The robot's claw was composed of two pieces of metal that could translate parallel to the axis of the arm. Movements were controlled by a computer. To ensure that the robotic kinematics would be clearly non-biological only one degree of freedom was moved at a time (vertical or horizontal or rotational). In addition, the movements of the robot were slow, of constant velocity, and were later accelerated to match the duration of natural human movements using video-editing (Adobe Premiere pro). This resulted in linear velocity profiles that differed very much from the bell-shaped acceleration profiles typical for biological motion. The human agent was instructed to perform the actions as naturally as possible, and the most natural of 5 repetitions of each action was selected. To reduce overt visual differences, the claw of the robot was spray-painted in yellow to resemble the skin color of the human hand, and the human wore a black jumper to resemble the color of the arm of the robot.

Common to all 4 types of movies was the basic setting: the table with the chequered red and white table cloth used in day 1. The stimuli are described in Table 1. All movies were matched for duration, such that for each complex action, there was a movie with the same duration in the other conditions. Simple and complex actions were also match such that, for each complex action, there was a simple action that contained a similar movement, but performed on a simpler object. Grasping the cocktail glass by the stem for instance was matched with grasping a thin cylinder, and closing the sugar box, with placing a small cylinder on top of a larger cylinder. Each movie was filmed twice, once with the agent entering from the right, and once with the agent entering from the left. Movies were recorded using a digital video camera, elaborated using AdobePremiere (www.adobe.com), presented using Presentation (www.neuro-bs.com), projected using an LCD projector on a

semi-opaque screen placed at the head end of the bore and seen through a mirror placed on the head coil. The duration of the movies ranged from 2.5 s to 4 s.

All conditions were presented in a block design, with four exemplars of each condition picked out at random to form 13.5 s blocks containing 4 different simple actions, 4 different complex actions, 4 different movements or 4 different static images, separated by 200 ms intervals of blank screen. Consecutive blocks were separated by a 10 s pause of blank screen with a fixation cross. The experiment was split into 4 runs with a total of 12 repetitions per condition. Half of the blocks depicted only agents entering from the right of the screen, and half only from the left. Subjects were instructed to watch the movies carefully, paying particular attention to the relationship between the agents and the objects.

Day 3: motor task

Subjects performed a single run of motor testing. Before scanning the subject was shown the T-shaped table that would be placed on his/her lap during scanning. The table contained 4 objects. The two lateral branches of the T contained two high-stemmed plastic glasses. The intersection of the T contained a plastic cup with a handle. The bottom of the T contained a plastic bowl with a plastic spoon. Subjects were then trained in their task. The task sequence was as follows. At the commencement of each trial subjects viewed a diagram of the table on the screen, with a pink rectangle at the left or right to indicate what hand to use, and a red cross in one of the four object locations indicated which objects they had to act upon. When the red cross turned to green, subject had to perform the action compatible with the object. For the glass, they had to reach for the glass, grasp it, bring it towards their face, then stopping half way along that movement, they had to place it back in its location. For the cup of coffee, they had to do the same action, but grasping the cup by the handle. For the bowl, they had to perform the same action as above, but with the spoon, as if drinking soup with a spoon. In all cases, they were instructed that in the scanner, they view of the table and object will be blocked by the head coil and their thorax, and that they need to train to perform the action without looking at their own action. In the scanner, subjects' heads and upper arms were firmly strapped onto the scanner bed to avoid that the actions would lead to significant head motion. They then underwent a number of training trials to check that they could perform the actions accurately. For all subjects, we ensured that they were unable to see their own actions due to the visual obstacles. The timing of the actions was rehearsed to last approximately 5 s, but an experimenter within the scanner room documented the beginning and end of each action using a button box to determine the actual duration of the action that was then used to define the design matrix of the experiment for data analysis. Within a single scanning session of 500 s, subjects performed eighteen ~5 actions with their right hand and 18 with their left. Their arms never crossed the table (i.e. right hand only grasped the right glass, and left hand only the left glass), and the 18 actions were composed of 6 actions involving each of the three objects (hand action execution — HandExe). Conditions were fully randomized with 13±2 s lapsing between the onset of two conditions.

fMRI

Scanning was performed using a Philips Intera 3T Quaser, a synergy SENSE head coil, 30 mT/m gradients and a standard single shot EPI with TE=30 ms, TA=TR=2 s, 39 axial slices of 3 mm thickness, with no slice gap and a 3×3 mm in plane resolution acquired to cover the entire brain and cerebellum.

General data processing

Data were preprocessed using SPM2 (www.fil.ion.ucl.ac.uk/spm). Data from the three days were slice time corrected. EPI images from all sessions were aligned to the first volume of the second day of scanning. High quality T1 images were co-registered to the mean EPI image and segmented. The co-registered gray matter segment was normalized onto the gray matter template and the resulting normalization parameters applied to all EPI images. Smoothing using 6×6×6 mm FWHH was applied to all normalized EPI images. Data were then analyzed by applying a general linear model separately for each individual, either voxel-by-voxel, using SPM, or to the mean signal of the voxels contained in a region of interest (ROI), using MarsBar (http://marsbar.sourceforge.net). The contrast values obtained for each subject were then analyzed at the second level using *t*-tests to implement a random effects analysis. All conditions were modeled using a box-car function convolved with the hemodynamic response function (HRF). Additional predictors of no interest were modeled to account for translation and rotation along the three possible dimensions as measured during the realignment procedure.

