Reading motor intention through mental imagery

**Daniel Lewkowicz[1,2], Yvonne Delevoye-Turrell[1,2], David Bailly[3,4],
Pierre Andry[3,4] and Philippe Gaussier[3,4]**

Abstract
Motor imagery is defined as a dynamic state during which the representation of a given motor act is internally rehearsed without overt motor output. Some evidence in experimental psychology has suggested that imagery ability is crucial for the correct understanding of social intention. The present study was conducted first to confirm that the nature of the motor intention leads to early modulations of movement kinematics. Secondly, we tested whether humans use imagery to read an agent's intention when observing the very first element of a complex action sequence. Results revealed early variations in movement kinematics between three different social actions and further showed that human agents can use these early deviants to anticipate above chance level the end-result before seeing the second half of the sequence. Response times in the observation task were similar in duration to those measured in the true production task, suggesting the use of motor imagery for trial categorization. Nevertheless, in a third study, the use of an artificial (neural network) classifier demonstrated that classification within the first 500 ms is possible without cognitive imagery processing. Hence, our results suggest that low-level motor indices afford intention reading without need for motor imagery but that human agents may use imaging beyond simulation to create an embodied sense of interactivity.

1 Introduction

Motor imagery is a cognitive process in which a subject imagines that he/she is producing a movement without actually performing it and without even tensing the muscles. It is a dynamic state during which the representation of a specific motor action is internally activated without any motor output (Jeannerod, 1994; Lotze & Cohen, 2006). A fast-growing number of studies are indicating that brain areas engaged in the actual performance of movements are also active during motor imagery (Dechent, Merboldt, & Frahm, 2004; Gerardin et al., 2000; Grezes & Decety, 2000; Hallett, Fieldman, Cohen, Sadato, & Pascual-Leone, 1994; Hanakawa et al., 2003; Jeannerod, 2001; Kimberley et al., 2006; Lotze et al., 1999; Sirigu, Cohen, Duhamel, & Pillon, 1995; Stephan et al., 1995). Besides the overlap in neural activation between imagery and execution, there are also similarities in the behavioral domain. For instance, the time to complete an imagined movement is similar to that needed for actual execution of that movement. This phenomenon known as mental isochrony has been reported in various tasks, e.g., handwriting (Decety, 1993; Decety & Michel, 1989), object manipulation (Frak, Paulignan, & Jeannerod, 2001), tapping on targets of different dimensions (Sirigu et al., 1995) and even whole body actions like walking through doorways of different widths (Decety & Jeannerod, 1995).

These findings have led to a theoretical position termed the simulation hypothesis (Jeannerod, 2001) that states that movement execution and motor imagery are driven by the same basic cognitive mechanism, with motor imagery being conceived as an "offline" operation of the motor areas in the brain. Through mental simulation, motor imagery would not only produce internal images of our body in action from a first-person perspective but it would also contain elements

[1]Université Lille Nord de France, Villeneuve d'Ascq, France
[2]URECA EA 1059
[3]Université Cergy-Pontoise, Cergy-Pontoise, France
[4]ETIS—ENSEA, CNRS UMR 8051

Corresponding author:
Yvonne Delevoye-Turrell, Laboratory URECA EA 1059, Université Lille Nord de France, Rue du Barreau, 59653 Villeneuve d'Ascq, France.
Email: yvonne.delevoye@univ-lille3.fr

of kinesthetic sensations, which act as a substitute for the sensory feedback that would arise if the movement were executed (Naito et al., 2002). Hence, motor imagery is a cognitive activity that requires and uses mental simulation with in addition an *embodied* aspect, which is created through the voluntary activation of those sensory-motor areas in the brain that would be activated if we were the true actor of a movement (Mulder, 2007).

What would then happen to our own movements when watching someone doing another movement? Brass et al. (2000) conducted a study in which subjects were instructed to perform a finger movement as rapidly as possible. Results indicated that the initiation times of the movements were faster when the to-be-performed action was identical to the movement that was observed, suggesting a priming effect of similarity between executed and observed actions (Brass, Bekkering, Wohlschläger, & Prinz, 2000). More recent studies have confirmed this effect for hand posture (Urgesi, Candidi, Fabbro, Romani, & Aglioti, 2006) and for sequence learning (Pascual-Leone, 2001; Zijdewind, Toering, Bessem, van der Laan, & Diercks, 2003). More importantly for the purpose of the present study, observing an action primes the very muscles needed to perform the same action (Craighero, Bello, Fadiga, & Rizzolatti, 2002; Fadiga, Craighero, Buccino, & Rizzolatti, 2002; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995). Hence, observed actions activate in the observer's brain the same mechanisms that would be activated, were that action intended or imagined by the observer (Gallese & Goldman, 1998).

The question asked in the present contribution is the nature of the information that is used during action observation to permit intention reading. Indeed, chronometric studies have shown that viewers decide whether the presented stimulus is a left or a right hand, by engaging implicit motor imagery, a simulation that reflects in fact the time constraints of limb rotation (Parsons, 1994, 2001). It has also been shown that humans use implicit imagery when asked to decide how they would grasp a handle appearing in a variety of orientations with the subjects' choices being in strict agreement with the grips chosen during actual grasping (Johnson, 2000). However, in these studies the exact nature of the information used to simulate, to imagine and then, to select a response is not defined or discussed. In the present study, we will consider the possibility that body kinematics are characterized by early deviants that reflect the social intention of the agent. We will further demonstrate that these early deviants can be used as non-contextualized cues to anticipate the end-point motor intention of an observed actor.

Optimal control models of biological movement have been more successful than any other class of models to predict empirical findings as diverse as movement corrections from unexpected changes or responses to

global perturbations (Shadmehr & Mussa-Ivaldi, 1994), but also modeling the structure of motor variability (Gordon, Ghilardi, & Ghez, 1994; Messier & Kalaska, 1999; van Beers, Haggard, & Wolpert, 2004) and of generic motor laws (Lacquaniti, Terzuolo, & Viviani, 1983). Optimal control theory helps to understand why humans move in a specific way by showing that the movements are optimal with respect to an ecologically valid criterion. But many experimental cases have shown that this optimal control is not a rule and that in many cases, naturalistic movements do not respect the optimal control laws. Some have named these exceptions task-irrelevant factors (Herbort & Butz, 2011). A much-studied example is that of the end-state comfort effect (e.g., Rosenbaum, Chapman, Weigelt, Weiss, & Van der Wel, 2012). In this situation, the initial discomfort of the arm posture is tolerated for the sake of final comfort, because the end-state comfort is critical for good execution of future task demands (Rosenbaum & Jorgensen, 1992; Rosenbaum et al., 1990; Short & Cauraugh, 1997).

In the case of complex sequential actions (as reaching to use an object), it seems legitimate to ask whether the constraints of subsequent elements would have an effect on those previous elements that are embedded in the same sequence. Thus, in an optimal point of view, the optimization criteria would not be based on performing correctly each sub-element of the sequence but rather, it would be to follow a global optimal criteria that will thus constrain the entire motor sequence and its performance. In a pick and place task, Lewkowicz & Delevoye-Turrell (under revision) manipulated the constraints set upon the final target of a two-sequence action. They reported that both first and second sub-elements of the sequence were affected by specific end-position constraints (distance, accuracy, stability). Moreover, they found that *coupled planning* (i.e., embedding the two movements together) occurred only if subjects were able to anticipate the consequences of both sub-elements prior to the initiation of the entire motor sequence. Thus, it was suggested that despite the fact that the reach-to-grasp sub-element possesses its own specific constraints (object size, relative distance and timing), the criteria set for the second sub-element (required speed, accuracy and efficiency) would constitute a global and dominant speed–accuracy trade-off property that would back-propagate to modulate very early on the planning and/or the execution of the very first sub-element of the sequence.

