

“You never fail to surprise me”: the hallmark of the Other

Experimental study and simulations of perceptual crossing

Charles Lenay¹, John Stewart¹, Marieke Rohde² & Amal Ali Amar¹

¹Centre Pierre Guillaumat, France / ²MPI for Biological Cybernetics, Germany

Classically, the question of recognizing another subject is posed unilaterally, in terms of the observed behaviour of the other entity. Here, we propose an alternative, based on the emergent patterns of activity resulting from the interaction of both partners. We employ a minimalist device which forces the subjects to externalize their perceptual activity as trajectories which can be observed and recorded; the results show that subjects do identify the situation of perceptual crossing with their partner. The interpretation of the results is guided by comparable evolutionary robotics simulations. There are two components to subjects' recognition capacities: distinguishing mobile and fixed entities; and behaving so as to interact with their partner rather than with a mobile lure. The “Other” is characterized by the feature that there is sufficient regularity in the interactions to encourage the formation of anticipations; but sufficient indetermination that the actual behaviour is consistently surprising.

Keywords: Recognition of other; perceptual crossing; evolutionary robotics

1. Introduction

We use the term $\langle\langle$ perceptual crossing $\rangle\rangle$ to designate the situation where two perceptual activities of the same kind meet each other, as when two people catch each others' eye, mutual touch, and generally kinesthetic or acoustic interactions (proto-conversation with babies, dialogue, choral singing etc.) (Argyle & Cook 1976; Bailenson et al. 2001). Perceptual crossing is a very frequent kind of event in the context of encounters in everyday life, in childhood from the earliest age of post-natal life, in many animal species and even across species. It is known that

these situations play an important role in social regulation (Argyle & Dean 1965), that they are crucial in child development (Murray & Trevarthen 1985; Reddy 2003; Fogel 2006), and that they carry strong emotional qualities (Nadel & Muir 2005; Fogel & Garvey 2007). Furthermore, if we take into account the diversity of technological systems which allow for real-time interaction, this theme of “perceptual crossing” can be further generalized to include all those situations where, by means of a common mediation, each participant is able to perceive the perceptual activity of their partner.

It is a major question in cognitive science to understand how it is possible to distinguish between the presence of another subject and that of any other object or programme. In this paper, we aim to explore the role of perceptual crossing in enabling subjects to make such distinctions in a two-dimensional simulated environment. Our focus is on the embodied dynamics of interaction, on how it leads to the detection of agency, and how it compares to robotic agents evolved for the same behaviour.

Classically, the problem of the recognition of another person is brought back to the question of adopting an “intentional stance” (Dennett 1987); the question is then to identify the criteria used by subjects in deciding if the perceived entity is intentional. There are then several possibilities. One is to adopt an elaborate cognitive plane, and to refer to behaviour that can be considered as “rational” (Csibra et al. 1999). Another is to consider that this recognition arises from modular mechanisms directly linked to perception, with “innate” upstream filters which function quasi-automatically to decide whether the object is “self-propelled” (Premack 1990). Other approaches relate to the striking capacities for imitation displayed by new-born infants (Meltzoff & Decety 2003). The capacity for imitation is sometimes attributed to “mirror neurons” which fire both when an animal (or human) acts and when the animal observes the same action performed by another (Rizzolatti & Craighero 2004; Keysers 2010). Thus, the neuron “mirrors” the behaviour of the other, as though the observer were itself acting. However, despite the excitement generated by these findings, to date no widely accepted models have been put forward to describe how mirror neuron activity supports cognitive functions such as imitation (Dinstein et al. 2008). The general point we wish to make is that in spite of their diversity and points on which they disagree, all these approaches have an important feature in common: the question of recognition is posed to only one of the partners. The perceptual situation is asymmetrical: there is an observer and the other entity which is observed. Thus, all these theories suppose that recognition of the other is based on a sufficiently precise perception of his behaviour (Thomas, Press, Haggard 2006), without taking into account the active engagement of the observer in the situation.