Analysis of the static images (day 1)

Data from each scanning session of the static images were analyzed using a general linear model with a separate predictor for the left and right versions of each stimulus category convolved with the HRF. Parameter estimates for the right and left versions of each category over all runs were then added together to provide a single estimate of the brain activity to each stimulus category. At the second level, these estimates were compared against zero or against each other using *t*-tests.

Analysis of the movies (day 2)

Data from each scanning session of the movies were analyzed using a general linear model with a separate predictor for the left and right versions of each stimulus category convolved with the HRF. Parameter estimates for the right and left versions of each category over all runs were then added together to provide a single estimate of the brain activity to each stimulus category. Brain activity to the control stimuli (static and movement) was either compared against zero or against those of the other agent using linear contrasts. Activity to the simple and complex actions was either compared against the static control stimulus or against each other using linear contrasts at the first level. At the second level, these parameter estimates or contrast values were compared against zero using *t*-tests.

Analyses of motor task (day 3)

Data were analyzed using a general linear model with two experimental conditions (actions performed with the right and action performed with the left hand—HandExeR, HandExeL), modeled as box-car function convolved with the HRF. Inspection of the realignment parameters revealed that head motion within the motor run was below one voxel of translation and 3° of rotation in all

subjects. We nevertheless included all 6 motion parameters (3 translations, 3 rotations) as covariates in the analysis. For each subject, this resulted in a parameter estimate for actions of the right and actions for the left hand. These parameter estimates were analyzed separately at the second level by testing all 16 parameter estimates for actions with the right hand against 0, and doing the same for the left hand. The two t-maps were then combined using a maximum function that corresponds to a logical or, and thresholded at t 3.73 ($p_{unc} < 0.001$) to identify voxels involved in performing actions with the right or the left hand.

Masking

As this manuscript is concerned with the activation of the MNS, all visual results were masked inclusively with the motor execution results of day 3 and exclusively with these of the scrambled pictures of day 1 (Fig. S2a,b): at the second level of analysis, a voxel had t 3.73 during motor execution and $t < 3.73$ during the vision of scrambled pictures. This restricted results to areas also involved in motor execution, and not responding to meaningless visual patterns (Fig. S2c).

Subject-by-subject maps

Here each single subject was analyzed entirely at the first level of analysis requiring the same probability cutoffs ($p_{unc} < 0.001$) used at the second level. For the contrast Hsa–Hst for instance, a voxel was considered mirror if (Hsa–Hst 3.11) and (max(HandExeR,HandExeL) 3.11) and (scrambled pictures < 3.11), where Hsa–Hst and scrambled pictures refers to the t-value of these contrasts and max(HandExeR,HandExeL) the maximum t-value of the action execution with the right or left hand. This lead to a map of 1 and 0's, where a voxel was mirror or not. The 16 single subject maps were then summed together to lead to a map where each voxel value could range between 0 (none of the subjects had a mirror voxel there) to 16 (all 16 subjects had a mirror voxel there). This was performed separately for all 4 contrasts of interest (Hsa–Hst; Hca–Hst; Rsa–Rst; Rca–Rst; Fig. 4).

Statistic thresholds

All significant differences are reported at the $p_{unc} < 0.001$ level, uncorrected for multiple comparisons, which corresponded to a $t_{(16-1)} = 3.73$ for the group analysis. The minimum cluster size was set at $k = 10$. Unless otherwise specified only clusters surviving also an FDR (False Discovery Rate) corrected threshold of $p_{FDR} < 0.05$ are reported.

Anatomical descriptions

Anatomical description were performed based on the probabilistic cytoarchitectonic maps of the brain mapping group in Juelich, Germany (http://www.fz-juelich.de/ime/ime_brain_]mapping) as implemented in the SPM anatomy toolbox (http://www.fz-juelich.de/ime/spm_anatomy_toolbox). In that approach, a maximum probability map is created of all cytoarchitectonically identified brain areas, including BA6,44,45,1,2,3,4 and the parietal operculum (OP1–4). Brain areas BA1–3 will be described as SI, OP1–4 as SII, and all other areas by the BA number. Outside of these areas, locations are described anatomically as precentral gyrus etc. This means that a reference to precentral gyrus means that the

activation was in a sector of the precentral gyrus that did not fall within any of the cytoarchitectonically identified areas.

Results

Motor execution (day 3)

In order to restrict our analyses to the MNS, we identified the brain areas involved during motor execution. These brain areas included: M1 and BA6/44 in the frontal lobe; SI, SII, IPL (SMG) and SPL in the parietal lobe and the middle temporal gyrus in the temporal lobe (abbreviations explained Table S1). Additional clusters of activation were observed in the visual areas, probably due to the visual instructions of the motor execution task. To exclude these latter activations, we used areas activated by the sight of scrambled pictures as an exclusion mask for the motor results and used the remaining activations as an inclusive mask for all further visual analyses (Fig. S2c). Results masked with the motor execution will simply be referred to as 'masked' (see supplemental online material section 2 for details).

Activations to human actions (day 2)

Fig. 2a shows that, after masking, the observation of the static hand behind the object (Hst) caused only small and relatively weak activations in bilateral posterior temporal clusters, centered on the MTG but extending into the ITG, the middle and inferior occipital gyrus and bilateral parietal activations in the SPL extending slightly into the IPL. A small frontal activation was also observed in the right anterior BA44.

The observation of human movements not involving objects, after masking (Hm, Fig. 2b), again caused only relatively weak activations in the MNS. The basic activations in the posterior temporal and parietal lobe resembled those of Hst, extending slightly further rostrally into SI, with an additional cluster in the right SMG also extending into SI. In the frontal lobe, differences were more pronounced, with the movements activating the dorsal aspects of the right BA44, extending dorsally into the precentral gyrus.