Motor intention may also be a parameter that modifies early sequence kinematics. For example, Ansuini and colleagues (2008) measured the prior-to-contact grasping kinematics for reach-to-grasp movements performed toward a bottle filled with water. By comparing hand shaping across tasks involving different subsequent actions—pour the water into a container; throw the bottle; move the bottle from one spatial location to

another—the authors demonstrated how the prior intention in grasping the object strongly affected the positioning of the fingers during the reaching and the contact phase of the action (Ansuini, Giosa, Turella, Altoè, & Castiello, 2008; Ansuini, Santello, Massaccesi, & Castiello, 2006). In another series of studies, Becchio et al. investigated the effects of social context on reach-to-grasp actions. They found initial adjustments reflecting specific planning strategies (Becchio, Sartori, Bulgheroni, & Castiello, 2008a) as well as online adjustments (Sartori, Becchio, Bulgheroni, & Castiello, 2009) when performing under social context (Becchio, Sartori, & Castiello, 2010).

Hence, there is growing evidence that both end-point constraints and social contexts affect movement kinematics early on. Recently, it was suggested that these deviants are meaningful and may be used to read motor intention. For example, when observing actions performed under social context (or not), Castiello and collaborators demonstrated that humans can successfully use kinematic cues of reach-to-grasp movements to predict the final goal of the action (Sartori, Becchio, & Castiello, 2011). However, other contextual cues may have been used, e.g., the face, body movements and gaze orientation to infer the goal of the action and its social underlying intention. To study this aspect, Stapel and collaborators (2012) investigated specifically the contribution of contextual cues and found that both sources were important: intention reading was more accurate if the observed actions were placed within a meaningful context (Stapel, Hunnius, & Bekkering, 2012). Nevertheless, humans can exploit subtle movement cues alone as suggested by an elegant study using point-light displays of simple reach to grasp movements (Manera, Becchio, Cavallo, Sartori, & Castiello, 2011). However, in this later study, choices were to be made between voluntary actions that were performed under different speeded conditions (cooperation vs. competition; fast vs. slow), which may have induced artificial and significant kinematic differences.

In the present study, we wanted to examine the capacity of humans to read motor intention: (1) during voluntary motor sequences performed under constant speeded constraints, and to discuss (2) the need of motor imagery in the process. Complexity was manipulated both on the fact that the motor sequence was constituted of two elements (elt1: reach to grasp; elt2: lift to place) and was performed during a true social-interactive game with a conspecific. As such, we recorded sequential actions during an ecologically inspired task (*Jungle Speed*) in a face-to-face game using a unique manipulated object. One important criterion in the design construction was to select a first common action (i.e., the reach movement) that led to different subsequent situations, which had direct impact on the game's progression.

Confronting Jacob and Jeannerod's (2005) *reading motor intention* hypothesis, we hypothesized that human agents are able to read motor intention through the simple observation of arm kinematics of the first element of a two-sequence action. This is possible due to the fact that arm kinematics of the reach to grasp movements reveal specific deviants in function of goal intention from an ideal optimized trajectory. Finally, if motor imagery is not necessary for intention reading, and if low level motor simulation is sufficient, then an artificial neural network (ANN) should be able to learn from the deviants and predict as well as humans, the motor intention of an observed agent. In the following section, we first describe the methods used to make the observation videos (Part A), which were then played to human agents (Part B) and used as input parameters to an ANN (Part C).

2 Method

2.1 Part A: creating the stimuli

Two adults participated in the study, one as the experimenter and the other as the subject. Both participants were right handed as verified with the Edinburgh Handedness Inventory (Oldfield, 1971). They had no prior knowledge of the experiment and provided informed consent before participating in the experimental session that lasted approximately 90 min. The subjects' movements only were recorded using (1) a video camera (Sony Handycam) and (2) 4 Oqus infrared cameras (Qualisys system). To provide the means to analyze arm kinematics, infrared reflective markers were placed on the index (base and tip), the thumb (tip), the wrist (scaphoid and pisiform) of the subject, as well as on the object. Care was taken as to provide no contextual information within the video clips (torso, gaze, face expression), i.e., only the hand and the target object were fully in view. Cameras were calibrated before each session, allowing the system to reach a standard deviation smaller than 0.2 mm, with a 200-Hz sampling rate. Three different positions were indicated on the tabletop by black tape and symbolized three specific locations that will be referred to in the next section as the placing positions: "Play", "Me", "You" (Figure 1a).

