World Scientific
www.worldscientific.com

HUMAN–HUMANOID INTERACTION: IS A HUMANOID ROBOT PERCEIVED AS A HUMAN?

ERHAN OZTOP[*,†,‡], DAVID W. FRANKLIN[†,§], THIERRY CHAMINADE[†,¶]
and GORDON CHENG[*,†,‖]

[*]JST-ICORP Computational Brain Project
and
[†]ATR Computational Neuroscience Laboratory,
2-2-2 Keihanna Science City,
Soraku-gun, Kyoto, 619-0288, Japan
[‡]erhan@atr.jp
[§]dfrank@atr.jp
[¶]tchamina@atr.jp
[‖]gordon@atr.jp

As humanoid robots become more commonplace in our society, it is important to understand the relation between humans and humanoid robots. In human face-to-face interaction, the observation of another individual performing an action facilitates the execution of a similar action, and interferes with the execution of a different action. This phenomenon has been explained by the existence of shared neural mechanisms for the execution and perception of actions, which would be automatically engaged by the perception of another individual's action. In one interference experiment, null interference was reported when subjects observed a robotic arm perform the incongruent task, suggesting that this effect may be specific to interacting with other humans. This experimental paradigm, designed to investigate motor interference in human interactions, was adapted to investigate how similar the implicit perception of a humanoid robot is to a human agent. Subjects performed rhythmic arm movements while observing either a human agent or humanoid robot performing either congruent or incongruent movements. The variance of the executed movements was used as a measure of the amount of interference in the movements. Both the human and humanoid agents produced a significant interference effect. These results suggest that observing the action of humanoid robots and human agents may rely on similar perceptual processes. Our findings suggest that experimental paradigms adopted from cognitive psychology can be used to derive measures for quantifying the degree of the implicit perception of a robot as human.

Keywords: Social robotics; perception of action; motor interference; uncanny valley.

1. Introduction

Recent advances in humanoid robotics[1–3] have brought certain social issues into consideration. It is generally accepted that humanoids will become part of our

daily lives, as have computers, the internet and email. Therefore, the study of how humanoids are perceived by human peers becomes an important issue if we expect them to interact with humans in a natural fashion.[4] There are some studies addressing the question of what kind of form[5] and functionality[6–8] a human-like robot should have in order to be socially accepted. The former relies on the introspective judgments of humans to decide how human-like robots are perceived. The latter approach relies on the implicit assumption that if the robot has cognitive mechanisms similar to humans' (e.g. gaze following) then they must be readily accepted as human-like, and mainly focuses on building such robotic systems.

Here, we propose a different approach, to test whether a humanoid robot could be treated as a human by the *perceptual system* of a human observer. We deliberately use the term "perceptual system of human" rather than simply "human." This is because our paradigm uses the *implicit* behavioral *effect* caused by the observation of others' behavior to investigate how the humanoid robot is perceived. In this way, we avoid invoking higher level cognitive systems that are involved in answering introspective questions (e.g. "Do you think that this is a human-like movement?" "Is this movement natural?"). In order to determine a suitable task that would enable us to tap into the implicit processing of human subjects' perception of humanoid robots, we first need to review current knowledge about perception of actions in humans.

1.1. *Experimental psychology: Motor resonance*

A number of recent studies have demonstrated that the perception of another's action and the actions executed by the self are mediated by common brain areas. Experimental psychology experiments have demonstrated that perceiving a simple action, such as a finger tapping,[9] grasping,[10] or arm movements[11] facilitates the concomitant execution of the same action and curbs the execution of a different one (see Ref. 12 for a review). This implies that observing the actions of other individuals and executing actions are not entirely distinct processes. These results led to the hypothesis that some cognitive representations are used both in the observation and in the execution of actions. This phenomenon was called motor resonance, and is believed to explain a number of social behaviors such as contagion of behaviors and synchronicity within a group.[13]

An experimental paradigm was recently introduced to investigate sensory-motor interference, which derives from motor resonance, in face-to-face behaviors, thus reproducing a socially valid interaction.[11] It taps into implicit perceptual processing,[12] which therefore can be used to assess without introspection how human-like a robot is perceived by humans. In this paradigm, subjects and the experimenter were face-to-face, and both produced arm movements that were either horizontal or vertical. Experimental conditions were defined by the congruency between the two movements. Variance in the movement was increased when subjects observe incongruent movements of human demonstration. However, the effect

that has been found in this study was specific to watching human actions, and could not be reproduced by the observation of robot actions. We were surprised that the results were negative, i.e. the robot movement did not have a similar effect as the human movement. We hypothesized that the results could be due to the robot's form (i.e. not a humanoid) and movement generation capability (i.e. not biologically realistic motion). Thus, we recreated the experimental setup using human subjects and our humanoid robot that could perform human like movements. Our experiment, when contrasted with Kilner *et al.*'s[11] can uncover whether this interference is due to the knowledge of the human nature of the other agent, or due to an automatic processing of certain features of the stimuli — (human-like) form and (biological) motion.

1.2. *Neurophysiology: Mirror systems*

In macaque monkeys, 'mirror neurons' were found in reciprocally connected ventral premotor and parietal cortices using electrophysiology. These neurons are activated both when monkeys perform a goal-directed action and when they see (or infer) the same action performed by an experimenter,[14−16] or hear the auditory result of an action.[17] Human functional neuroanatomy studies identified brain regions responsible for motor resonance.[18−21] These studies consistently show that the premotor and parietal cortices are activated during observation of actions.[22] Therefore, both studies in human and non-human primates point to the premotor and parietal cortices as substrates of motor resonance.

In addition, ventral premotor and parietal cortices are involved in imitation: in the left hemisphere, the premotor cortex appears to be more involved in the goal related aspects of the action[23] whereas the parietal cortex is more involved in body movement.[24] Finally, one study showed that cortices involved in producing a specific action, pointing or writing, are specifically recruited when understanding the goal of another individual's pointing or writing actions.[25]

1.3. *Computational approaches to motor resonance*

Movement imitation requires various complex mechanisms for mapping an observed movement onto one's own motor planning. From a robotics point of view, imitation requires the solution of several hard sub-problems including action recognition, pose estimation and tracking, body correspondence, coordinate transformation from external to egocentric space, matching of observed movement to a previously learned one, suitable movement representations for imitation, modular motor control.[26] Although the exact nature of biological motor primitives is not known, in robotic imitation usually hand-tuned (perceptuo-)motor primitives are used as a means of shared representations, which form a basis set of motion serving as a vocabulary for imitation and perception of actions.

More recently, models that may address motor interference have been proposed for motor control, imitation and action understanding. The MOSAIC model is a

control architecture composed of multiple forward-inverse model pairs that can be used for both motor control and action recognition.[27–29] Each pair competes to contribute to overall control. The inverse models with better predicting forward models become more influential in the control. During action observation and imitation, forward model predictions let the system dynamically choose the best controller to match an observed action. Therefore, motor interference can be explained by postulating that observing an action injects bias to the control affecting the performance by increasing the influence of modules controlling congruent movements, and decreasing the influence of modules controlling incongruent movements. The Mental State Inference model[30] posits that the common mechanism that is shared by both execution and observation is a forward prediction system. The forward model in this model is proposed to cancel out delays during action execution. During action observation, it is used in a mental simulation loop for inferring the intentions encompassed in the observed action. A possible explanation of the motor interference within this model is that observing an action (unconsciously) activates mental simulation mechanisms, causing a crosstalk between the ongoing execution and the observed action.