Figs. 2c and d show the masked results of comparing the two goal directed actions against the static baseline (Hsa–Hst, Hca–Hst). Both stimuli robustly activated the entire MNS. Common to both cases (by requiring t(Hsa–Hst) 3.73 and t(Hca–Hst) 3.73, both masked, $p_{unc} < 0.001$ and $p_{FDR} < 0.05$; Fig. 2f) is a relatively right–left symmetric network involving a temporal node, a parietal node and two fontal nodes. The temporal node is located in the posterior MTG. The parietal involvement is extensive, starting dorso-caudaly in the SPL and extending rostroventrally all the way to the parietal operculum (SII), passing through the IPL (SMG) and SI. The frontal nodes are composed of a dorsal cluster on each side, including BA6 and extending rostrally into the SFG/MFG and the Precentral gyrus. The ventral cluster is located in BA44, and, in the left hemisphere, extends ventro-medially into cytoarchitectonically unassigned regions of the IFG *pars opercularis* and dorso-caudaly into cytoarchitectonically unassigned regions of the precentral gyrus. Additional activations surviving masking were found bilaterally in the MCC and in the cerebellum.

A direct comparison of the simple and complex actions (Hca–Hsa or Hsa–Hca) revealed that the complex actions activated only the left lingual gyrus more strongly than the simple actions (−24 −62 −12, $k = 78$ voxels, peak $T = 7.32$), while none of the areas
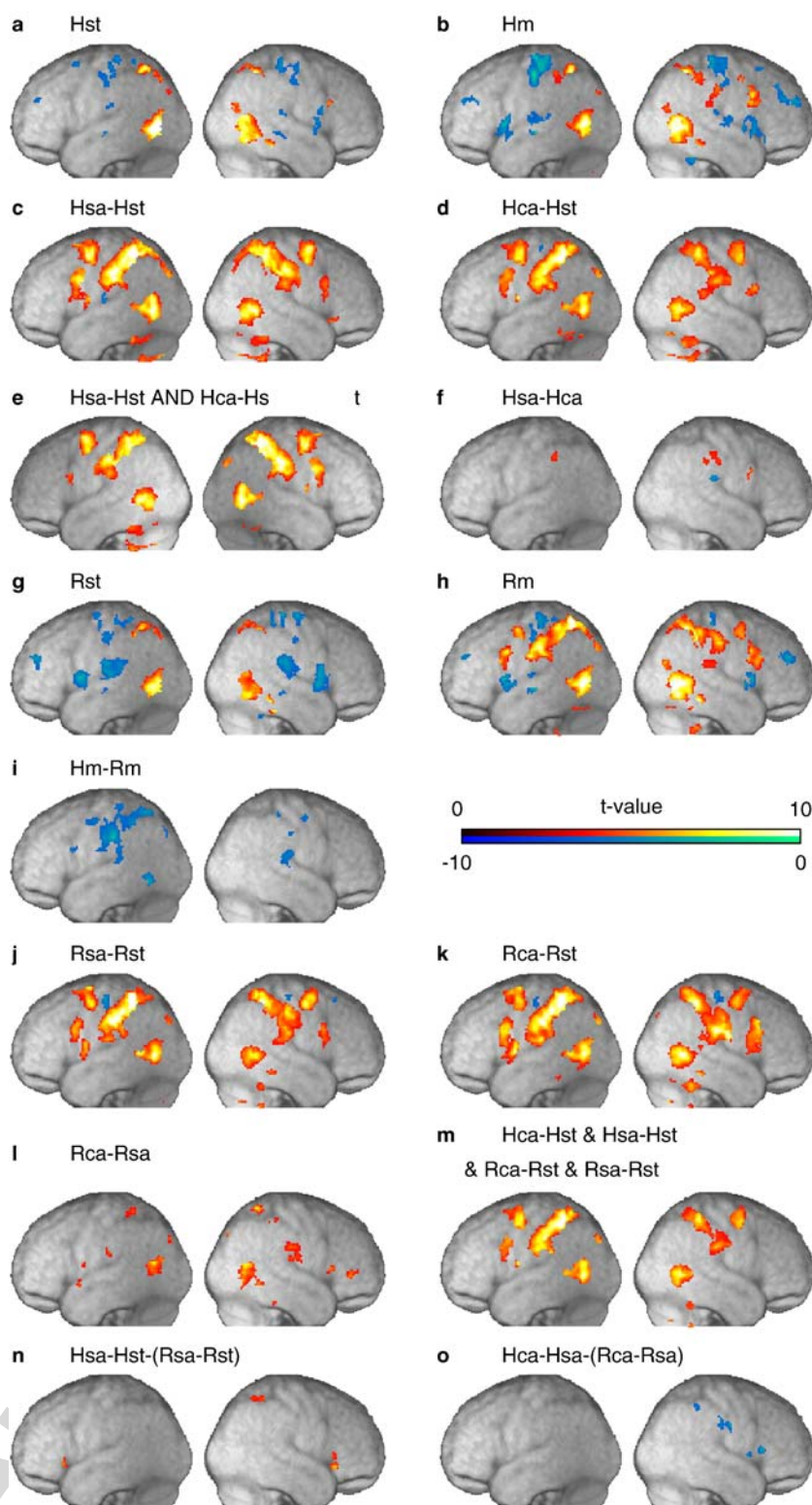
Fig. 2. Main results of day 2, as described in the text, inclusively masked with motor execution and exclusively with the viewing of scrambled images. Abbreviations as in Table 1 ($p_{unc} < 0.001$ and $p_{FDR} < 0.05$, $k = 10$).

outside the MNS responded more strongly to the simple actions. When masked, no significant differences, using FDR correction at 0.05, were found, while, without FDR correction (Fig. 2f), small differences between the conditions become apparent: a cluster of

27 voxels responded more to the Hca in the right SII (46 −26 22, blue in Fig. 2g) and conversely, the IPL (bilaterally) and the right BA44/precentral were more strongly activated by the Hsa (red in Fig. 2g).