2.1.1 The game.
Both the subject and the experimenter were seated at a table, facing each other. The starting position for both participants was a point placed at midline 80 cm in front of the body. The object that was to be manipulated was a wooden dowel (width 2 cm; height 4 cm) that was placed precisely 20 cm in front of the starting position of the subject. The subject's task was to reach and grasp the dowel between thumb and index finger in order to move it from the initial position to one of three placing positions, during an adapted version of the *Jungle Speed* game (Asmodee editions.).
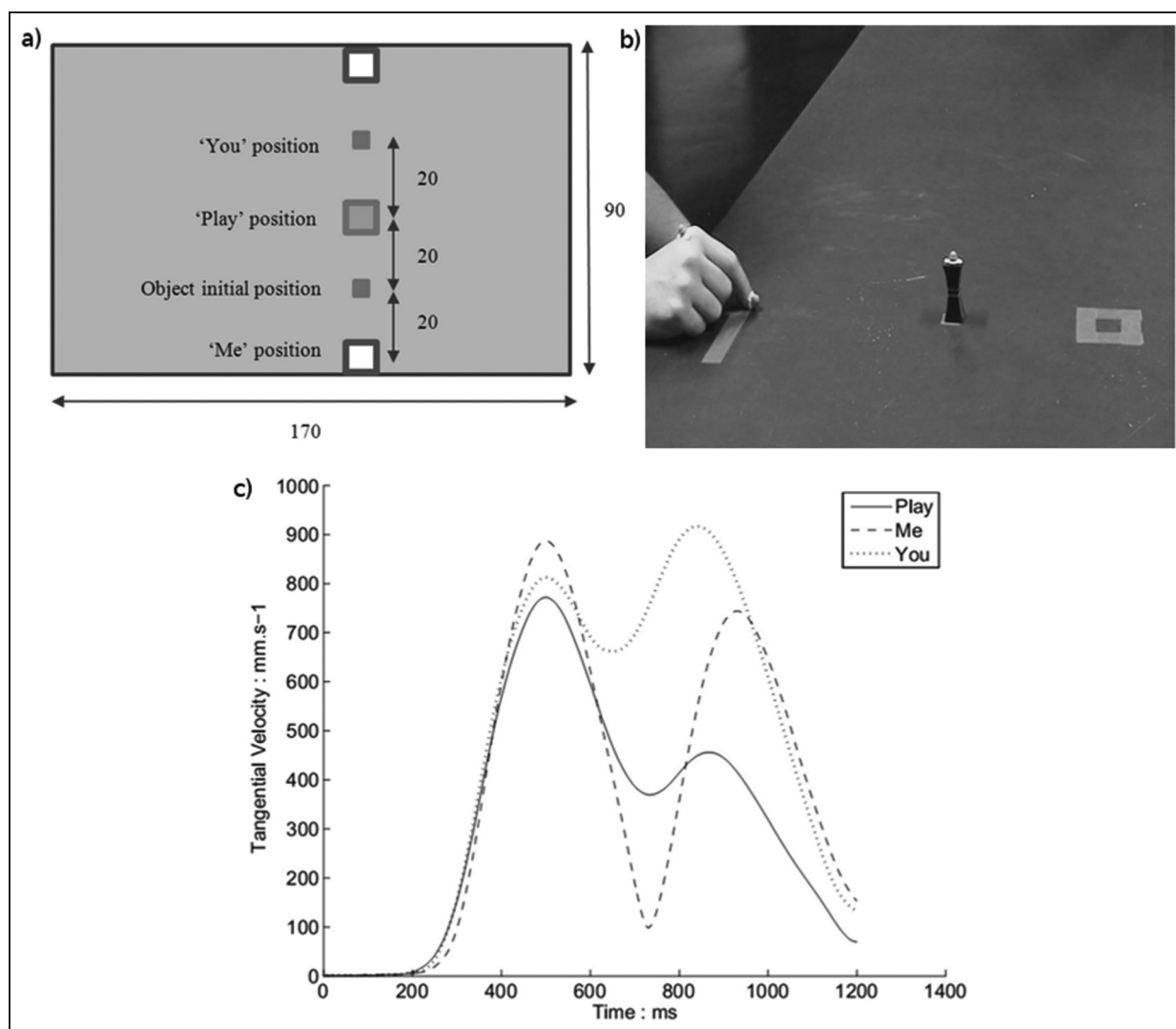
Figure 1. (a) Experimental set-up showing the "Play", "Me" and "You" final position and their respective distances from the initial object position. The white squared areas are the starting hand position for both the subject (bottom) and the experimenter (top). (b) Example of stimuli. This close view was used for the movie clips to avoid any contextual effect (no body, no head). (c) Typical velocity profiles observed in the "Play", "Me" and "You" conditions are presented with full, dashed and dotted lines, respectively. Note that total movement time, the magnitude and the time to peak velocity of the first element of the sequence are significantly affected by motor intention (i.e., the experimental conditions).

A trial was defined as a series of three moves: initiate, compete, reward. The initiate-move required subjects to pick and place the dowel on the "Play" position. Then, at a "go" signal (given by the computer), both participants were required to perform the compete-move, i.e., to reach for the dowel as quickly as possible. For each win, 1 point was scored. The reward-move was performed by the subject who picked up the dowel to place it on the "Me" position (if he/she had won the point) or on the "You" position (if the point was attributed to the experimenter). Each trial started with the dowel placed by the experimenter at the initial position and with the participants pinching index and thumb together, with the fingertips set upon the starting position (Figure 1a). Time pressure was set only on the

compete-move, which was not recorded. A block ended when one of the two players reached a total of 20 points. The game consisted of four blocks of approximately 40 trials.

2.1.2 The recordings. The best 16 video recordings of each category ("Play", "Me" and "You") were selected based on their mean bit rate quality (>4500 kbps) and kept for future use as stimuli. Each sequence included a 1-s time interval before the initial movement onset, and was cut exactly one frame before the index finger contacted the object. Movies were compressed with FFdshow codec (MJPEG) at 50 frames/s with a screen resolution of 720×576 pixels (Figure 1b). Video clips were synchronized to the recordings of arm kinematics.

2.1.3 Analyzing arm kinematics. Positional data points were filtered using a dual fourth-order Butterworth low-pass filter (fc = 15 Hz; forward and backward) and tangential 3D instantaneous velocities were calculated. A threshold of 20 mm/s was used to determine the onset of movement. All velocity trajectories were bell shaped and consisted of two "bells": (1) the first bell corresponds to the *reach to grasp* element, which will be referred to in the following as the first element of the sequence; (2) the second bell corresponds to the *lift to place* element, which will be referred to in the following as the second element of the sequence (Figure 1c). The amplitude of peak velocity of the first element (APV1) was extracted using the local maxima (first zero-crossing of acceleration). The end of the first element was determined as the time of occurrence of the local minima (second zero-crossing of acceleration) between the first and the second element-peaks. The duration of the first element (MT1) was calculated as the time interval between the onset and the end of the first element. Each parameter was submitted to a repeated-measure analysis of variance (ANOVA) with *Block* and *Category* as within factors; an alpha level of significance was set to .05.

Results revealed an absence of Block effect on APV1, $F(2,30) = 3.056$, $p = .062$, $\eta_p^2 = .17$ and MT1, $F(2, 30) = 2.727$, $p = .082$, $\eta_p^2 = .15$, suggesting similar movement properties across block repetitions, both for amplitude of peak velocity (Block 1: $M = 838$, $SD = 31$ mm/s; Block 2: $M = 824$, $SD = 26$ mm/s; Block 3: $M = 847$, $SD = 36$ mm/s) and for movement duration (Block 1: $M = 440$, $SD = 35$ ms; Block 2: $M = 459$, $SD = 28$ ms; Block 3: $M = 462$, $SD = 24$ ms). More importantly, results showed a global effect of Category both on APV1, $F(2,30) = 58.463$, $p < .001$, $\eta_p^2 = .80$, and on MT1, $F(2,30) = 55.821$, $p < .001$, $\eta_p^2 = .79$ indicating that when subjects reached for the dowel, peak velocities were lower for the "Play" ($M = 796$, $SD = 27$ mm/s) than for the "Me" category. Furthermore, movement durations were longer for the "Play" ($M = 481$, $SD = 34$ ms) than for the "You" category. The reach to grasp movement in the "Me" category showed both higher peak velocities ($M = 893$, $SD = 37$ mm/s) and longer movement durations ($M = 479$, $SD = 22$ ms) than that observed in the "You" category (APV1: $M = 820$, $SD = 28$ mm/s, MT1: $M = 402$, $SD = 19$ ms). As seen in Figure 2, these results lead to an overlap of 30% across conditions only. Hence, through the use of these two parameters alone, a classification system should be able to categorize in a predictive matter these three motor sequences, which possess distinct motor intentions.