2. Hypotheses

2.1. *Interference between perceiving and executing actions*

Under the hypothesis that there is a shared neural substrate between perceiving and executing actions, it is predicted that observing an action should facilitate the execution of the same action, and curb the execution of a different action; the aim of the present experiment is to design a paradigm using this prediction to investigate how robots are perceived by humans. Results reported in the literature[11] are equivocal. Though a strong interference effect is found when subjects perform an arm movement incongruent to the one they are producing (e.g. vertical versus horizontal) defined by the increase of variance of the movement, no facilitation effect is associated with congruent movements. We reproduced this experiment with various changes to the paradigm in order to test its validity in the investigation of interaction between humans and robots. If increased variance in the movements is produced by the incongruency between the executed and observed movements, we will confirm that the present paradigm is suited to investigate the shared substrate between perception and action, and evaluate the competence of the interaction between humans and other agents (e.g. robots).

2.2. *Influence of the interacting agent on the interference*

Interestingly, Kilner *et al.*'s study did not find any interference effect using a robot. They concluded that the visuomotor interference is specific to interactions between humans.[11] We aimed at testing this claim with two improvements: we used a humanoid robot, that is, its general shape is similar to the shape of a human; and

the robot reproduced the movements of one of the experimenters. The form and motion of the robot and experimenter are similar. Therefore, if human–humanoid interactions produce comparable interference effects to human–human interactions, we can conclude that the effect is explained by an implicit perceptual mechanism sensitive to the general humanoid form and biological motion of the other agent. In this case, the interference effect can be used to evaluate the quality of the human–humanoid interaction. In contrast, if the interference effect is absent when subjects interact with the humanoid robot then this effect must be explained by the contextual knowledge of the human nature of the agent.

2.3. *Influence of the task on the interference*

In the original study,[11] subjects were instructed to track the finger of the experimenter or the tip of the robotic arm. Could the fact of tracking the finger create some interactions between motor control of eye and hand movements? To ensure that the interference did not result from motor/motor interaction, we designed two conditions: in the tracking conditions, subjects were instructed to follow the other agent's behavior. In a fixation condition, subjects were required to fixate a dot located on the forehead of the experimenter and on the part of the humanoid robot representing the head.

2.4. *Influence of the orientation of arm movements on the interference*

In the original study, the movements were specified in the vertical and horizontal directions. We thought the vertical and horizontal movements are not subject to exactly similar conditions due to gravity: in the vertical movements, gravity is parallel to the axis of the movement whereas in the horizontal movements, it is oriented perpendicular to the axis of movement. This could provide additional sensory information to the subjects about the accuracy of their movements, which may limit their variability. In order to reduce this possible effect, we reproduced the experiment with movements that were rotated by 45 degrees to the left and right, where similar forces due to gravity are present during both movements.

3. Methods

3.1. *The humanoid robot*

We used the humanoid robot "DB"[3] to produce diagonal reaching movements. DB is a hydraulic anthropomorphic robot with legs, arms (with hands without fingers), a jointed torso, and a head (Fig. 1).

DB was designed by the Sarcos company for the Kawato Dynamic Brain Project, and built by Sarcos to be 1.85 m tall with a human-like appearance. The robot contains 25 linear hydraulic actuators and five rotary hydraulic actuators having
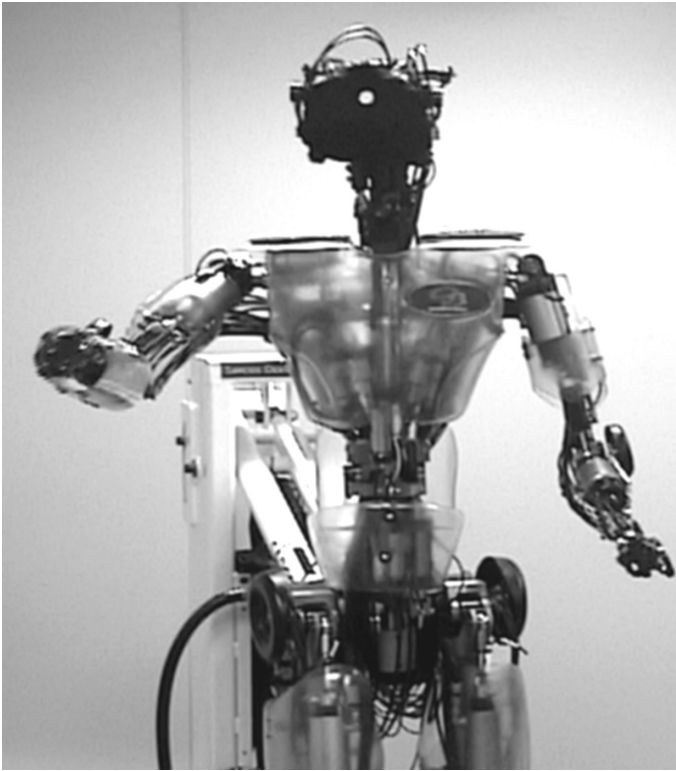
Fig. 1. DB, the humanoid robot used in our experiment.

30 degrees of freedom: three in the neck, two in each eye, seven in each arm, three in each leg, and three in the trunk (Fig. 1).

Our interest was to use the robot's right arm to produce diagonal movements. Thus, we commanded only the right arm and the torso joints to generate the movement. The robot was mounted from the back eliminating the need to deal with balancing issues. The task of the robot was to track the end point Cartesian trajectories captured from human motion (see next section for the data collection details) which were periodic top-left to bottom-right (L) and top-right to bottom-left (R) reaching movements involving elbow, shoulder and some torso movements.

The controller was implemented on the real-time operating system VxWorks using several parallel Motorola PowerPC processors in a VME rack within the environment provided by SL Simulation and Real-Time Control Software Package (http://www-clmc.usc.edu/publications/S/schaal-TRSL.pdf). To produce trajectories for the robot, we first extracted the main direction of human trajectories using PCA and applied a linear transformation such that the points in the trajectory lie within the workspace of the robot as much as possible. Not all the points were reachable, so the extent of the movements were scaled by 0.95 in each direction

of the robot coordinate frame, allowing the robot to stay within a predefined task space most of the time.