Here, by contrast, we wish to explore the possibility that an essentially symmetrical situation, in particular that of perceptual crossing, makes it possible for the two subjects to recognize each other mutually. This approach is interactionist because we do not focus on the stimuli-response characteristics of one agent, but on the reciprocal recognition of each other. The collective dynamics of the two perceptual activities mutually influence each other continually in closed-loop interaction, with the result that the interaction is not controlled by either of the individual participants alone. In such an interactionist view, it is possible that recognition of another subject is not derived cognitively by either of the participants but emerges from the interaction dynamics. In order to test this hypothesis, we have already carried out an initial study in deliberately minimalist one-dimensional environment (Auvray, Lenay & Stewart 2009) in which two blindfolded subjects had to distinguish between interactions with a fixed lure, a mobile lure, and the receptor field of their partner. The results showed that a clear majority of clicks were produced when the two participants were effectively in front of each other, i.e. in a situation of perceptual crossing (62%); this is significantly much greater than the chance level of 33% ($\chi^2_1 = 50$, $p < 0.01\%$).

These results were sufficiently encouraging to lead to the study presented here. In order to better understand this recognition of perceptual crossing and to verify the generality of the capacity of subjects to privilege the face-to-face situation, we have therefore carried out a new experiment, this time in a 2-dimensional space. This situation, which is closer to that in video games and other virtual meeting-places, greatly increases the degrees of freedom for exploration and the corresponding perceptual strategies. In these conditions, is it still possible for subjects to succeed in the task of recognizing encounters with the other subject? And if so, will the analysis of the strategies employed provide evidence for a dynamic attractor for the perceptual activities, as was found previously in the one-dimensional version of the experiment (Auvray, Lenay & Stewart 2009)?

The interactionist approach taken here is an embodied and dynamical approach in the sense that it does not presume explicit manipulation of representational symbols inside an agent's brain in order to achieve a certain goal. Instead, much simpler circuits can bring about the same behaviour, exploiting ecological factors and the dynamics of the closed-loop interaction with other agents and with the environment. Evolutionary robotics (e.g. Nolfi & Floreano 2000; Harvey et al. 2005) is a simulation modelling technique that makes it possible to explore the space of possible solutions in closed loop interaction without making strong assumptions about the underlying circuitry. In order to study the principles of interaction in the virtual task environment in silicio, we simulated the task posed to the human subjects using such an Evolutionary Robotics approach (Di Paolo et al. 2008; Rohde & Di Paolo 2008; Rohde 2010).

We compare three different agent bodies in simulation: a wheeled agent, an agent that generates a velocity vector in Euclidean space and a simulated arm. By comparing the solutions evolved for different bodies, we can identify general principles and body specific factors in possible solutions. This exploration in simulation guides the analysis of the human data, pointing out possible factors and variables of interest; for instance, whether human motor control in this task is dominated by properties of the body and similar to the arm, or whether it is more world-centred and similar to the Euclidean agent. If humans draw heavily on cognitive and symbolic capacities in this task, we would expect them to behave in altogether a different fashion from such simple agents.

2. Materials and methods

2.1 Experimental study in humans

10 pairs of two subjects (mean age 28.9, 16 males, and 4 females,) participated in the study. The experiment was conducted using a tactile feedback platform TACTOS (Hanneton et al. 1999; Sribunruangrit et al. 2002). Using the computer mouse with their dominant hand, participants move a receptor field in a 2-dimensional virtual space, and receive tactile sensory feedback on the index finger tip of the other hand from 2 commercially available Braille display units (1 cm^2 , with 2×8 computer-controlled pins). The size of the receptor field was 24×24 units, and the whole virtual space was 200×200 units (pixels), which corresponds approximately to 20 cm^2 of mouse movement. In order to avoid complications from encounters with boundaries, the edges of the space were joined in both dimensions so that topologically the space was a torus. Whenever an object was encountered, the subject received a tactile stimulation on the index finger of the other hand. Three types of object (each 24×24 units in size) could be encountered: (i) the receptor field of the other subject; (ii) a fixed lure; and (iii) a mobile lure which, unknown to the subjects, was actually attached to the receptor field of the other subject at a fixed distance (28 pixels in x and y), so that it had exactly the same movements as the receptor field (Figure 1). The mobile lures were arranged so that it was impossible for both participants to interact with the other participant's lure at the same time. The fixed lures (same surface area) were placed with their centre points at (155,155) for subject1 and at (45,45) for subject 2. The idea behind this experimental setup is that the mobile lure and the receptor field of the partner should have exactly the same objective behaviour (Figure 1). The aim here is to examine whether the subjects are nevertheless able to discriminate the presence of the receptor field of their partner, i.e. the situation of perceptual crossing which corresponds to the moment

when the receptor fields of the two subjects cross each other so that they perceive each other mutually.