2.2 Part B: human prediction of ongoing actions

In the present study, the short video clips were presented to a panel of human subjects to test whether human agents are able to predict the goal of a
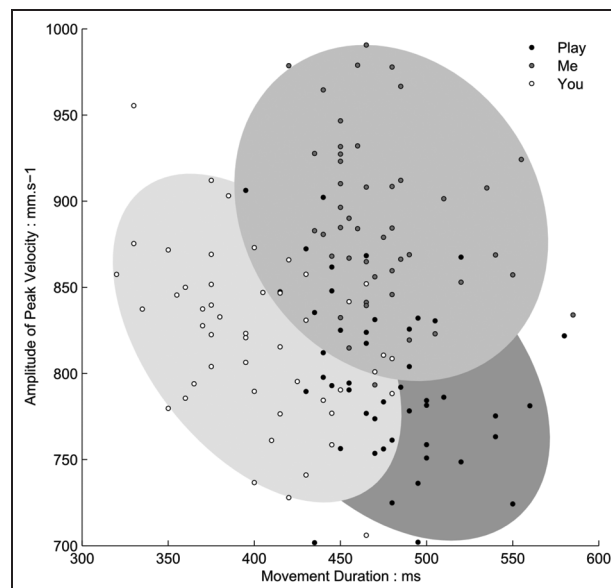

Figure 2. The spatial representations of the first element of the sequence where the 144 stimuli that were recorded using the Qualysis 3D motion capture system. The *X*-axis codes movement duration. The *Y*-axis codes maximum amplitudes of peak velocity of the first element of the sequence. In the three "Play", "Me" and "You" categories, the 95% confidence ellipses are plotted in dark, gray and white colors, respectively. The overlapping areas were estimated as following: three categories=6.8%, two categories=24.6%, no overlap=68.6%.

sequential action when shown only the first element of a sequence, i.e., the reach to grasp element. To test the hypothesis that a simple artificial classifier could also learn to discriminate between social categories (because the biological kinematics are affected by the agent's intention), the 3D instantaneous velocities of the arm kinematics corresponding to each short video clip were fed as input parameters to a classical feedforward neural network (NN) with one hidden layer. These experiments are presented in the following sections, Part B and Part C, respectively.

2.2.1 Participants. Twenty-six young adults (mean age: 21.82 ± 2.76 years, range = 18–29 years) participated in the study. All subjects were right handed (Oldfield, 1971) and had no prior knowledge of the experimental goals. Subjects provided informed consent before participating in the experimental session that lasted approximately 45 min.

2.2.2 Apparatus and software. Participants were seated comfortably facing a table in a dark and silent room. For each trial, participants started by placing their hand on response keys that were delimited by tape placed directly on the number-pad (numbers 2, 5 and 8 were used). Stimuli were presented on a laptop

computer with MATLAB software (Mathworks) in the PsychToolbox environment. Analogical scales (10-cm lines) were used for self-evaluation of performance levels.

2.2.3 Experimental procedure.

The participants' task was to answer after each video clip presentation whether the social intention of the sequence was "let's Play" (key5), "for Me" (key2) or "for You" (key8). A 1-s blank screen was displayed in between two trials. Participants were instructed to give their answers as fast and as accurately as possible. They were obliged to provide an answer within a 4-s time window; otherwise, the trial was cancelled and presented at the end of the block. A feedback message was given when responses were too slow. Each block consisted of the random presentation of a series of 48 stimuli, i.e., 16 different video clips for each of the three categories (Play; Me; You). At the end of the block, a 5-min pause was systematically taken. At this occasion, participants filled in an analogical scale to provide a subjective judgment of the performance they thought to have achieved on a scale from 0 (very poorly) to 10 (very well).

2.2.4 Dependent variables and statistical analyses.

For each trial and participant, response times were calculated as the time interval between the presentation of the last frame of the video and the subjects' key press. Mean percentages of correct responses, mean response times and mean self-evaluation scores were calculated for each category and submitted to a repeated-measure ANOVA with *Block* and *Category* (Play; Me; You) as within factors. We also conducted sub-analyses. (1) For the percentages of correct responses, scores for each category were compared with the reference constant, i.e., the random answer value of 0.33, with a single sample *t*-test. (2) To gain an indicator of motor imagery, response times were compared with the true movement duration presented on the video (MT1 + MT2) for each category. In all of these analyses, the alpha level of significance was set to .05.

2.2.5 Response times.

Statistical analyses revealed an absence of Block effect, $F(2,50) = 1.401$, $p = .256$, indicating that participants answered as fast in Block 1 ($M = 878$, $SD = 382$ ms), in Block 2 ($M = 848$, $SD = 315$ ms), and in Block 3 ($M = 944$, $SD = 316$ ms). An absence of main effect was also found for Category, $F(2,50) = 2.621$, $p = .083$, indicating that participants answered as fast for "Play" ($M = 900$, $SD = 294$ ms), "Me" ($M = 866$, $SD = 294$ ms), and "You" categories ($M = 905$, $SD = 300$ ms). The sub-analysis revealed an absence of difference between response times and movement times for "Play" ($M = 900$ vs 898 ms), $t(25) = 0.25$, $p = .980$, "Me" ($M = 866$ vs 964 ms),
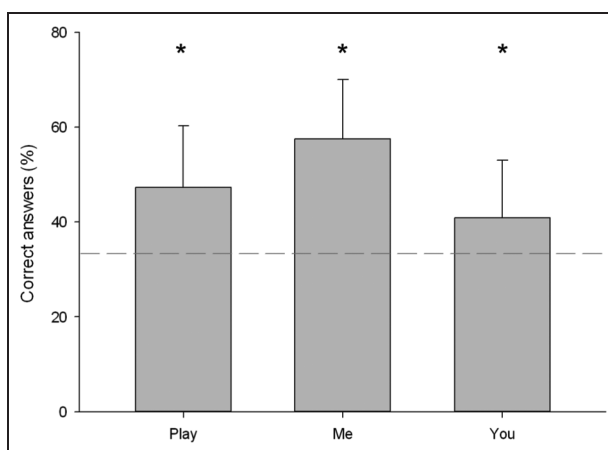

Figure 3. Mean percentage of correct responses (standard deviations as error bars) obtained in the experiment reported in Part B. The dotted lines represent the random baseline level of 33.33%. Note: *α=.05; **α=.01.

$t(25) = 1.694$, $p = .103$ and "You" categories ($M = 905$ vs 907 ms), $t(25) = 0.032$ $p = .975$. These results indicate that the participants waited a time interval before giving their response. However, it is here difficult to conclude on the use of motor imagery using the classic approach of comparing trial durations in true and imagined situations, respectively, as there was an absence of Category effect on movement time.

2.2.6 Percentages of correct responses.

There was an absence of Block effect on classification performances, $F(2,50) = 0.102$, $p = .903$. However, a main effect of Category was obtained, $F(2,50) = 16.022$, $p < .001$, $\eta_p^2 = .39$. *Post hoc* Scheffé analyses further indicated that participants were more accurate for trials in the "Me" category ($M = 57.53$, $SD = 13.02\%$) than in the "You" ($M = 40.87$, $SD = 12.12\%$) and in the "Play" category ($M = 47.27$, $SD = 13.04\%$). Single sample tests confirmed that each category was above random ("Me": $t(25) = 5.463$, $p < .001$; "Play": $t(25) = 9.914$, $p < .001$; "You": $t(25) = 3.169$, $p = .004$). These results are presented in Figure 3.

2.2.7 Self-evaluation.

The main effect of Block did not reach significance for the self-evaluation scores, $F(2,50) = 2.674$, $p = .079$. Across blocks, participants self-evaluated their own performances as only slightly better than chance ($M = 40.90$, $SD = 13.15\%$). As correct classification was obtained for all categories, statistical analysis confirmed that participants significantly under estimated their true performance level, $F(1,25) = 11.680$, $p = .002$, $\eta_p^2 = .32$; this was especially true in Block 3 in which performance level was subjectively reported as being the worst ($M = 38.17$, $SD = 19.30\%$ vs. $M = 48.80$, $SD = 12.46\%$).