A velocity-based Cartesian controller was used to track the transformed trajectories shown in Fig. 2(b). The control scheme used is shown in Fig. 3, where the inverse kinematics is solved using Jacobian pseudo-inverse with null space optimization.[31] Tracking achieved was satisfactory for our purposes as the robot movements were smooth and human-like for a human observer. Figure 4 shows the tracking in x-, y- and z-coordinates for the two trajectories used in the experiments where the z-axis is parallel to the gravity direction pointing up, and the y-axis is parallel to

(a)

Top-Right to Bottom Left   Top-Left to Bottom Right
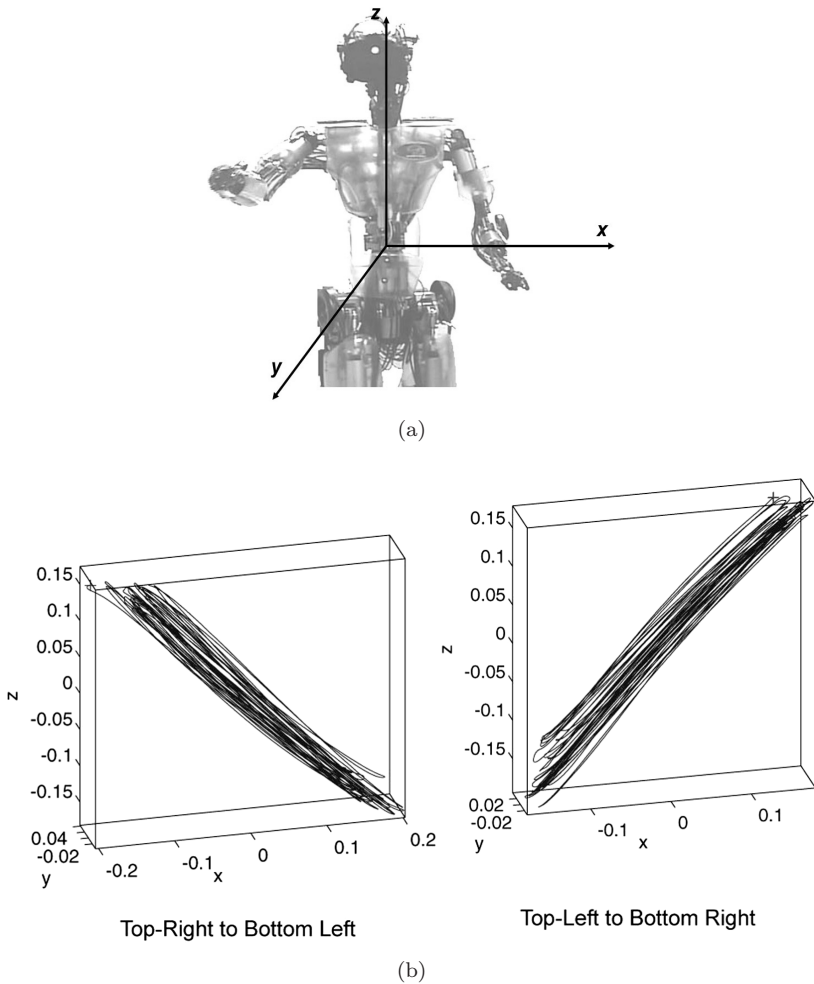
(b)

Fig. 2. (a) The coordinate system of the endeffector position is illustrated (arrows point towards the positive directions). (b) The transformed human movement trajectories for each of the two movements studied.
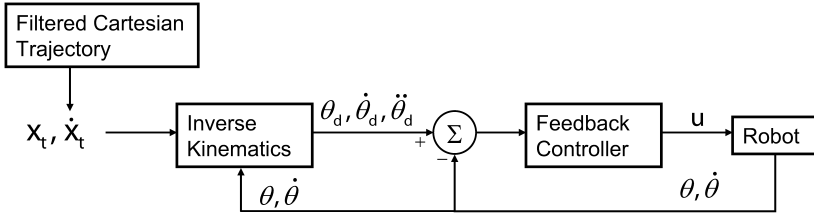
Fig. 3. Cartesian tracker used in generating the arm movements of the humanoid robot.

the floor pointing away from the robot, and finally the x-axis extends laterally, pointing towards the left side of the robot. Note that the plane of interest was the plane spanned by x- and z-axes.

3.2. *Experimental paradigm*

Eight naïve subjects (2 left-handed; 6 right-handed) (6 males; 2 females) ranging from 20–35 years of age performed rhythmic arm motions with their right arms in front of their bodies while standing. Subjects were instructed to make rhythmic arm movements across their bodies from either the top-right to the bottom-left (R) or from the top-left to the bottom-right (L) with respect to their own torso at 0.5 Hz. Subjects performed these movements while standing approximately 2 m away from either a humanoid robot or another human that was performing similar arm movements (Fig. 5). During each trial, the subjects were instructed to produce one of the two movements (R) or (L). The movements produced by the two agents were either spatially congruent (C; same direction) or incongruent (I; orthogonal to the subject's motion). The subjects were also instructed to be in phase with the other agent's movements. During each task, the subject was either instructed to track the other agents' hand movements with their eyes (T) or to fixate the view of their eyes on a marked point located on the head of the agent (F). Each trial was initiated by the robot or human agent. During each 30 second trial, the kinematics of the endpoint of the subject's right index finger was recorded at 250 Hz using an OPTOTRAK 3020 (Northern Digital, Waterloo). The experiments with the humanoid robot were performed on a separate day from those with the human experimenters.