To check if the small differences between simple and complex actions in the MNS could be due to differences in the objects themselves, possibly through the activity of canonical neurons (Grezes et al., 2003; Murata et al., 1997; Rizzolatti and Fadiga, 1998), we compared the sight of our simple and complex objects alone, collected on day 1. The comparison revealed no significant differences anywhere in our mask, indicating that the differences between simple and complex actions must arise from the interaction of the hand with the object.

The lack of differences between the sight of simple and complex objects justifies the use of a single static control (Hst, involving the complex objects) to contrast the activations of both the complex and simple actions during day 2.

Interestingly, during the vision of all stimuli, a reduction in BOLD signal was observed within the mask in M1, extending into the caudal most aspect of dorsal BA6 ($t < -3.73$, $p_{unc} < 0.001$). This reduction was strongest during the vision of complex actions.

Activations to robotic actions (day 2)

The vision of the robotic agent behind the object (Rst) activated a network very similar to the corresponding human condition (Fig. 2g), indeed a masked direct comparison of the two stimuli (Hst–Rst) revealed no significant differences.

Fig. 2h shows that the vision of the robotic movement not involving an object (Rm) caused a pattern of activation that was substantially stronger than the human counterpart. A masked direct comparison (Hm–Rm, Fig. 2i) revealed that while none of the voxels responded more strongly to the human movements (which would be shown in warm colors), much of the left MNS showed a marked preference for the robotic movements (cold colors). This involved mainly the parietal lobe (SI, SII, IPL/SPL, and precuneus), and the frontal lobe (a dorsal cluster in BA6/precentral gyrus and a ventral cluster in BA44). Smaller clusters preferring the robotic movement were also found in the right hemisphere, involving mainly SI,SII, BA4 (i.e. M1), BA6 and BA44.

Figs. 2j,k show the results of contrasting the robotic simple and complex actions against the static control. The MNS was activated in both cases. Fig. 2l shows that part of this system (bilateral MTG, SPL, SI and SII, cingulate cortex, right border between BA44/45 and IFG pars triangularis) preferred the complex actions. As for humans, an unmasked comparison of Rca–Rsa revealed that the ventral visual areas (fusiform and lingual gyri in particular) were more active during the complex actions, suggesting that given the more informative nature of the object, the processing of the complex actions draws upon a more detailed analysis of the objects.

How do the mirror activation for robotic actions compare to those for human actions? Visual comparison with the corresponding human stimuli of suggests (i) that a basic circuit common to all these cases exists, and (ii) that subtle differences do exist between the human and robotic stimuli.

(i) Fig. 2m shows the masked result of applying a logical 'and' function between the two robotic and the two human actions after subtraction of their static controls. This core MNS, common to all human and robotic actions and also involved in the execution of similar actions, included bilateral MTG activations, bilateral parietal clusters extending from the SPL to SII, passing through the IPL (SMG), SI and frontal clusters. The bilateral dorsal frontal activation includes BA6 and extends into MFG/SFG and the precentral gyrus and the ventral cluster, restricted to the left hemisphere, includes BA44 and extends into the precentral gyrus.

Additional clusters common to all cases were found in the right cerebellum and the left insula and MCC. The core MNS demonstrates the fact that our brain maps the sight of both robotic and human actions onto our own motor representations, irrespectively of whether the robot engages in actions that are typically human (Hca) or actions that are less linked to human behavior (Hsa).

(ii) To investigate the differences between the human and the robotic actions, we compared the activation to the human and robotic simple actions (Hsa–Hst)–(Rsa–Rst). Using FDR correction, this comparison reveals no differences. Without FDR correction, but at $p_{unc} < 0.001$, a number of smaller clusters located in bilateral IFG *pars orbitalis* and right SPL, BA6, BA45 responded more strongly to the human simple actions (Fig. 2n). Robotic simple actions did not activate any voxels more than the human counterpart. For the complex actions, no significant differences between the robotic and the human counterpart were found, both with and without FDR correction. A direct analysis of interaction ((Hca–Hsa)–(Rca–Rsa); Fig. 2o) showed no significant differences when FDR correction was used. At $p_{unc} < 0.001$, the same comparison demonstrated the existence of subtle differences: a network composed of right parietal (SPL, SI), SMA and BA44/45 benefited more from a complex object if the agent was robotic.