2.3 Part C: predicting sequential actions using artificial neural networks

In the following section, we describe the simple feedforward NN that was developed in the present study to discuss the possibility to anticipate the motor intention of an agent without the need of high order cognitive imagery processes.

2.3.1 Architecture and learning procedure.
A simple classification NN was constructed with N neurons (1–23 neurons) as inputs, three hidden neurons and three output neurons (one for each category). The N size is the number of samples taken from the kinematics, which may be used for the classification task, i.e., the sub-selection of the total movement duration. Activation functions for the output layers were symmetrical and sigmoid, between −1 and 1. It is the case that the output vector can be interpreted as a logical value (+1) when the associated category is activated; the logical value (−1) is attributed otherwise. In the present case, a single component was set to +1.

For each trial, only the section for which there was a hand-movement was considered, i.e., from the start of the first element to the end of the second element of the sequence. This time interval was determined in reference to the 20-mm/s threshold criteria used in part B of the present study. The instantaneous velocity in 3D was then calculated between the two subsequent frames of the wrist's position. A *sampling* parameter was used to compute the average velocity across 10 frames. This *sampling* parameter is a constant time window of 50 ms without overlap that echoes the perceptual binding of motion events. Thus, the procedure afforded a vector of N samples that were then used as input parameters for the network. Finally, a training-set (25%) and a test-set (75%) were randomly picked from the 144 different kinematic recordings. For each possible size of the input vector (i.e., time window for kinematic recognition), 20 different networks were trained to obtain a classification performance. The results for mean responses and variances across the 20 networks are described in the result section as the NN success rate (this value is always lower than the best performing network).

By varying the amount of data fed as input parameters (1–23), we computed the classification performance from multiple time windows (50–1150 ms). The learning procedure that was chosen is a back-propagation algorithm using the FANN library (Nissen, 2003). Target error (to stop the learning) was set to mean standard error (MSE)<0.001, with a maximum number of epochs set to 10,000, and 300 iterations between each test for the evaluation of target global error. In a preliminary study, we used networks with an augmented number of hidden units and good performances were also obtained. Nevertheless, as an
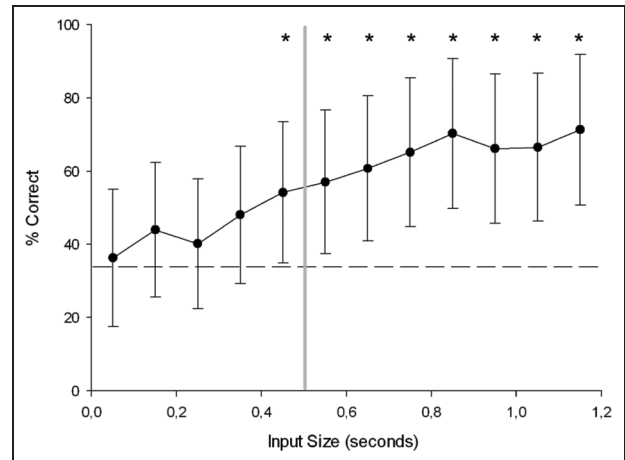

Figure 4. Mean percentage of correct classifications (standard deviations as error bars) obtained with artificial neural networks. The horizontal axis codes the input size (step1: 50 ms to step23: 1150 ms). The vertical axis codes the mean and standard deviation values of the 20 networks. The dotted lines illustrate random baseline. The vertical gray bar indicates the end of the first motor element of the sequence.

ANN with three hidden units succeeded in learning the classification task, we opted in the present report for the simplest system in order to minimize risks of over learning.

2.3.2 Classification results in function of time.
The artificial classifier was able to converge in most cases. The classifier succeeded in discriminating between categories for input sizes above nine, i.e., with at least 450 ms of movement information. For the input size of nine, single sample t-tests confirmed that all categories were recognized above chance level, $p<.001$: "Play" category ($M = 55.70$ $SD = 8.08\%$); "Me" category ($M = 56.70$ $SD = 4.16\%$), and "You" category ($M = 50.33$ $SD = 5.63\%$). Figure 4 presents the detailed results obtained for 12 different input sizes, between one (50 ms) and 23 (1150 ms). From the input size of five (250 ms) to nine (450 ms), only two categories were successfully recognized while the other remained below chance level. Below the input size of five, only one category was correctly classified. Depending on initial conditions and learning procedures, we observed that the categories that were more rapidly identified could change. Hence, further work is required for further interpretation of these category-specific effects. Nevertheless, the crucial point to note here is the fact that by 450 ms all categories were classified above chance level, a point in time that occurred before the end of the first element of movement sequence (Figure 4). This finding confirms the capacity of a simple network to categorize motor intention through the use of low-level kinematics, before the initiation of the second element of the motor sequence.

3 Discussion

Recent studies in robotics, cognitive sciences, and motor control have described high-skilled robotic systems that are able to solve complex tasks such as navigation, object recognition and even fine object manipulations. The development and application of e.g. the optimal theories of motor control (Doya, 2000; Todorov, 2004; Todorov & Jordan, 2002) have provided the means to create systems that have even mastered the problem of gravity, allowing for the execution of complex physical motor sequences such as walking and dancing. In contrast, robotic systems still reveal limited adaptive capacities with humans, especially for rhythmic situations for which the *robots lack crucially of interactivity*. The absence of interactivity may be because the optimal theories of motor control are no longer sufficient (Sisbot, Marin-Urias, Broquere, Sidobre, & Alami, 2010) to account for the behavioral data that have been reported in experimental psychology in cases of human social interactions. Indeed, an increasing number of studies are reporting that for humans acting alone (motor interaction), the motor actions performed do not follow kinematics that are similar to those used when the person acts in collaboration/competition with a partner (social interaction; Becchio et al., 2010). These behavioral deviants that are observable in the early stages of motor execution may play a functional role and be used by conspecifics as social cues to infer motor intention. Thus, they need to be considered when creating cybernetic systems that afford true human-robot interactivity (Andry, Gaussier, Moga, Banquet, & Nadel, 2001; Gaussier, Moga, Quoy, & Banquet, 1998).

In the present contribution, we report experimental data confirming that motor intention modifies movement kinematics within the first hundreds of milliseconds. More specifically, our findings demonstrated first that the three different motor intentions that were created using a simplified version of the *Jungle Speed* game modified the kinematics of the first (reach) element of the motor sequence. Second, human agents were able to classify rapidly (<1 s) and above chance level (>40%), the trial category through the observation only of the reaching movement of the sequence. Response times were long (> 800 ms), which may have permitted the use of motor imagery by our human participants. Nevertheless, using a classic feedforward NN, results indicated that motor imagery is maybe not necessary since the ANN was able to categorize trials through the use of low-level kinematics within the first 450 ms of the sequence. In the following section, we discuss these findings in more detail and open the discussion on the need to use motor imagery by human agents in order to experience true interactivity.