The experimenters made movements at 0.5 Hz while listening to a 1 Hz beep on headphones. They kept their eyes closed to avoid reciprocal interference from the observation of the subjects' movements. Each subject performed movements in front of the robot and two experimenters, where the order of experimenter presentation was randomized across subjects. A session of the experiment, containing eight conditions (R/L) × (C/I) × (T/F) was performed in a random order of presentation with the first experimenter, and in a reversed order with the second experimenter. The robot's arm movements were based on data recorded from one of the experimenters which were collected in the same way as for the subjects and were digitally
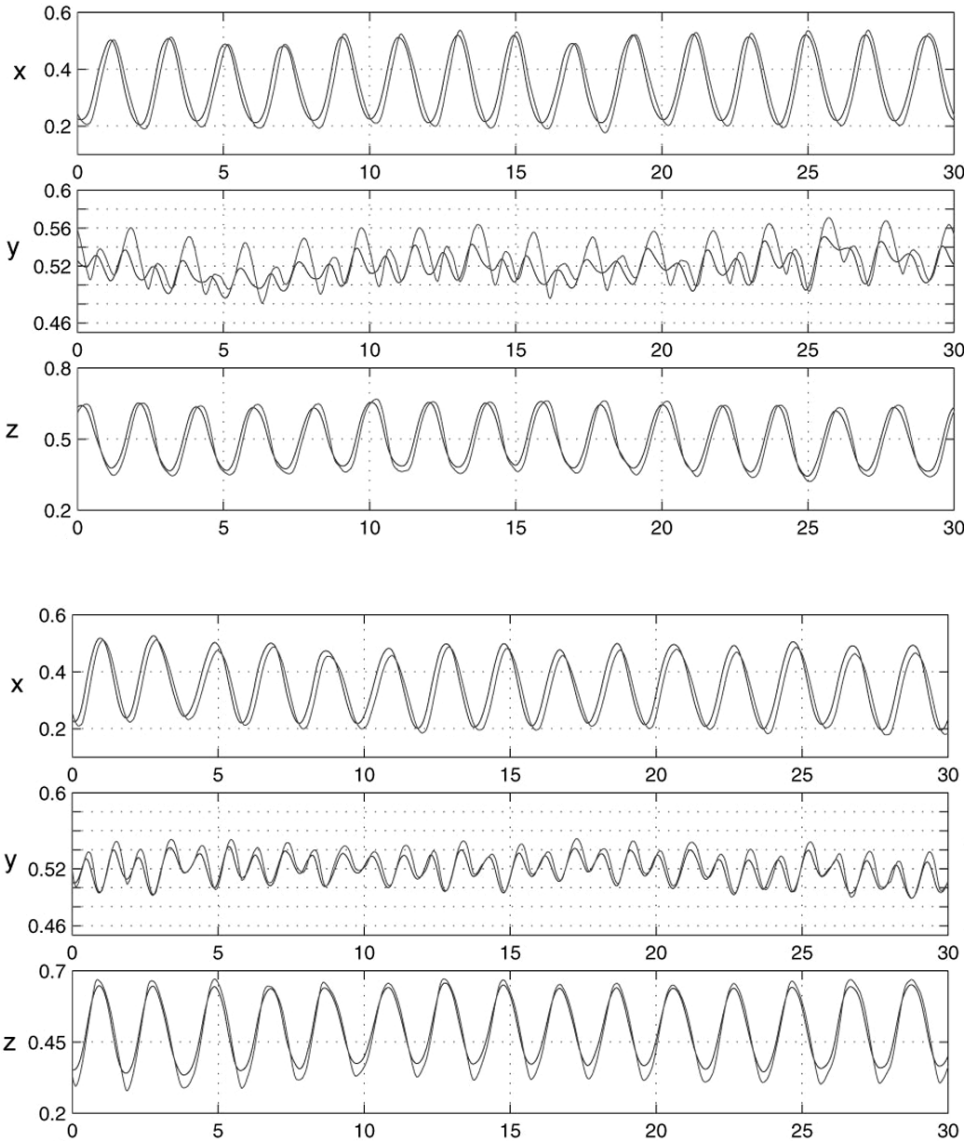
Fig. 4. Bottom: top-right to bottom-left movement tracking (R). Top: top-left to bottom-right tracking (L); Horizontal axis is the time, the vertical axes are x, y, and z of the robot coordinate frame (z pointing up, x pointing right, y pointing away in front of the robot; the origin is at the waist).
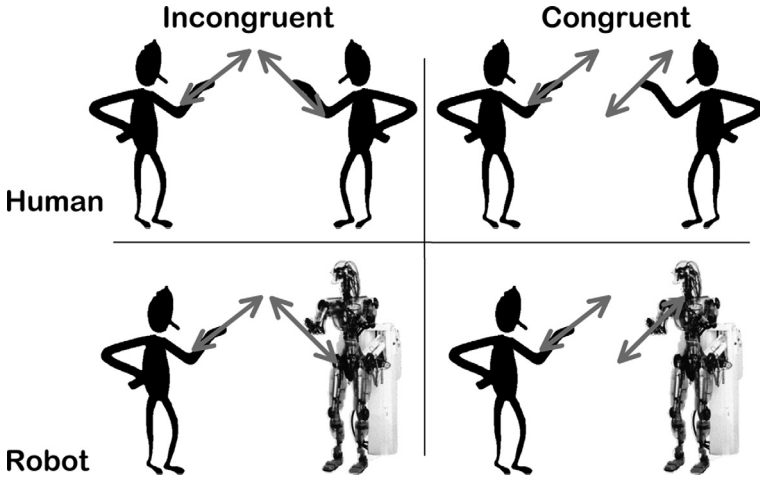
Fig. 5. The Experimental paradigm: subjects performed diagonal rhythmic movements starting either from top-left (L) or top-right (R) while tracking the hand of either a human agent or a humanoid robot performing either congruent or incongruent movements.

low-pass filtered using a zero-lag, fifth-order Butterworth filter with a cut-off frequency of 25 Hz. The humanoid robot has a pair of cameras mounted at the top of the head that could be perceived as human eyes. It is not possible to close them to reproduce the closed eyes of the experimenter in the human–human interaction conditions. In order to avoid a confounding factor, a black cloth was used to cover the robot's cameras.

3.3. *Analysis*

The three-dimensional trajectories of each marker were digitally low-pass filtered using a zero-lag, fifth-order Butterworth filter with a cut-off frequency of 25 Hz. Each movement was segmented from the surrounding movements using the 3D curvature $(c(t))$[32]:

$$c\left(t\right) = \frac{\sqrt{\|\dot{\mathbf{x}}(t)\|^2 \, \|\ddot{\mathbf{x}}(t)\|^2 - (\dot{\mathbf{x}}(t)^T \ddot{\mathbf{x}}(t))^2}}{\|\dot{\mathbf{x}}(t)\|^3}, \tag{1}$$

where $\mathbf{x}$ is a vector of the three-dimensional position of the hand. This measure of curvature is very low when the velocity of the movement is high, but becomes very high as velocity slows and the movement changes direction. By using the measure of curvature, each upward or downward movement was segmented from the surrounding movements for further analysis. The beginning and end of each movement was removed if the curvature was above 100 [1/m], preventing small drifts in the hand location at the extremes of movement from influencing the results.

For analysis, the three-dimensional kinematic data of the finger was projected onto two orthogonal planes: vertical and horizontal. The vertical plane was defined as the plane spanned by the x-and z-axes, whereas the horizontal plane was defined as the plane spanned by the x-and y-axes. The previous work on human–human interaction[11] projected the data only upon the vertical plane. As the vertical plane is the plane in which the movements are either congruent or incongruent, it is possible that variability of the movements is only affected in this plane, in contrast to the orthogonal horizontal plane. However, as the movements, and therefore their variability, are actually produced by muscles (whose forces are not constrained to a particular plane), it is possible that this variability would also be found in the horizontal plane. By projecting the data onto both planes, the specificity of the plane of movement affected by the experimental manipulation can be examined.

The previous work found significantly increased variability in the subjects' movements in the incongruent condition. In order to quantify the effects of the conditions on the behavior of the subjects, we calculated the signed area of each movement defined as the deviation from the straight-line joining the start and end of each segmented movement.[33] This area was calculated separately in the vertical and horizontal planes. This measure gives an estimate of the curvature of each trial. To estimate the variability of the behavior within each 30-second trial, the variance of the signed area of movement segments was calculated. This gives an estimate of the amount by which this curvature changes from one trial to the next during the 30-second trial. In order to normalize the data to compensate for possible confounding effects due to movement length, the variance was divided by the mean absolute signed area during this trial. In particular, the measure of the variance V for each condition was estimated as

$$V_{\text{cond}} = \frac{\text{var}(e_{tr=1:n})}{\text{mean}\,|e_{tr=1:n}|}, \tag{2}$$

where e is the signed area for all trials in that condition.

The variability in movements could also exist in terms of the length of movements. In order to test this, a similar quantification was also performed on movement length. Movement length was calculated as the straight line segment joining the start and end points of each movement as determining from the curvature. Similar to the previous computation [Eq. (2)], the variance of the movement lengths in each condition was divided by the mean movement length in the condition.