For a more quantitative analysis of the sensitivity of the MNS to robots, we defined the main clusters identified in Fig. 2e (i.e. dorsal premotor, ventral premotor, parietal and MTG in both hemispheres), and extracted the mean parameter estimates in these 4 areas of each hemisphere in our 8 experimental conditions. Results are shown in Fig. 3, with the time course of the activations shown in Fig. S3. Overall, the most striking finding is the relative similarity of human and robotic stimuli, both in terms of parameter estimates and time courses, despite the fact that the regions of interest where selected based on the brain activations for the human stimuli (thereby biasing the results systematically in favor of the human stimuli). Results were examined using analyses of variance and Newman–Keuls post-hoc comparisons. In all areas, for the static control condition, there were no significant differences between robots and humans. For the movement condition robotic stimuli caused stronger activations in the dorsal premotor, left ventral premotor and left parietal cortex (all $p < 0.01$), while only in the right ventral premotor cortex the human movement caused stronger activations (all $p < 0.01$). A subtle but consistent effects was observed when comparing the simple and complex actions for the two agents. Analysing the brain activity for the 2 types of actions in a 4 regions (dPM, vPM, parietal, MTG) $\times 2$ hemisphere (L/R) $\times 2$ agents (human/robot) $\times 2$ actions (simple/complex) repeated measurement ANOVA revealed an interaction of agent and action ($p < 0.01$) that did not depend on the region or hemisphere (both p 0.5). This interaction is of a cross over type: while the mirror activations *decrease* for the human stimuli when a complex object is introduced, they *increase* for the robotic actions. In addition, to verify that responses during the observation of actions where specific to goal directed actions, we compared, in all eight regions of interest separately, brain activations during the observation of movements not directed at goals with the complex actions (i.e. Hca–Hm, Rca–Rm). In all, but the right middle temporal cluster, the difference was significant (all $p_{unc} < 0.005$, Newman–Keuls post hoc; Fig. 3).

Finally, to investigate whether the activations found in the MNS to all 4 types of actions were robust, we examined the presence of activations in the MNS for all four basic contrasts (Hsa–Hst, Hca–
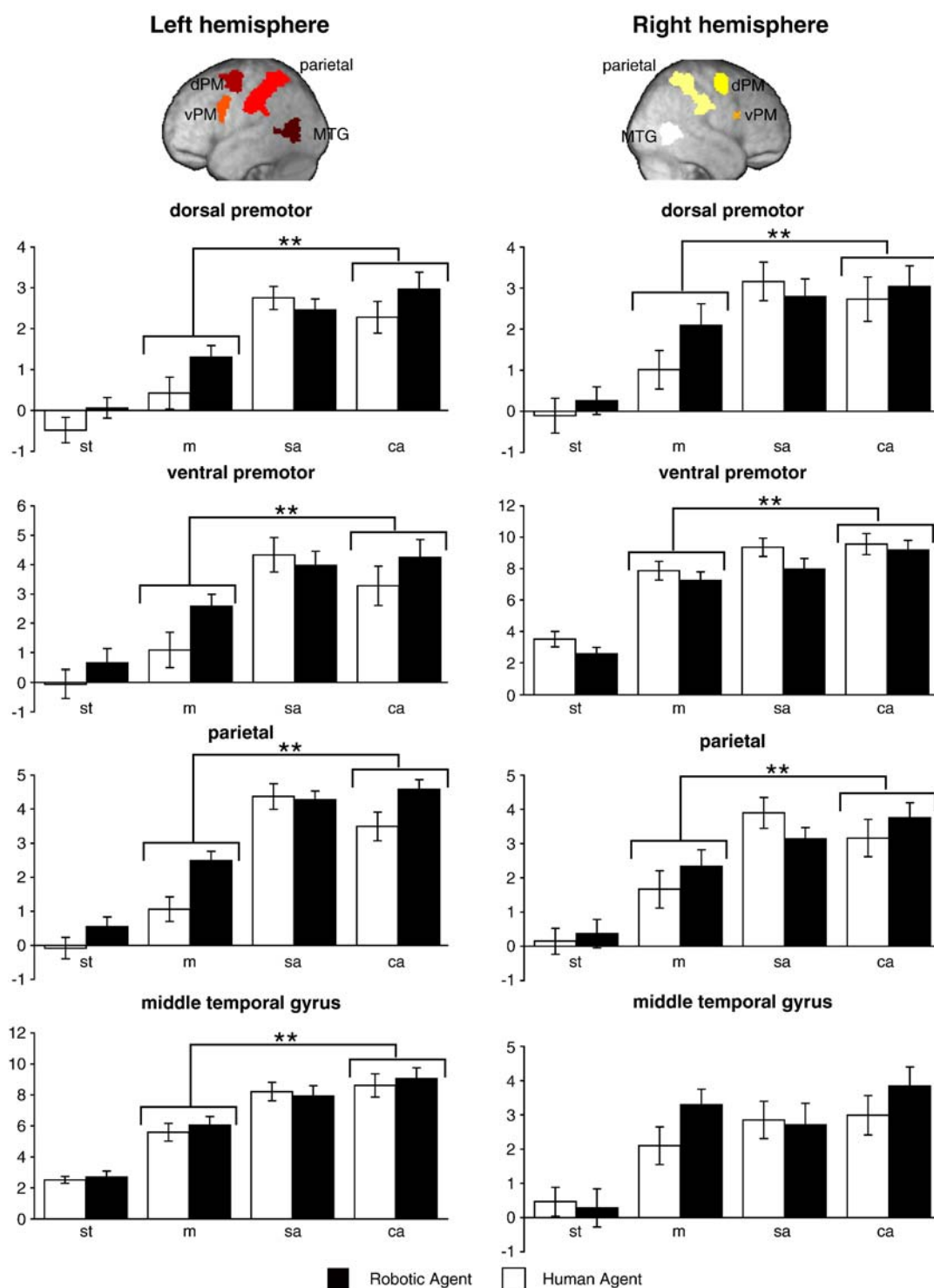
Fig. 3. Contrast estimates for the eight movie conditions in the 8 main mirror regions as defined in Fig. 2e. The parameter estimates are shown in arbitrary units on the y-axis (see Fig. S2 for approximate conversion in % signal change). The x-axis separates the four stimulus classes, with robotic and human agents shown separately. The 8 regions of interest are shown in arbitrary colors. Error bars refer to standard error of the mean. Abbreviations: st=static control, m=movement, sa=simple actions, ca=complex actions. $**p_{\text{Newman–Keuls}} < 0.005$.