3.1 Kinematics reflecting motor intention

In the abundant literature of manipulative actions, the effects of end-point constraints on the early parts of movement kinematics have been investigated extensively. In non-social situations, multiple parameters have been reported to modify and shape hand trajectory in two-element sequences such as second-target distance (Gentilucci, Negrotti, & Gangitano, 1997), end-target orientation (Haggard, 1998; Hesse & Deubel, 2010; Seegelke, Hughes, Schütz, & Schack, 2012), and second-action type (Armbrüster & Spijkers, 2006; Johnson-Frey, McCarty, & Keen, 2004; Marteniuk, MacKenzie, Jeannerod, Athenes, & Dugas, 1987; Mason, 2007; Weiss, Jeannerod, Paulignan, & Freund, 2000). In social tasks, final-goals have also been reported as having an effect on reach-to-grasp kinematics such as giving vs. placing an object (Becchio et al., 2008a), cooperative vs. competitive actions (Becchio, Sartori, Bulgheroni, & Castiello, 2008b; Georgiou, Becchio, Glover, & Castiello, 2007), absence vs. presence of social request (Ferri, Campione, Dalla Volta, Gianelli, & Gentilucci, 2011; Sartori et al., 2011), and even verbal communicative vs. non-communicative intentions (Sartori, Becchio, Bara, & Castiello, 2009). The kinematic effects reported in the present study are consistent with this literature and suggest that when planning a sequential action with multiple motor elements, the requirements of the endpoint element are back propagated to constrain the way the very first element of the sequence will be planned and performed. Thus, it is possible to suggest that low-level motor components may contain early indices that reflect the end-point motor intention of an agent.

3.2 Reading motor intention …

In the present study, each trial started with the similar movement of reaching to grasp an object, whether the action was then to give, keep or displace the object to another position on the table. Indeed, subjects initiated their move with their hand placed on the starting pad of the playing area, and reached for and grasped the wooden-peg that was always located at the same position on the table. However, the second part of the sequence was specific and directly related to motor intention: lift the wooden peg to take it ("Me" category), to give it ("You" category) or to place it on the table ("Play" category). Thus, any kinematic deviants observed on the first part of the sequence may be related to the social intention of the second part. By measuring two basic motor parameters (*peak velocity* and *movement duration*), we showed that it was possible to dissociate the three types of social interaction categories (*Figure 2*). We then tested the fact that human observers could use these deviants to classify observed

actions above chance level. The video clips were created in order to show the first element only, without any contextual cues; care was also taken to cut the end of the reaching action, one frame before object contact, in order to avoid providing any cues on movement direction of the second element of the sequence. Even if participants found the task very difficult and thought to have responded randomly, our findings demonstrated that classification is possible and that in certain cases, the participants' performance can be extremely precise (up to 67% of correct classification for the best of participants). But how is this possible?

3.3 … through mental imagery

According to the simulation hypothesis (Jeannerod, 2001), both observing and imagining an action activates the same neural correlates than actual execution. This possibility has since been confirmed in behavioral experiments that have shown that the time to complete an imagined movement is similar to that needed for actual execution of that same movement. When human subjects were required to classify observed actions, response times were found to be significantly longer than classic simple reaction times (>500 ms). They were in fact as long as the durations of the actual second motor element, with an absence of time differences. These findings suggest that subjects were performing motor imagery of the second motor element in order to simulate the motor intention of the observed agent, motor image on which they based their decision. It has been proposed that cognitive high-level functions like motor simulation would also be the basis for the emergence of intention understanding in more complex situations like role playing, theory of mind or empathy (Gallese, 2001; Gallese, Keysers, & Rizzolatti, 2004; Iacoboni, 2009). It would be interesting in a future study to include personality and individual characteristics to assess how emotional valence and social similarity may modulate the capacity to read intention through action observation.

3.4 An alternative low-level hypothesis

It is nevertheless possible that the understanding of motor intention is based on more low-level cue readings. Indeed, despite a total absence of contextual cues within the video clips (body, head, eyes), we demonstrated in the present study that participants were able to read motor intention significantly above chance level. Hence, it might be that the subjects' responses are guided only by the slight kinematics deviances from the optimal trajectory. This would trigger a stimulus–response type of mechanism that would take place after years of interactive experiences; none cognitive by nature, it would require little resources but could lead to high recognition errors, especially in poorly known environments. As a first investigation of this alternative low-level hypothesis, we reported here a second method of investigation (presented in part C) for which we used a very simple NN classifier and we showed that this NN was able to categorize the three categories of social interaction above chance level. The NN stabilized within the first 450 ms, which suggested that the classification was terminated *before* the end of the first motor element of the sequence. The performance level reached by the NN was similar to that observed in human individuals suggesting that stimulus–response coding could be sufficient for intention reading. However, it is to note that the NN was fed with extracted tangential velocity samples as input. Thus, the NN on the one hand and the humans on the other reached similar performance levels using different strategies to solve the puzzle of motor goal inference. Future studies are now required to investigate further the nature of the information that provides the best classification of motor goals/intentions. In addition, it will be important to determine whether human individuals could reach similar performance levels using direct coding of the low-level kinematic parameters (see the resonance theory by Viviani, 2002) or whether the kinematic deviants are simply a byproduct and hence, even for the simplest actions, humans need to engage in a cognitive simulation process to understand motor intention (for a debate see e.g. Kilner, Friston, & Frith, 2007).

It is to note that correct classification of the three social categories was far from being perfect, reaching in the best of cases 67% of correct identification. Hence, kinematics can be used for predicting ongoing actions but cannot be the only source used by human agents to judge motor intention. It has been shown that during natural sequential tasks (i.e., preparing a sandwich), eye movements are stereotyped and predictive (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Pelz, Hayhoe, & Loeber, 2001), with the eyes preceding the hand movements in a systematic way (Johansson, Westling, Bäckström, & Flanagan, 2001). It is thus possible that using both gaze position and the hand kinematics, an observer is able to increase the efficiency of intention reading (Bekkering & Neggers, 2002).

3.5 Perspectives for interactive and social robotics

The application of our work would be to develop robots that afford true interaction, i.e., being able (1) to read motor intention in human kinematics in order to adapt but also (2) to move with biological realistic kinematics, in order to allow others to understand the intention of the robot. The Aibo robot designed by Sony is a good example of the limitations of current approaches. The robot is quite impressive during the first minutes of functioning (smooth movements, nice physical design …) but rapidly people stop trying to interact with Aibo. Following the data presented here,

we hypothesize that a humanoid robot could become interactive if it moved following the laws of biological movement with action sequences that integrate back-propagation of terminal intention. Such a phenomenon would provide the means for human agents to read intentionality and thus, gain in understanding the goal of the robot's movements. Furthermore, including social deviants in the motor kinematics within early steps of motor sequences would also allow safe interaction with large industrial robots by affording humans the possibility of anticipating false moves in joint actions that share similar work spaces.

Implementing robots with the architecture necessary to "afford intentionality" would need to integrate the different brain regions that are known to play a role in motor planning and motor-sensory predictive mapping. De Rengervé et collaborators (De Rengervé, Hirel, Andry, Quoy, & Gaussier, 2011) have recently reported on such an architecture, which included amongst other areas, the cerebellum and the basal ganglia. Tested on both software and hardware, this neural architecture has demonstrated its efficiency on data collected in a hydraulic robotic arm. With a series of imitation trials, this system demonstrated the capacity to learn how to perform sequential actions that respected biological laws, i.e., to perform movements with kinematics that mirror those performed by human agents. As such, this robot arm suggests increased interactivity with human agents affording augmented interaction in both time and space (none published results). Ongoing studies are now being conducted to assess whether this interactivity is associated to an increase in the capacity of human collaborators to read the robot's intention.