The statistical significance of the movement variability was examined using an analysis of variance (ANOVA) with main effects of agent (H/R), laterality of movement (R/L), motion congruency (C/I), eye motion (F/T) and experimenter, and a random effect of subjects. This was performed on the variability of path trajectory in the vertical and horizontal planes (ver-V and hor-V, respectively), and on the variability of the movement length (len-V). The ANOVA considered only first and second level interaction effects, significant at the alpha 0.05 level. A t-test was used to directly compare variability in the congruent and incongruent movements

under particular conditions. An ANOVA is a statistical model which determines what portion of the observed variance of the data can be attributed to each factor in the model. By using this statistical test, it is possible to determine whether each factor, for example, motion congruency, can explain some portion of the variance in the overall data. If our example factor of motion congruency has a probability value less than 0.05, then we can conclude that motion congruency significantly affects the motion studied independent of any other factor. Interaction effects describe how two or more factors influence the variance in the data in a non-additive sense. A full treatment of the ANOVA and its interpretations can be found in most introductory statistics textbooks.

4. Results

The subjects' movements were examined using three variability measures: the normalized variance of the movement lengths (len-V), the normalized variance of the areas projected on the vertical plane (ver-V), and the normalized variance of the areas projected on the horizontal plane (hor-V).

4.1. *Variability based on movement length (len-V)*

The results of the ANOVA on the variability of movement length showed no main effect of congruency ($p = 0.7345$) or agent ($p = 0.1882$) (Table 1). The only significant effect of interest was that of laterality, which indicates that subjects tended to make shorter movements in one of the two directions of movement. This means that there was no increased variability in the movement length produced by either the congruency of the movements or by a particular agent. This indicates that neither congruency nor agent type increased variability in the main axis of the movement.

In order to further confirm that there is no effect on movement length variability due to congruency or incongruency of movements, a *t*-test was performed. There were no significant differences between congruent and incongruent movements for either the human or the robot agents under either of the eye conditions (tracking or fixating). The results are shown in Fig. 6. Clearly, no increased variance was found in the movement lengths. Therefore, if increased variance is found in either the vertical or horizontal planes, this effect is not produced by changes in the movement lengths. It must instead be produced by increased variability in the trajectory of the movements.

4.2. *Variability based on signed area on the vertical plane (ver-V)*

In the vertical plane, the ANOVA on the variability of the behavior (variance of the signed area for each trial) showed a significant main effect of congruency ($p = 0.015$) (Table 2). Incongruent movements exhibited higher variability than those of congruent movements (see Fig. 7). This indicates that when subjects faced an agent (either

Table 1. Results from the ANOVA examining the variability of the movement length (len-V). The "Effect" column labeling is as follow: Agent → agt; Laterality → lat; Congurency → cng; Tracking-Fixation → eye; Experimenter → exp. The products (*) represent the interaction effects; for example, agt*cgn indicates the interaction effect between Agent and Congruency.

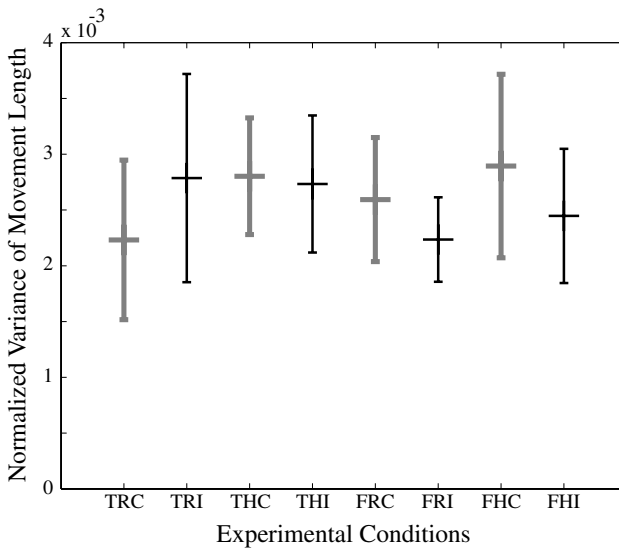| Effect	d.f.	F	Prob > F
agt	1	1.7440	0.1882
lat	1	5.3647	0.0216
cng	1	0.1153	0.7345
eye	1	0.1749	0.6763
exp	1	2.4952	0.1158
agt*lat	1	0.0242	0.8765
agt*cng	1	0.8645	0.3536
agt*eye	1	0.0037	0.9515
agt*exp	1	2.1311	0.1459
lat*cng	1	0.0025	0.9604
lat*eye	1	0.4298	0.5128
lat*exp	1	0.4639	0.4966
cng*eye	1	2.6894	0.1026
cng*exp	1	0.0745	0.7852
eye*exp	1	8.6719	0.0036
Error	**197**		
Total	**254**		

Fig. 6. No effect of congruency on the variability of the movement length (len-V). The mean and standard error of the mean (SEM) of the variability in moment length for all subjects across the conditions. The congruent conditions are shown in gray, whereas the incongruent conditions are shown in black. The conditions are defined by the letters (T: eye track, F: eye fixate; R: robot, H: human; C: congruent movement, I: incongruent movement). No significant effects were found using *t*-tests.

Table 2. Results from the ANOVA examining the variability of the signed area of the movement projected onto the verticle plane (ver-V). The conventions used are the same as in Table 1.

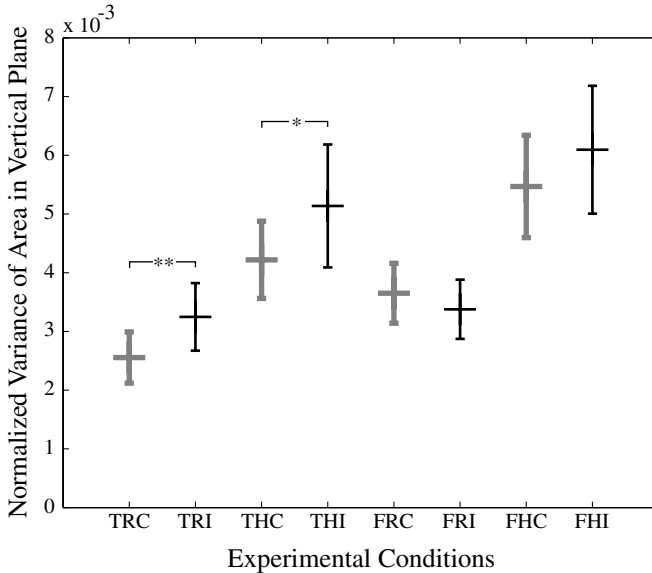| Effect	d.f.	F	Prob $> F$
agt	1	98.4891	0.0000
lat	1	6.0857	0.0145
cng	1	5.9823	0.0153
eye	1	17.98	0.0000
exp	1	0.7163	0.3984
agt*lat	1	4.078	0.0448
agt*cng	1	1.7394	0.1887
agt*eye	1	1.3116	0.2535
agt*exp	1	1.9457	0.1646
lat*cng	1	0.183	0.6693
lat*eye	1	2.2338	0.1366
lat*exp	1	1.4557	0.2291
cng*eye	1	2.5141	0.1144
cng*exp	1	1.79	0.1825
eye*exp	1	1.5425	0.2157
Error	**197**		
Total	**254**		

Fig. 7. Congruency has a significant effect on the variability in the vertical plane (ver-V). The mean and SEM of the variability of the trajectory in the vertical plane for all subjects across the conditions. The conditions and colors are shown exactly as in Fig. 6. Significant effects from a t-test are shown ($** : p < 0.05; * : p < 0.1$).

human or humanoid robot) performing incongruent movements, the variability of their movements was increased. While there was also a significant main effect for agent, there was no interaction effect of agent versus congruency ($p = 0.18$). This illustrates that the same effect of congruency was seen for both interactions with human agents and with the humanoid robot. The main effect of eye was significant, indicating that fixation versus tracking exhibited different behaviors.