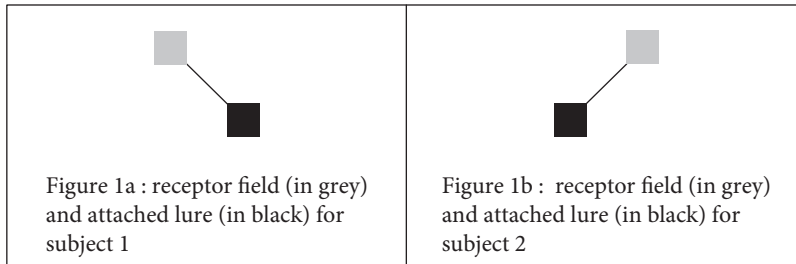


Figure 1. The receptor field and the fixed lure are each 24×24 pixels

Procedure: Subjects from each pair were blindfolded and placed in separate rooms so that their tactile encounters in the virtual space were the only form of communication. Subjects were told that they could encounter one of three possible objects in the virtual environment: the other participant, a mobile object and a fixed object. Subjects were unaware that the mobile object was attached to the other participant. The task for the subjects was to click on the left mouse button whenever they considered that they were in contact with the other subject. Each pair of subjects was observed in three blocks of 4 minutes interaction.

2.2 Simulation model

We modelled the task posed to the human participants using an Evolutionary Robotics approach (Harvey et al. 2005; Rohde 2010). In Evolutionary Robotics, what is selected is not a control function or an input-output mapping, but instead embodied behaviour in a closed sensorimotor loop. This means that the circuits underlying the required behaviour are underspecified and that the genetic algorithm can exploit non-linear or non-obvious properties of the interaction with the environment or other agents. A plurality of possible strategies that realize one and the same task are generated, and these can be tested against the human data. Many of these solutions rely on dynamical subtleties and run counter to human intuitions. The objective of the Evolutionary Robotics simulation in this study is to catalyze the generation of hypotheses through comparative analysis, as if one were dealing with different species. This kind of modelling is more conceptual or theory-driven (cf. Rohde 2010) than the typical data-driven modelling in psychology, where the objective is to fit and describe a data set with the smallest possible number of parameters: the simulation instead guides the analysis of human behaviour, as was the case for the previous 1-dimensional version of the task (Di Paolo, Rohde & Iizuka 2008).

The task was very similar to the human task. The artificial agents were evolved to move around a virtual plane that wraps around in both dimensions. All objects in this plane are circular and have a diameter of 20 units (see Figure 2). The attached lures shadow the trajectories of each of the agents, and the only sensory signal that the agents receive is a touch signal SG. The “dark” agent can perceive all “light” objects, but not the “dark” ones, and vice versa. Technical details and a full analysis of the results have been published in Rohde and Di Paolo (2008) to which the reader is referred for more details.

The performance criterion (fitness function) in this case was to find a source of stimulation and correctly classify whether it is another agent (a clone of the agent itself) or an inanimate object. One difference from the human scenario was that agents only had to make that decision at the end of a trial, not continually while moving around. The other difference was that there was no learning for the humans, since the complete set of experimental sessions are performed without any feedback on whether the clicks are correct or not.

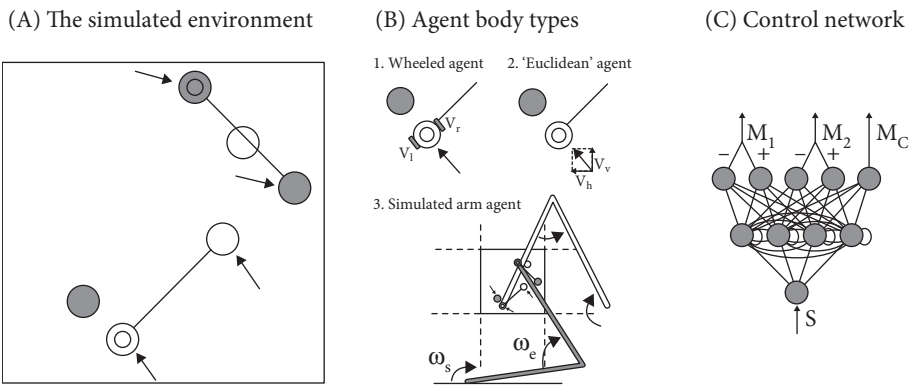


Figure 2. Schematic diagram of the simulation model and control network. (A) The simulated environment with the two agents (double circles), the attached lures (circles attached with a line) and the fixed lures (circles). (B) The different agent bodies evolved. (C) The control network

In order to explore the role of embodiment and sensorimotor dynamics on solutions to the task, three different types of agents were evolved:

1. A two-wheeled agent that generates a velocity signal for each wheel (Figure 2B1)
2. An agent that generates a velocity vector in Euclidean space, like a joystick (Figure 2B2).
3. A simple simulated arm with two segments of length 400 units that is steered through angular velocity signals to the elbow and the shoulder joint (Figure 2B3).