4 Conclusion

We have reported experimental data demonstrating that it is possible to read motor intention through the simple observation of kinematic deviants. Classification capacities were significantly above chance level and provided human subjects the means to dissociate between three different socially oriented actions. We argue in the present study that reading intentionality may not depend on a high-level cognitive function as suggested in the psychological literature. Internal simulations may not be systematically required and understanding other intentions may, in certain cases, relate to a direct coding of those kinematic deviants that back propagate from end-point to early on during sequence execution. This direct coding would emerge through years of learning, during interactions with adult conspecifics. As a first step to support this hypothesis, we report in the present study simple NNs that were able, after learning the meaning of kinematic deviants, to classify the three categories of actions to the *same degree of accuracy as the human participants, without the* *need of complex cognitive processes*. These preliminary results stress the importance of further development of the optimal theories of motor control to include the more cognitive aspects of social context.

References

Andry, P., Gaussier, P., Moga, S., Banquet, J. P., & Nadel, J. (2001). Learning and communication via imitation: An autonomous robot perspective. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 31*(5), 431–442.

Ansuini, C., Giosa, L., Turella, L., Altoè, G., & Castiello, U. (2008). An object for an action, the same object for other actions: Effects on hand shaping. *Experimental Brain Research, 185*(1), 111–119.

Ansuini, C., Santello, M., Massaccesi, S., & Castiello, U. (2006). Effects of end-goal on hand shaping. *Journal of Neurophysiology, 95*(4), 2456–2465.

Armbrüster, C., & Spijkers, W. (2006). Movement planning in prehension: Do intended actions influence the initial reach and grasp movement? *Motor control, 10*(4), 311–329.

Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008a). The case of Dr. Jekyll and Mr. Hyde: A kinematic study on social intention. *Consciousness and Cognition, 17*(3), 557–564.

Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008b). Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition, 106*(2), 894–912.

Becchio, C., Sartori, L., & Castiello, U. (2010). Toward you: The social side of actions. *Current Directions in Psychological Science, 19*(3), 183–188.

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science, 13*(4), 370–374.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44*(2), 124–143.

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40*(5), 492–502.

De Rengervé, A., Hirel, J., Andry, P., Quoy, M., & Gaussier, P. (2011). On-line learning and planning in a pick-and-place task demonstrated through body manipulation. In *2011 IEEE International Conference on Development and Learning (ICDL)* (Vol. 2, p. 1–6). Présentéà 2011 IEEE International Conference on Development and Learning (ICDL), IEEE.

Decety, J. (1993). Analysis of actual and mental movement times in graphic tasks. *Acta psychologica, 82*(1), 367–372.

Decety, J., & Jeannerod, M. (1995). Mentally simulated movements in virtual reality: Does Fitt's law hold in motor imagery? *Behavioural Brain Research, 72*(1), 127–134.

Decety, J., & Michel, F. (1989). Comparative analysis of actual and mental movement times in two graphic tasks. *Brain and Cognition, 11*(1), 87–97.

Dechent, P., Merboldt, K.-D., & Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Cognitive Brain Research.*

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology, 10*(6), 732–739.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience, 15*(2), 399–402.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*(6), 2608–2611.

Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., & Gentilucci, M. (2011). Social requests and social affordances: How they affect the kinematics of motor sequences during interactions between conspecifics. *PLoS ONE, 6*(1), e15855.

Frak, V., Paulignan, Y., & Jeannerod, M. (2001). Orientation of the opposition axis in mentally simulated grasping. *Experimental Brain Research, 136*(1), 120–127.

Gallese, V. (2001). The shared manifold hypothesis. From mirror neurons to empathy. *Journal of Consciousness Studies, 8*(5–7), 5–7.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493–501.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8*(9), 396–403.

Gaussier, P., Moga, S., Quoy, M., & Banquet, J. P. (1998). From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. *Applied Artificial Intelligence, 12*(7–8), 701–727.

Gentilucci, M., Negrotti, A., & Gangitano, M. (1997). Planning an action. *Experimental Brain Research, 115*(1), 116–128.

Georgiou, I., Becchio, C., Glover, S., & Castiello, U. (2007). Different action patterns for cooperative and competitive behaviour. *Cognition, 102*(3), 415–433.

Gerardin, E., Sirigu, A., Lehéricy, S., Poline, J.-B., Gaymard, B., Marsault, C., … Bihan, D. L. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cerebral Cortex, 10*(11), 1093–1104.

Gordon, J., Ghilardi, M. F., & Ghez, C. (1994). Accuracy of planar reaching movements. I. Independence of direction and extend variability. *Experimental Brain Research, 99*(1), 97–111.

Grezes, J., & Decety, J. (2000). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human brain mapping, 12*(1), 1–19.

Haggard, P. (1998). Planning of action sequences. *Acta Psychologica, 99*(2), 201–215.

Hallett, M., Fieldman, J., Cohen, L. G., Sadato, N., & Pascual-Leone, A. (1994). Involvement of primary motor cortex in motor imagery and mental practice. *Behavioral and Brain Sciences, 17*(02), 210–210.

Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Van Gelderen, P., & Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *Journal of Neurophysiology, 89*(2), 989–1002.

Hayhoe, M., Shrivastava, A., Mruczek, R., & Pelz, J. (2003). Visual memory and motor planning in a natural task. *Journal of Vision, 3*(1).

Herbort, O., & Butz, M. V. (2011). The continuous end-state comfort effect: Weighted integration of multiple biases. *Psychological Research, 76*(3), 345–363.

Hesse, C., & Deubel, H. (2010). Advance planning in sequential pick-and-place tasks. *Journal of Neurophysiology, 104*(1), 508–516.

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology, 60*(1), 653–670.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: a critique. *Trends in Cognitive Sciences, 9*, 21–25.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences, 17*(02), 187–202.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage, 14*(1), S103–S109.

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–hand coordination in object manipulation. *The Journal of Neuroscience, 21*(17), 6917–6932.

Johnson, S. H. (2000). Imagining the impossible: Intact motor representations in hemiplegics. *Neuroreport, 11*(4), 729–732.

Johnson-Frey, S., McCarty, M., & Keen, R. (2004). Reaching beyond spatial perception: Effects of intended future actions on visually guided prehension. *Visual Cognition, 11*(2), 371.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing, 8*(3), 159–166.

Kimberley, T. J., Khandekar, G., Skraba, L. L., Spencer, J. A., Van Gorp, E. A., & Walker, S. R. (2006). Neural substrates for motor imagery in severe hemiparesis. *Neurorehabilitation and Neural Repair, 20*(2), 268–277.

Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica, 54*(1), 115–130.

Lotze, M., & Cohen, L. G. (2006). Volition and imagery in neurorehabilitation. *Cognitive and behavioral neurology, 19*(3), 135–140.

Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., … Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *Journal of Cognitive Neuroscience, 11*(5), 491–501.

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research, 211*(3–4), 547–556.

Marteniuk, R. G., MacKenzie, C. L., Jeannerod, M., Athenes, S., & Dugas, C. (1987). Constraints on human arm movement trajectories. *Canadian Journal of Psychology, 41*(3), 365–378.

Mason, A. H. (2007). Performance of unimanual and bimanual multiphased prehensile movements. *Journal of Motor Behavior, 39*(4), 291–305.