The effects of interest were further investigated by performing t-tests contrasting the congruent and incongruent movements for the human and robot agents (Fig. 7). Under the tracking condition, both the human and robot observation caused significantly higher variability for incongruent movement than congruent movements ($p < 0.1$). This further confirms that the humanoid robot had the same type of motor interference effect as did the human which had previously not been seen for an industrial robot.[11] No significant effects were seen for either the human or robot agent under the fixation condition, and the variability of the movements was higher in the fixation conditions compared to corresponding tracking conditions.

4.3. *Variability based on signed area on the horizontal plane (hor-V)*

In the horizontal plane, the ANOVA on the variability of the behavior (variance of the signed area for each trial) did not have a significant main effect of congruency ($p = 0.065$) although the value was borderline (Table 3). This is not surprising since the muscles involved in the manifestation of the interference in ver-V also

Table 3. Results from the ANOVA examining the variability of the signed area of the movement projected onto the horizontal plane (hor-V). The conventions used are the same as in Table 1.

OUTSource	d.f.	F	Prob $> F$
agt	1	41.3173	0.0000
lat	1	10.0596	0.0018
cng	1	3.4483	0.0648
eye	1	4.9646	0.0270
exp	1	1.4236	0.2342
agt*lat	1	0.2038	0.6522
agt*cng	1	0.0192	0.8900
agt*eye	1	0.463	0.4970
agt*exp	1	0.65	0.4211
lat*cng	1	0.7589	0.3847
lat*eye	1	0.0181	0.8931
lat*exp	1	0.0215	0.8835
cng*eye	1	1.6432	0.2014
cng*exp	1	0.6498	0.4212
eye*exp	1	8.584	0.0038
Error	**197**		
Total	**254**		

contribute to the production of the motion in the horizontal plane. Presumably, with more subjects the variability under these conditions may have become significant. However, we can conclude that the effect of incongruent movements being preformed in front of the subjects is stronger in the plane in which this incongruency exists (i.e. the vertical plane). Again, while there was also a significant main effect for agent, there was no interaction effect of agent versus congruency ($p = 0.89$). This illustrates that the absence of the main effect of congruency is not due to different effects for the human or the robot agents. The main effect of eye was also significant, indicating that fixation versus tracking exhibited different behaviors.

These effects were further investigated by performing t-tests contrasting the congruent and incongruent movements for the human and robot agents (Fig. 8). No significant differences were found with the t-test, although in general the data exhibited the same trends as found in the vertical plane. Clearly, there was no difference in the relative behaviors for movements while viewing either the human and robot agents making movements. Under the tracking condition, both the human and robot data have slightly larger variability in incongruent movements compared to congruent movements. In contrast to the vertical plane, the effect was only significant when subjects interacted with the human, although the p-value for the robot was close to significant at the level of 0.1 ($p = 0.17$). No significant effects were
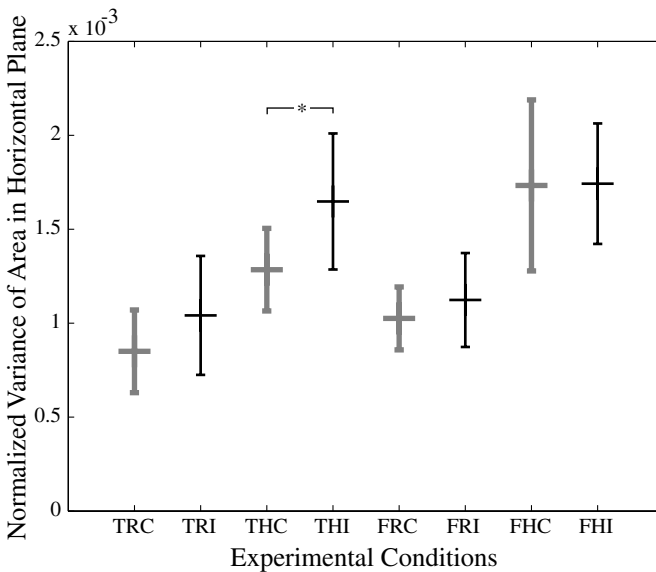

Fig. 8. Congruency has no significant effect on the variability in the horizontal plane. The mean and SEM of the variability of the trajectory in the horizontal plane for all subjects across the conditions. The conditions and colors are shown exactly as in Fig. 6. The overall trend of the data is similar to that in the vertical plane (Fig. 7), but the only significant effect at $p < 0.1$ was found when subjects interacted with the human (*).

seen for either the human or robot agent under the fixation condition, although the values were increased from similar tasks under the fixation condition.

5. Discussion

5.1. *An objective tool to assess human perception of robots*

Robot designers currently rely on the assumption that anthropomorphic robots would be more competent in social interactions because humans would interact intuitively with them.[4] However, this assumption is widely unexplored, especially considering the conflicting hypothesis called the "uncanny valley" introduced by Masahiro Mori.[4,34] For this roboticist, the relation between human empathic reaction and anthropomorphism of the robot does not show a monotonic increasing curve, but presents a sharp chasm, indicating a strong negative reaction, before reaching the exact human-likeness. Robots imperfectly attempting to reproduce human behavior would cause a negative reaction leading to the rejection of the robot as an interacting partner.

Despite the consequences of the "uncanny valley" hypothesis on the design of interacting robots, it has not been investigated scientifically, perhaps due to a lack of an objective tool to test the human reaction. Instead, designers usually rely on intuition or surveys, which can suffer from subjective biases. It is thus desirable to create a paradigm in order to assess the human reaction to the perception of robots without conscious introspection.

We contend that the experimental framework and tools developed in behavioral sciences to investigate social interactions among humans[12] could be adapted to test human interactions with robots. One prominent hypothesis in cognitive neuroscience is the existence of a common framework for execution and perception of action. A paradigm developed to investigate this common mechanism measures how an observed action interferes with the production of an action. Within this framework, we adapted a motor interference study[11] that is well suited for examining full body interactions, and modified it: first, we used diagonal movements rather than vertical and horizontal movements to avoid the possible confounding factor of gravity. Second, we introduced a fixation condition that required the subjects' eyes not to move during the experiment.

5.2. *Validation of the interference effect*

For both fixation and tracking conditions, the interference was measured as the change in the variability of the subjects' arm movements while observing congruent and incongruent movements made by either a human agent or a humanoid robot. In the analysis, we considered three types of variability measures. First, two measures were based on the signed areas of the projected movements onto vertical (ver-V) and horizontal (hor-V) planes under the hypothesis that interference should be found on the vertical plane, on which congruency was controlled.