Hst, Rsa–Rst, Rca–Rst) in our single subjects. Fig. 4 shows the result of those single subject analyses by showing for each voxel how many of the 16 subjects showed mirror properties in that voxel. As can be seen, MNS for both human and robotic actions could be robustly demonstrated in a vast majority of our subjects.

Discussion

In this experiment, we compared the brain activity in the MNS while subjects observed humans and robots perform various actions. The aim of the experiment was to identify whether, to be
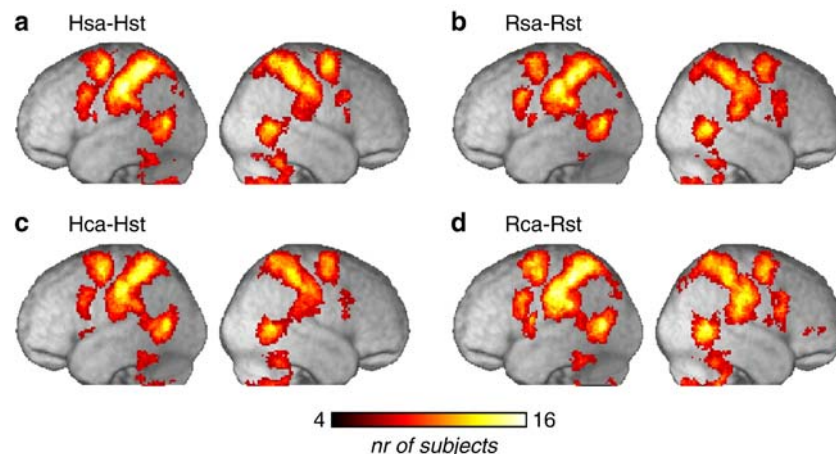
Fig. 4. Summary maps of subject-by-subject definitions of mirror activity. The maps represent for each voxel how many subjects satisfied the criteria for mirror voxel in that voxel. Only voxels where at least 5 subjects showed mirror activity and clusters of at least 10 voxels are shown.

activated, the human MNS requires a tight match between the kinematics of the observed actions and the actions within the observer's motor vocabulary. As implied goals of motor acts (Umilta et al., 2001) and even actions (Fogassi et al., 2005; Iacoboni et al., 2005) may affect the response of the MNS, we examined two types of actions: complex action with distal action goals richly embedded in human behavior and clearly suggesting a further stream of actions (e.g. grasping a cocktail glass) and simple motor acts less suggestive of further actions (e.g. moving wooden block around).

In our experiment all actions, be they human or robotic, simple or complex, activated a significant section of the areas involved in the motor execution of similar actions, in particular the temporal, parietal and frontal areas, typically considered to compose the MNS. In contrast to previous findings (Tai et al., 2004), our data thus show that the MNS can respond to robotic actions. Given that we subtract a static control condition in which the complex objects were seen in front of the agents, these activations cannot be due to activity in canonical neurons alone (Grezes et al., 2003; Murata et al., 1997). In addition, the mirror regions responding to the sight of robotic actions responded more during the observation of goal-directed actions than similar movements not directed at goals. This supports the idea of goal matching and shows that these responses cannot reflect unspecific visual responses to any movement.

To understand why Tai et al. (2004) failed to find robotic activations in the MNS, we examined data collected during the preparation of the main experiment. In that block design study, 12 healthy volunteers were scanned while viewing (a) a robot perform 5 *different* actions within a block, as in our main experiment, or (b) 5 times the exact same action as in Tai et al. (2004). We found significant activations in the MNS in (a) but not in (b) (Gazzola et al., 2004). These findings further support the main results of our experiment: the observation of actions done by another agent activates the MNS irrespectively of the agent provided that different movies are shown within a block. If the same movies are repeated over and over again, the effect is much reduced, a finding in accord with the observation that repeating identical movies causes strong habituation in the MNS (Hamilton and Grafton, 2006).

The lack of significant differences between MNS activity during the observation of robotic and human actions is intriguing, as it places new constraints on models of how mirror neurons

develop, and what they code. The first studies investigating the plasticity of the MNS (Bangert et al., 2006; Calvo-Merino et al., 2005) showed that the MNS responds more strongly to actions for which the observer has motor expertise. Given that we have more motor expertise with human actions than with robotic actions, one might predict stronger mirror activation to the sight of human actions. It is important though to differentiate experience with the goal of an action from experience with the detailed kinematics of an action. During acquisition of piano or dance skills, students develop familiarity with both the kinematics and the goal of the actions, while in our experiment, observers were equally familiar with the immediate goal of the human and robotic agent: grasping and manipulating objects. Keeping in mind that a lack of significant difference between human and robotic agents does not prove that brain responses are equal in these conditions, the similarity of the activations in our study may nevertheless suggest that the difference in familiarity with the kinematics of the robotic actions had little impact on mirror activations. This corroborates our recent finding that individuals born without hands or arms activate the MNS to the same extend as typically developed individuals while observing hand actions of others (Gazzola et al., submitted for publication). Together these experiments suggest that if observers are highly familiar with the goal of an action, deviations in kinematics can be ignored, and a full-blown resonance of the observer's actions can be measured. Situations where the observer is not familiar with the goals may lead to differences in mirror activations depending on the degree of correspondence between observed kinematics and those of the observer.