Messier, J., & Kalaska, J. F. (1999). Comparison of variability of initial kinematics and endpoints of reaching movements. *Experimental Brain Research, 125*(2), 139–152.

Mulder, T. (2007). Motor imagery and action observation: Cognitive tools for rehabilitation. *Journal of Neural Transmission, 114*(10), 1265–1278.

Naito, E., Kochiyama, T., Kitada, R., Nakamura, S., Matsumura, M., Yonekura, Y., & Sadato, N. (2002). Internally simulated movement sensations during motor imagery activate cortical motor areas and the cerebellum. *The Journal of Neuroscience, 22*(9), 3683–3691.

Nissen, S. (2003). Implementation of a fast artificial neural network library (fann). *Report, Department of Computer Science University of Copenhagen (DIKU), 31*.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 709.

Parsons, L. M. (2001). Integrating cognitive psychology, neurology and neuroimaging. *Acta Psychologica, 107*(1–3), 155–181.

Pascual-Leone, A. (2001). The brain that plays music and is changed by it. *Annals of the New York Academy of Sciences, 930*(1), 315–329.

Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research, 139*(3), 266–277.

Rosenbaum, D. A., Chapman, K. M., Weigelt, M., Weiss, D. J., & Van der Wel, R. (2012). Cognition, action, and object manipulation. *Psychological Bulletin, 138*(5), 924–946.

Rosenbaum, D. A., & Jorgensen, M. J. (1992). Planning macroscopic aspects of manual control. *Human Movement Science, 11*(1–2), 61–69.

Rosenbaum, D. A., Marchak, F., Barnes, H. J., Vaughan, J., Slotta, J. D., & Jorgensen, M. J. (1990). Constraints for action selection: Overhand versus underhand grips. In *Attention and performance 13: Motor representation and control* (p. 321–342). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Sartori, L., Becchio, C., Bara, B. G., & Castiello, U. (2009). Does the intention to communicate affect action kinematics? *Consciousness and Cognition, 18*(3), 766–772.

Sartori, L., Becchio, C., Bulgheroni, M., & Castiello, U. (2009). Modulation of the action control system by social intention: Unexpected social requests override preplanned action. *Journal of Experimental Psychology. Human Perception and Performance, 35*(5), 1490–1500.

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition, 119*(2), 242–252.

Seegelke, C., Hughes, C. M. L., Schütz, C., & Schack, T. (2012). Individual differences in motor planning during a multi-segment object manipulation task. *Experimental Brain Research, 222*(1–2), 125–136.

Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *The Journal of Neuroscience, 14*(5), 3208–3224.

Short, M. W., & Cauraugh, J. H. (1997). Planning macroscopic aspects of manual control: End-state comfort and point-of-change effects. *Acta Psychologica, 96*(1), 133–147.

Sirigu, A., Cohen, L., Duhamel, J. R., & Pillon, B. (1995). Congruent unilateral impairments for real and imagined hand movements. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 6*(7), 997–1001.

Sisbot, E. A., Marin-Urias, L. F., Broquere, X., Sidobre, D., & Alami, R. (2010). Synthesizing robot motions adapted to human presence. *International Journal of Social Robotics, 2*(3), 329–343.

Stapel, J. C., Hunnius, S., & Bekkering, H. (2012). Online prediction of others' actions: The contribution of the target object, action context and movement kinematics. *Psychological Research*, 1–12.

Stephan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., & Frackowiak, R. S. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *Journal of Neurophysiology, 73*(1), 373–386.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience, 7*(9), 907–915.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience, 5*(11), 1226–1235.

Urgesi, C., Candidi, M., Fabbro, F., Romani, M., & Aglioti, S. M. (2006). Motor facilitation during action observation: Topographic mapping of the target muscle and influence of the onlooker's posture. *European Journal of Neuroscience, 23*(9), 2522–2530.

Van Beers, R. J., Haggard, P., & Wolpert, D. M. (2004). The role of execution noise in movement variability. *Journal of Neurophysiology, 91*(2), 1050–1063.

Viviani, P. (2002). Motor competence in the perception of dynamic events. *Common Mechanisms in Perception and Action: Attention and Performance, 19*, 406–442.

Weiss, P., Jeannerod, M., Paulignan, Y., & Freund, H.-J. (2000). Is the organisation of goal-directed action modality specific? A common temporal structure. *Neuropsychologia, 38*(8), 1136–1147.

Zijdewind, I., Toering, S. T., Bessem, B., Van der Laan, O., & Diercks, R. L. (2003). Effects of imagery motor training on torque production of ankle plantar flexor muscles. *Muscle & Nerve, 28*(2), 168–173.

About the Authors

Daniel Lewkowicz received an MS in neuroscience from the University of Toulouse III for a work on single-cell recordings in primary visual cortex V1 of macaque monkey. He is currently a PhD student at the University of Lille Nord de France and participating in the interdisciplinary project INTERACT that aims to build a humanoid robot that can sustain intuitive interaction through observation and imitation. His main research interests focus on the mathematical description of biological movements to better understand the functional association between kinematics and intentionality.

Yvonne Delevoye-Turrell received a PhD in behavioral neurosciences from the University of Birmingham (UK) in 2001. Her work has since focused on the study of the cognitive aspects of rhythmic interactions within the theoretical framework of embodiment. From 2001 to 2009, she has worked in the field of schizophrenia to study the behavior associations between motor planning, motor fluency and the pre-conscious experience of the sense of agency. She is now professor at the University of Lille Nord de France. Her research interests are focused on the study of the timing properties of motor planning and to assess how cognitive functions (attention, spatial memory, intentionality) can modulate the experiences of our body in action.

David Bailly received his undergraduate degree in engineering from the ENSEA school in 2007. He then went on to obtain an MS in intelligence systems and neural networks at the University of Cergy-Pontoise, France. Finishing a PhD thesis in the area of bio-inspired robotics, he is specializing in the fields of vision and motor control. He is now participating in the project INTERACT, conducting experimental studies to model the activity of those specific cerebral structures important for the visual control of action using emulated neural networks and robots.

Pierre Andry received an MS in artificial intelligence from the University of Paris VI (Pierre et Marie Curie) in 1999, and a PhD in computer science investigating the role of imitation in learning and communication skills of autonomous robots in 2002. He is now a researcher at the ETIS lab, University of Cergy-Pontoise. His main topics of interest are epigenetic robotics, i.e., the way autonomous robots develop in order to adapt to their physical and social environment. This concern issues such as imitation, emotions, pre-verbal interactions processes, turn-taking and emergent behaviors.

Philippe Gaussier received a PhD in computer science from the University of Paris XI (Orsay) for work on the modeling and simulation of a visual system inspired by mammalian vision. From 1992 to 1994, he conducted research in neural network (NN) applications and in the control of autonomous mobile robots at the Swiss Federal Institute of Technology. He has edited a special issue of the journal *Robotics and Autonomous Systems* on 'moving the frontier between robotics and biology'. He is now professor at the Cergy-Pontoise University in France and leads the neuro-cybernetic team of the image and signal processing laboratory (ETIS). His research interests are focused on the modeling of the cognitive mechanisms involved in visual perception, motivated navigation and action selection. He is also interested in the study of the dynamical interactions between individuals with a particular research in the fields of imitation and emotions.