The third measure was based on the length of movements in three-dimensional space (len-V).

In the fixation condition, we did not find any significant effects of congruency for both human and robot observation. This was true for all the three types of movement variability measures. However, the eye movements clearly had an impact on the behavior of the subjects, as demonstrated by the significant effect of eye on the movement variability in the vertical and horizontal planes. As noted earlier, measures of variability in the fixation conditions generally show an increase when compared to the measures in the tracking conditions considered in the same tasks. This increase of variability in the fixation conditions indicates a reduction in control performance for the actions when subjects are required to perform a movement with their arm while fixating a certain point on the body of another agent performing movements. Thus, in the fixation conditions the effect of interference might have been shadowed by the poor motor control, explaining the absence of significant effect of congruency on the variability of movements.

However, in the tracking condition there was a significant effect of congruency on the subjects' variability of behavior both for human and robot observation when the movements were projected on the vertical (ver-V) but not on the horizontal plane (hor-V). The change in variability using vertical projection (ver-V) was positive in the incongruent condition for both agents (human and humanoid robot) (Fig. 7). In addition, by using length variance (len-V) as the movement variability, we showed that the effect of congruency was not due to the spatial extent of the executed movement as there was no significant effect of congruency on the subjects' variability for both agents.

For human–human interactions, our tracking condition results are highly similar to the previous report.[11] Therefore, our study strengthens the conclusion for human–human interaction since our experiments were designed to eliminate one possible confounding factor, namely, the differential effect of gravity on subjects' movements.

However, in Kilner *et al.*'s study,[11] no increase in variability was found when subjects were observing a robot performing the actions. In sharp contrast to the above-mentioned study, we observe a significant effect of the incongruency for human–robot interactions. Similar to human–human interactions, the observation of incongruent movements increases the variability of the subjects' behaviors. The discrepancy between our study and the earlier study can be explained by the differences between the humanoid robot used in our study, which has a human-like appearance and produced human-like movements contrasting with the industrial robot used in Kilner *et al.*'s study,[11] which consisted of metal shafts which moved linearly with constant velocity. In addition, the overall trend of increased variability in the interactions with humans compared to interactions with robots might indicate that interacting with a human may have a general effect of increasing the movement variability in addition to the increase that may be caused by the incongruency present in the movement observation.

Altogether, our results provide guidelines for future robot experiments to assess the human-likeness perception of robots, which can offer principles for robot design and control. First, we showed that the inference effect can be obtained with different trajectories (diagonal–diagonal versus vertical-horizontal). Thus, the experimental design can possibly be adapted to accommodate a humanoid robot's complex motions (e.g. the motion of the body and the arm of a piano playing robot). Second, tracking conditions should be preferred as the interference effect is limited to these conditions. Finally, interference is systematically found in the vertical plane, which therefore, should be preferred to measure the interference effect.

5.3. *Future work*

Further studies controlling the form of the robot, for instance, using more human-like robots (influence of the face appearance), and its motion, using a gradient of movements ranging from least human-like to human-like, are needed to develop quantitative measures for human perception of human-likeness of robots. An important issue is to answer the question: "What are the contributions of motion and shape to the human-likeness perception?" as these two characteristics were varied together by Kilner *et al.* and in the present experiment. Do motion and shape have their independent uncanny valleys, as originally proposed by Mori?[34] Intuition suggests that the answer to the latter question is negative. A perfect shape replica requires a perfect motion to go with to elicit the impression of human-likeness, whereas an animated character with an approximate human motion easily creates the feeling of "behaves like a human!" Nevertheless, as humanoid robots are still far from perfect replicas of humans, both in terms of overall shape and of their motor control, the question of the effect of the motion characteristics on perception becomes crucial, and can be stated as follows: considering a given robotic design, which type of motion elicits an interference effect to match the level of interference caused by human motion?

Take, for example, the humanoid robot used in the present experiment. Its design is clearly robotic, as noted by the absence of flesh, clear facial features, etc. If the responses to motion and shape are clearly independent then the interference should be different when linear motions, like the ones used by Kilner *et al.* with their industrial robot, are used for the humanoid robot instead of the more natural motions based on motion captured data used in the present experiment. On the other hand, it is possible that the interference effect relies on the expectancies we have when interacting with another agent. If we expect a human to move with a human motion, it is clear that when faced with a robot our expectations are different, in which case linear motion would not have a significant impact on the interference effect. Hopefully, this hypothesis can easily be tested with the paradigm presented here, and future experiments should aim at answering exactly these questions. Given a robot with a humanoid form or a human (e.g. a pantomime artist), what happens when the arm follows a "robotic" movement trajectory? Turning around the situation,

given a biology realistic arm motion, what would subjects' implicit reaction be when the form of the demonstrator is hidden?

6. Conclusion

The motor interference which is explained by the motor resonance hypothesis is not specific to human–human interactions but can also be observed in human–humanoid interactions. Our study, together with previous studies, suggests that a combination of the form and the motion of an agent is an important factor for the social competence of the interaction. Our next step should be to separate the relative contribution of form and motion to this interference. This study forms the first step in this direction as it not only validates an effective experimental paradigm to assess a human's implicit reaction to a humanoid robot but also provides a method to derive a quantity from the subject motion to measure the level of this reaction.

References

1. B. Adams, C. Breazeal, R. A. Brooks and B. Scassellati, Humanoid robots: A new kind of tool, *IEEE Intell. Syst.* **15**, 25–31 (2000).
2. G. Cheng, A. Nagakubo and Y. Kuniyoshi, Continuous humanoid interaction: An integrated perspective — gaining adaptivity, redundancy, flexibility — in one, *Robot. Autonom. Syst.* **37**, 161–183 (2001).
3. C. G. Atkeson, *et al.*, Using humanoid robots to study human behavior, *IEEE Intell. Syst.* **15**, 46–56 (2000).
4. T. Fong, I. Nourbakhsh and K. Dautenhahn, A survey of socially interactive robots, *Robot. Autonom. Syst.* **42**, 143–166 (2003).
5. C. DiSalvo, F. Gemperle, J. Forlizzi and S. Kiesler, All robots are not created equal: The design and perception of humanoid robot heads, in *Proc. Conf. Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, London, England, 25–28 June, 2002 (ACM, 2002).
6. B. Scassellati, Foundations for a theory of mind for a humanoid robot, Foundations for a theory of mind for a humanoid robot, Ph.D. thesis, Dept. Elec. Eng. Comp. Sci., MIT (2001).
7. H. Kozima and H. Yano, A robot that learns to communicate with human caregivers, in *Proc. Int. Wksp. Epigenetic Rob.*, Lund, Sweden, 17–18 September, 2001 (LUCS, Lund, 2001).
8. C. Breazeal and B. Scassellati, Infant-like social interactions between a robot and a human caretaker, *Caretaker Adaptive Behavior* **8**, 49–74 (2000).