This comparison of different body types had a two-fold objective. Firstly, by comparing the strategies evolved for different bodies, it is possible to examine to what extent this task dictates solutions or dynamical interactions that evolve robustly even across a wide diversity of situations. Secondly, by comparing the evolved behaviour of the robots to the behaviour of human participants, it is possible to examine whether the strategies that evolve across different agent bodies also hold for the human example, or whether humans share body-specific properties of evolved solutions. This comparison does not assume that humans are as simple as the evolved agents in performing this task; the agents have no intention to interact, and no concept of “the other”. The comparison serves to show to what extent humans employ simple dynamic interaction strategies in finding one another that do not require symbolic capacities or intentionality. On the other hand, if the human behaviour were to be different from that of the robots, this would point towards more complex strategies or task-external factors in humans.

For purposes of comparison, all three kinds of agent are controlled by structurally identical continuous-time recurrent neural networks (Di Paolo et al. 2008). The networks are fed a signal if the agent touches something (with a 100 ms transmission delay) and generate two motor signals that are implemented according to the agent’s body type. Network parameters (synaptic weights, time constants, biases, sensory and motor gains) are evolved using the evolutionary search procedure. Agents are evolved only to interact with clones of themselves, not across agent bodies. For each body type, the procedure has been initialized with ten different seeds to evolve a population of 30 agents across 1000 generations.

3. Results

3.1 Human performance

We first categorized the stimuli received according to their origin: fixed lure, mobile lure, or receptor field of the partner. We then categorized the clicks according to the most frequent type of stimulus during the two previous seconds.

Table 1. The percentages of clicks, percentage stimulations, and ration clicks/stimuli, in each of the three situations PC (Perceptual Crossing), FL (Fixed Lure) and ML (Mobile Lure)

	Perceptual Crossing	Fixed Lure	Mobile Lure
% clicks	63,4	13,2	23,5
% stimulations	50,4	31,4	18,3
Ratio clicks/stim	1,26	0,42	1,28

Overall, the subjects are quite successful: clicks in the situation of perceptual crossing, when the partners are effectively in contact with each other, are much more frequent than clicks on either the fixed or mobile lures. The ratio clicks/stimuli was significantly lower for clicks on the fixed lure (0.42), than clicks on the mobile lure or perceptual crossing (1.28 and 1.26 respectively) ($\chi^2_1 = 30$, $p < 0.01\%$). However, the subjects did not seem to distinguish between stimuli due to perceptual crossing and those due to the mobile lure (1.26 vs. 1.28). ($\chi^2_1 = 0.1$, $p > 50\%$). Thus, the difference between clicks on the mobile lure and on the receptor field of the partner (23.5 % vs. 63.4 %) seems to be due to strategies of displacement such that stimuli due to the mobile lure were less frequent than those due to perceptual crossing (18.3 % vs. 50.4 %).

Thus, at an initial descriptive level, it appears that the success of participants has two components. Firstly, the behavioural strategies are such that subjects are much more frequently in contact with each other (Perceptual Crossing, PC) than with the Mobile Lure (ML). Secondly, the subjects are able to distinguish the Fixed Lure (FL) from both the mobile stimuli, as evidenced by the lower click-rate in line three of the Table 1. This basic result is very similar to that obtained in the initial 1-dimensional study (Auvray et al. 2009).

3.2 Artificial agent performance and behaviour

Contrary to the human subjects, the artificial agents that were evolved did not make any perceptual discriminations in the task. Instead, they evolved to constantly reply “yes” or “no”, making behavioural distinctions between different kinds of objects. Given that agents were only evaluated on their output at the end of a trial, it was a perfectly viable behaviour to localise the fixed lure and stay with it, avoiding both the attached lure and the other agent; a constant output of “no” then yielded perfect fitness (up to 1.0). This strategy will be referred to as “no-interaction” throughout this paper.¹

Less successful agents searched the two-dimensional space for any source of stimulation and tried to stay in its vicinity. Unable to discriminate between live interaction and inanimate objects, these agents produced a constant “yes” output, which is on average more successful than a constant “no” since their trajectories allowed them to receive more stimulations due to the presence of the other subjects than those due to lures (average fitness of up to 0.4). Given that no other solutions were evolved, irrespective of agent body, these two types of behaviour seem to be locally optimal strategies for the task as it was posed to the artificial agents.