The present experiment should thus be seen as an empirical contribution towards the emerging notion that the human MNS can represent actions of which we understand the goal, but for which our own way of achieving the goal does not match with the kinematics of the observed movement (Buccino et al., 2004a). While we already knew that the goal of an action is an important element in activating the MNS (Fogassi et al., 2005; Iacoboni et al., 2005; Umilta et al., 2001), our finding goes a step further, by showing that the goal alone, without matching kinematics, is sufficient to activate our MNS. The direct-matching hypothesis of the MNS has often been interpreted as focusing on matching the details of observed and executed actions, but goal matching may be as important for the function of the MNS.

A subtle interaction between goal and agent/kinematics was apparent in our data (although it failed to survive correction for multiple comparison): while for complex actions, human and robotic agents did not differ significantly, for the simple actions, human agents determined marginally stronger activations than robotic ones. This suggests that the richness of the goal in the complex actions was sufficient to fully activate the MNS whatever the agent/kinematics. The more limited goals of the simple actions were still sufficient to determine robust activations in the MNS, but the tighter match with the observer's own kinematics could have further augmented activations in that condition. This observation is compatible with the primate literature. Mirror neurons vary along several dimensions (Gallese et al., 1996). Some, broadly tuned mirror neurons respond during the execution of particular actions (e.g. grasping with the hand), but respond to the sight of a variety of actions that achieve the same goal (different types of prehensions and even grasping with the mouth), with some of them even responding to the sight of actions performed by tools (Ferrari et al., 2005). These could provide the neural basis for the presence of activations in all of our conditions. For cases where the kinematics match (our human conditions), strictly congruent mirror neurons, requiring a tight fit of observed and executed actions, could add to the activity. Finally in cases where the goal of the actions suggests a series of actions to follow, even if these actions are not seen, occlusion mirror neurons and sequence selective neurons could add to activations (Fogassi et al., 2005; Iacoboni et al., 2005; Umilta et al., 2001). Within that framework, robotic simple actions would trigger only the basic MNS, all other conditions would, either by suggesting a sequence of motor acts, or by providing a tight fit with the observers kinematics, supplement this activity with additional mirror neurons.

An additional finding of our study was that for the human agent, the complex actions caused smaller activation than the simple actions in the MNS (although only without FDR corrections). At first sight, this finding is in contradiction with the finding of Iacoboni et al. (2005) that seeing an action in a contexts that suggests upcoming actions result in larger premotor activations: our complex actions suggest further actions (drinking etc.), yet they produced less activation. How can this discrepancy be explained? There is a fundamental difference between our experiment and that of Iacoboni et al.: in our case, the context of the actions was always the same–a table–and all that changes was the object on which the action was executed, in their case, the object goal of the action was always the same, but the context changed. It is likely that in our experiment, understanding the action could occur in two ways: in our simple cases, the object contained very little information about what will be done, while in our complex cases, the object gave valuable information about what will be done. It is thus likely that processing of the simple actions depended entirely on action related areas, while the processing of the complex actions could rely in part on processing the object. Matching that hypothesis, we found an increased mirror activity in the simple actions, and increased ventral visual activity in the fusiform areas in the complex actions. This interpretation receives independent confirmation from a study showing that while subjects observe a grasping actions, directing their attention to the object leads to an increase of activity in fusiform areas and a reduction in mirror areas, while directing attention to the action produced opposite effects (Shmuelof and Zohary, 2005). A similar increase in ventral visual areas was also observed for the robotic complex actions compared to the robotic simple actions, in accordance with that interpretation. For the robotic case though, mirror activity was actually larger for the complex actions,

suggesting that in that particular case, the ventral visual activations processing the object may have integrated with the processing of the action per se, possibly to compensate the incongruent kinematics.

A challenge for future behavioral research will be to tease apart the contribution of goals and kinematics to behavioral interference and potentiation effects that observing human and robotic movements can have on the observer's behavior (see Press et al., 2005). For instance, would studies that fail to find robotic interference (Kilner et al., 2003) find stronger effects if the robot would engage in meaningful actions?

A persisting enigma surrounding the MNS is the fact that activations within the premotor cortex do not lead to overt imitation of the observed behavior. In our dataset we observed during all visual conditions a clear reduction in the BOLD signal relative to baseline in a sector of M1 that was activated during the execution of hand actions. While it is generally difficult to interpret such reduction in BOLD signal, it could reflect the existence of a mechanism that prevents the overt motor output by inhibiting premotor output to M1 during action observation.

Finally, an interesting implication of our results, beyond neuroscience, is the fact that the understanding of actions of artificial devices can take advantage of the intricate brain mechanisms that humans have developed to understand other human beings without having to be particularly human-like. The strong activations in the MNS observed to the vision of robots in our experiments suggests that even crude industrial robots, particularly while engaging in meaningful human actions, can tap into our social brain—as long as their behaviors are not too repetitive. Once again, science seems to lag behind the arts: George Lucas was apparently well aware of the potential of robots to enter our social brains when he decided to have a set of robots star in his new science fiction saga. Now we know, that our MNS may be part of the reason why, when in Star Wars, C3PO taps R2D2 on the head in a moment of mortal danger, we cannot help but attribute them human feelings and intentions, even if their physical aspect and kinematics are far from human.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2007.02.003.

References

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H.J., Altenmuller, E., 2006. Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. NeuroImage 30 (3), 917–926.

Borroni, P., Montagna, M., Cerri, G., Baldissera, F., 2005. Cyclic time course of motor excitability modulation during the observation of a cyclic hand movement. Brain Res. 1065, 115–124.

Buccino, G., Binkofski, F., Fink, G.R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R.J., Zilles, K., Rizzolatti, G., Freund, H.J., 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. Eur. J. Neurosci. 13, 400–404.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C.A., Rizzolatti, G., 2004a. Neural circuits involved in the recognition of actions performed by nonconspecifics: an FMRI study. J. Cogn. Neurosci. 16, 114–126.