9. M. Brass, H. Bekkering, A. Wohlschlager and W. Prinz, Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues, *Brain Cogn.* **44**, 124–143 (2000).

10. M. G. Edwards, G. W. Humphreys and U. Castiello, Motor facilitation following action observation: A behavioural study in prehensile action, *Brain Cogn.* **53**, 495–502 (2003).

11. J. M. Kilner, Y. Paulignan and S. J. Blakemore, An interference effect of observed biological movement on action, *Current Biol.* **13**, 522–525 (2003).

12. S. J. Blakemore, and C. Frith, The role of motor contagion in the prediction of action, *Neuropsycholology* **43**, 260–267 (2005).

13. T. L. Chartrand and J. A. Bargh, The chameleon effect: The perception-behavior link and social interaction, *J. Pers. Soc. Psychol.* **76**, 893–910 (1999).

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

15. G. Rizzolatti, L. Fadiga, V. Gallese and L. Fogassi, Premotor cortex and the recognition of motor actions, *Cogn. Brain Res.* **3**, 131–141 (1996).

16. M. A. Umilta, *et al.*, I know what you are doing: A neurophysiological study, *Neuron* **31**, 155–165 (2001).

17. C. Keysers, *et al.*, Audiovisual mirror neurons and action recognition, *Exp. Brain. Res.* **153**, 628–636 (2003).

18. M. Iacoboni, *et al.*, Grasping the intentions of others with one's own mirror neuron system, *PLoS Biol.* **3**, 79 (2005).

19. G. Buccino, *et al.*, Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study, *Eur. J. Neurosci.* **13**, 400–404 (2001).

20. R. Hari, *et al.*, Activation of human primary motor cortex during action observation: A neuromagnetic study, *Proc. Nat. Acad. Sci. USA* **95**, 15061–15065 (1998).

21. M. Iacoboni, *et al.*, Cortical mechanisms of human imitation, *Science* **286**, 2526–2528 (1999).

22. J. Grezes and J. Decety, Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis, *Hum. Brain. Mapp.* **12**, 1–19 (2001).

23. L. Fadiga and L. Craighero, New insights on sensorimotor integration: From hand action to speech perception, *Brain. Cogn.* **53**, 514–524 (2003).

24. T. Chaminade, A. N. Meltzoff and J. Decety, An fMRI study of imitation: Action representation and body schema, *Neuropsychology* **43**, 115–127 (2005).

25. T. Chaminade, D. Meary, J. P. Orliaguet and J. Decety, Is perceptual anticipation a motor simulation? A PET study, *Neuroreport* **12**, 3669–3674 (2001).

26. S. Schaal, A. Ijspeert and A. Billard, Computational approaches to motor learning by imitation, *Philos. Trans. R. Soc. Lond. B.* **358**, 537–547 (2003).

27. D. M. Wolpert and M. Kawato, Multiple paired forward and inverse models for motor control, *Neural Net.* **11**, 1317–1329 (1998).

28. M. Haruno, D. M. Wolpert and M. Kawato, MOSAIC model for sensorimotor learning and control, *Neural Comput.* **13**, 2201–2220 (2001).

29. D. M. Wolpert, K. Doya and M. Kawato, A unifying computational framework for motor control and social interaction, *Philos. Trans. R. Soc. Lond. B.* **358**, 593–602 (2003).

30. E. Oztop, D. Wolpert and M. Kawato, Mental state inference using visual control parameters, *Brain Res. Cogn. Brain Res.* **22**, 129–151 (2005).

31. L. Sciavicco and B. Siciliano, *Modelling and Control of Robot Manipulators* (Springer, London, 2000).

32. S. Schaal, and D. Sternad, Origins and violations of the 2/3 power law in rhythmic three-dimensional arm movements, *Exp. Brain Res.* **136**, 60–72 (2001).
33. D. W. Franklin, R. Osu, E. Burdet, M. Kawato and T. E. Milner, Adaptation to stable and unstable dynamics achieved by combined impedance control and inverse dynamics model, *J. Neurophysiol.* **90**, 3270–3282 (2003).
34. M. Mori, The valley of eeriness, *Energy* **7**, 33–35 (1970) [in Japanese].

Erhan Oztop received his M.S. degree from Middle East Technical University in 1996, and his Ph.D. degree from the University of Southern California in 2002. From 2002 to 2003, he was a visiting researcher at the Advanced Telecommunications Research Institute, Computational Neuroscience Laboratories, Kyoto, Japan. He is now a researcher at Japan Science Technology, working for the ICORP Computational Brain Project, and he is still affiliated with the Advanced Telecommunications Research Institute. His research interests include computational modeling of the brain mechanisms of action understanding and visuomotor transformations, human–robot interaction, robotics, machine learning and cognitive neuroscience.

David W. Franklin is a researcher in the Department of Cognitive Neuroscience, Advanced Telecommunications Research Institute, Computational Neuroscience Laboratories, Kyoto, Japan. He received his M.Sc. and Ph.D. degrees from Simon Fraser University, Canada in 2000 and 2004, respectively. He is a member of the Society for Neuroscience and the author of over 40 technical publications. His research interests include neural control of human movement, muscle mechanics, biological and machine learning, motor control, haptics and human–machine interaction.

Thierry Chaminade received his Ph.D. in neuropsychology at the University Louis Lumières Lyon 2, France in 2003. His thesis work investigated the neurophysiology of action observation and imitation using human neuroimaging techniques performed in the INSERM Unit 280, Lyon, France and the University of Washington's Institute for Mind, Brain and Learning, Seattle, USA. In 2004, he was a visiting researcher at the Advanced Telecommunications Research Institute, Kyoto (Japan), where he started investigating social interactions with artificial agents. He is now a postdoctoral fellow at the Functional Imaging Laboratory at University College London, UK and a continuing collaborator of the Computational Neuroscience Laboratory of the Advanced Telecommunications Research Institute. His research interests include the understanding of the behavioral and neuronal bases of social interactions.

Gordon Cheng is the head of the Department of Humanoid Robotics and Computational Neuroscience, ATR Computational Neuroscience Laboratories, Kyoto, Japan. He is also the Group Leader for the newly initiated JST International Cooperative Research Project (ICORP), Computational Brain. Before taking up these positions, he held fellowships from the Center of Excellence (COE), Science, and Technology Agency (STA) of Japan. Both fellowships were taken at the Humanoid Interaction Laboratory, Intelligent Systems Division at the ElectroTechnical Laboratory (ETL), Japan. At ETL, he played a major role in developing a completely integrated humanoid robotics system. He received his Ph.D. in Systems Engineering from the Department of Systems Engineering, Australian National University and his Bachelor and Master degrees in Computer Science from the University of Wollongong, Australia. His industrial experience includes consultancy to and as a national systems manager for a major transport company. He was also the director of the company G.T.I. Computing, specializing in networking/transport management systems in Australia. His research interests include humanoid robotics, biomimetic of human vision, computational neuroscience of vision, action understanding, human–robot interaction, active vision, mobile robot navigation and object-oriented software construction. He is a society member of the IEEE Robotics and Automation and Computer Society. He is on the editorial board of the *International Journal of Humanoid Robotics*.