Comparing these evolved solutions to the task, three observations are worth mentioning.

1. It is remarkable that agents who manage to localise objects never evolve to recognise patterns in the input (i.e. “yes”/“no” answers contingent on source of stimulation), but instead always distinguish different objects behaviourally. Thus, active perceptual strategies seem more efficient than parsing and classifying inputs and point towards a potential relevance of closed-loop sensorimotor invariants in the task.
2. Given that “no interaction” behaviour was more efficient in the case of the artificial agents, we conclude that the explicit instructions for the robots and humans masked an implicit assumption: humans, but not necessarily the robots, actively seek the presence of other intentional beings.
3. Figure 3 shows how a “no interaction” agent *avoids* the shadow of the other agent on the basis of irregularity in the interaction. Agents – or humans – could use the same cues, i.e. irregularity to identify the other, if selection pressure favoured this kind of behaviour.

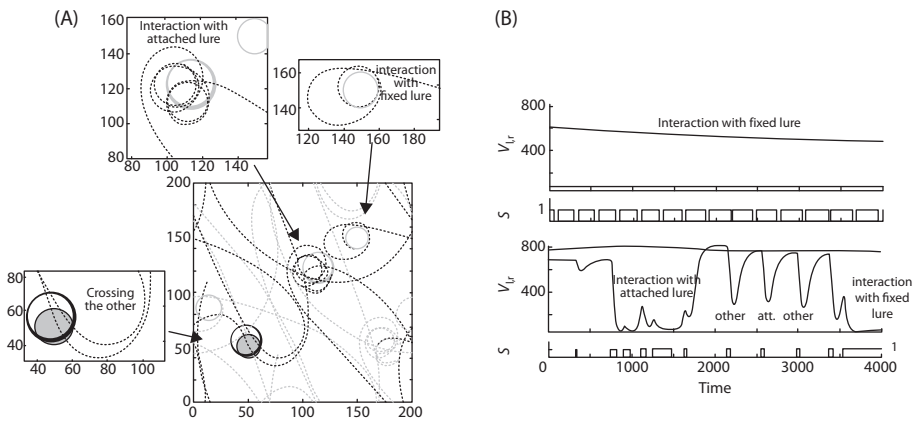


Figure 3. Example behaviour of a pair of wheeled agents evolved: (A) Trajectories of the interaction. (B) Velocity outputs and sensory inputs of the same behaviour (top agent 1, bottom agent 2). Agent 1 (solid line in (A)) finds the fixed lure immediately. Agent 2 (dotted line) mistakes the attached lure for the fixed object. However, registering irregularities in the interaction (cf. (B)), Agent 2 eventually moves away and finds the fixed object and remains there ($\llcorner$ no $\gg$ live interaction is the correct answer)

Confirming the results from the one-dimensional variant of the task, oscillatory interaction appears to provide a behavioural advantage. As the frequency of oscillations in simulations is driven by sensory latencies, the simulation predicts such a dependency for humans, too. The comparative evolution of different agent bodies predicts that, due to arm morphology, such oscillations

should be along one dimension at an angle perpendicular to the lower arm in the humans.

4. Analysis of human trajectories

In 3.1. we noted that human subjects are (a) more frequently in contact with each other (PC) than with the Mobile Lure (ML); and (b) that they distinguish the Fixed Lure (FL) from both the mobile stimuli. We now turn to an analysis of their trajectories in order to identify the strategies that enable them to do this.

4.1 The maintenance of contact: Oscillations

Figure 4 confirms that human subjects do tend to remain in the vicinity of each other. There is a clear concentration in the zone $[-25 +25]$, $[-25 +25]$ corresponding to Perceptual Crossing, and a subsidiary concentration around $[-75 -25]$, $[-75 -25]$ corresponding to the Mobile Lure. Overall, the rate of stimulation is 37.9%, which is substantially greater than the 24.5% expected on the basis of surface areas alone.