Buccino, G., Vogt, S., Ritzl, A., Fink, G.R., Zilles, K., Freund, H.J., Rizzolatti, G., 2004b. Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. Neuron 42, 323–334.

Calvo-Merino, B., Glaser, D.E., Grezes, J., Passingham, R.E., Haggard, P., 2005. Action observation and acquired motor skills: an FMRI study with expert dancers. Cereb. Cortex 15, 1243–1249.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., Fazio, F., 1997. Brain activity during observation of actions. Influence of action content and subject's strategy. Brain 120 (Pt. 10), 1763.

Fadiga, L., Fogassi, L., Pavesi, G., Rizzolatti, G., 1995. Motor facilitation during action observation: a magnetic stimulation study. J. Neurophysiol. 73, 2608–2611.

Fadiga, L., Craighero, L., Buccino, G., Rizzolatti, G., 2002. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. Eur. J. Neurosci. 15, 399–402.

Ferrari, P.F., Rozzi, S., Fogassi, L., 2005. Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. J. Cogn. Neurosci. 17, 212–226.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., Rizzolatti, G., 2005. Parietal lobe: from action organization to intention understanding. Science 308, 662–667.

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. Brain 119 (Pt. 2), 593–609.

Gallese, V., Keysers, C., Rizzolatti, G., 2004. A unifying view of the basis of social cognition. Trends Cogn. Sci. 8, 396–403.

Gazzola, V., Wicker, B., Gallese, V., Rizzolatti, G., Keysers, C., 2004. The anthropomorphic brain. In: Neuroscience, J.o.C. (Ed.), Annual meeting of the cognitive neuroscience society. MIT Press, San Francisco.

Gazzola, V., Aziz-Zadeh, L., Keysers, C., 2006. Empathy and the somatotopic auditory mirror system in human. Curr. Biol. 16, 1824–1829.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., Keysers, C., submitted for publication. Overcoming differences in embodiment: limb-aplasia and the mirror system.

Grafton, S.T., Arbib, M.A., Fadiga, L., Rizzolatti, G., 1996. Localization of grasp representations in humans by positron emission tomography: 2. Observation compared with imagination. Exp. Brain Res. 112, 103–111.

Grezes, 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, 928–937.

Hamilton, A.F., Grafton, S.T., 2006. Goal representation in human anterior intraparietal sulcus. J. Neurosci. 26, 1133–1137.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., Rizzolatti, G., 1998. Activation of human primary motor cortex during action observation: a neuromagnetic study. Proc. Natl. Acad. Sci. U. S. A. 95, 15061–15065.

Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., Rizzolatti, G., 1999. Cortical mechanisms of human imitation. Science 286, 2526–2528.

Iacoboni, M., Koski, L.M., Brass, M., Bekkering, H., Woods, R.P., Dubeau, M.C., Mazziotta, J.C., Rizzolatti, G., 2001. Reafferent copies of imitated actions in the right superior temporal cortex. Proc. Natl. Acad. Sci. U. S. A. 98, 13995–13999.

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 Biol. 3, e79.

Keysers, C., Gazzola, V., 2006. Towards a unifying neural theory of social cognition. Prog. Brain Res. 156, 383–406.

Keysers, C., Kohler, E., Umilta, M.A., Nanetti, L., Fogassi, L., Gallese, V., 2003. Audiovisual mirror neurons and action recognition. Exp. Brain Res. 153, 628–636.

Kilner, J.M., Paulignan, Y., Blakemore, S.J., 2003. An interference effect of observed biological movement on action. Curr. Biol. 13, 522–525.

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., Rizzolatti, G., 2002. Hearing sounds, understanding actions: action representation in mirror neurons. Science 297, 846–848.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., Rizzolatti, G., 1997. Object representation in the ventral premotor cortex (area F5) of the monkey. J. Neurophysiol. 78, 2226–2230.

Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., Orban, G.A., 2005. Observing others: multiple action representation in the frontal lobe. Science 310, 332–336.

Nishitani, N., Hari, R., 2000. Temporal dynamics of cortical representation for action. Proc. Natl. Acad. Sci. U. S. A. 97 (2), 913–918.

Oberman, L.M., McCleery, J.P., Ramachandran, V.S., Pineda, J.A., 2007. EEG evidence for mirror neuron activity during the observation of human and robot actions: toward an analysis of the human qualities of interactive robots. Neurocomputing. doi:10.1016/j.neucom.2006.02.024.

Press, C., Bird, G., Flach, R., Heyes, C., 2005. Robotic movement elicits automatic imitation. Brain Res. Cogn. Brain Res. 25, 632–640.

Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. Annu. Rev. Neurosci. 27, 169–192.

Rizzolatti, G., Fadiga, L., 1998. Grasping objects and grasping action meanings: the dual role of monkey rostroventral premotor cortex (area F5). Novartis Found Symp. 218, 81–95 (discussion 95-103).

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., Fazio, F., 1996. Localization of grasp representations in humans by PET: 1. Observation versus execution. Exp. Brain Res. 111, 246–252.

Shmuelof, L., Zohary, E., 2005. Dissociation between ventral and dorsal fMRI activation during object and action recognition. Neuron 47, 457–470.

Tai, Y.F., Scherfler, C., Brooks, D.J., Sawamoto, N., Castiello, U., 2004. The human premotor cortex is 'mirror' only for biological actions. Curr. Biol. 14, 117–120.

Umilta, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., Rizzolatti, G., 2001. I know what you are doing. A neurophysiological study. Neuron 31, 155–165.