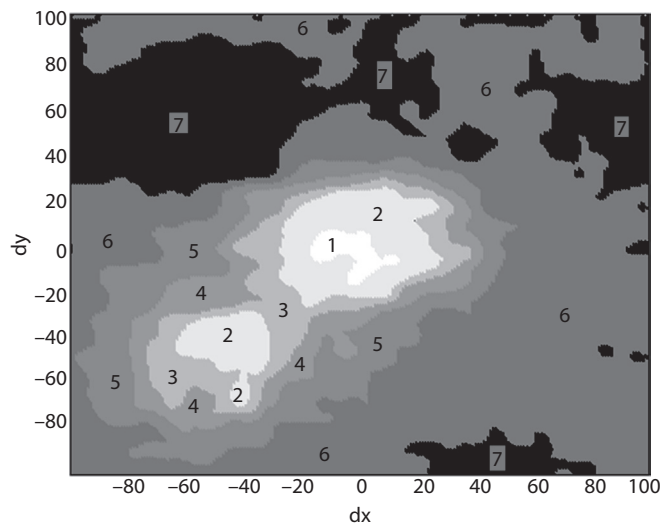


Figure 4. Contours of relative frequencies of positions in the $dx - dy$ plane, where dx is the distance in pixels between the positions of the two subjects on the X-axis, and dy the distance on the Y-axis. Number code for regions: $>3.5 = '1'$; $>2.5 = '2'$; $>2.0 = '3'$; $>1.5 = '4'$; $>1.0 = '5'$; $>0.5 = '6'$; $>0 = '7'$

What are the strategies which enable subjects to maintain contact with sources of stimulation? In the simpler 1-dimensional version of the experiment (Auvray et al. 2009) and in its robot model (Di Paolo et al. 2008), agents tended to remain in anti-phase oscillation with a partner once interaction had been established. In the 2D simulation model, such oscillatory interaction evolved in some but not in all agents. Freezing, once a source of stimulation is identified, is an alternative to rhythmic oscillatory interaction. However, rhythmical interaction was more successful; the reason is that by staying in rhythmic interaction with the boundary of an object, the agent can gain knowledge of the relative spatial position of the object; whereas by freezing, the agent knows only that the object is there, not where it is. This benefit was valid for all agent bodies.

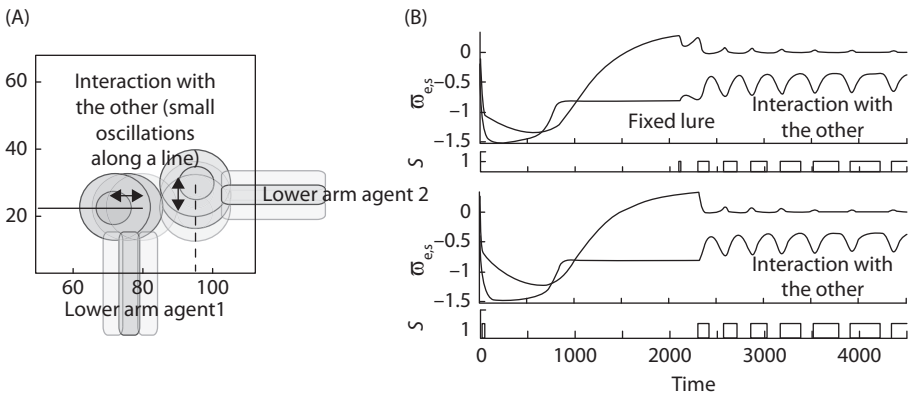


Figure 5. Example of trajectory (A) and motor outputs/sensory inputs of an oscillatory interaction in the simulated arm agent. Oscillation occurs perpendicular to the lower arm, due to oscillation in one of the motors, but not the other

In the simulations, the oscillations were constrained by bodily factors (see Figure 5). Firstly, the frequency of oscillations depended on the inbuilt sensorimotor latencies. Secondly, the simulation of a two-joint arm showed that oscillations were always along a line perpendicular to the lower arm. Agents were generally found to stabilise rhythmic interaction by freezing one of their two degrees of freedom (one wheel, one direction, one joint) and using the other motor output to maintain rhythmic interaction with the boundary of the other agent.

Oscillatory behaviour also occurred in the humans, although there were large individual differences. In order to characterize these oscillations, we have performed Fourier analyses of the time-series in both the X- and Y- axes. Examples of these analyses are shown in Figures 6 and 7.

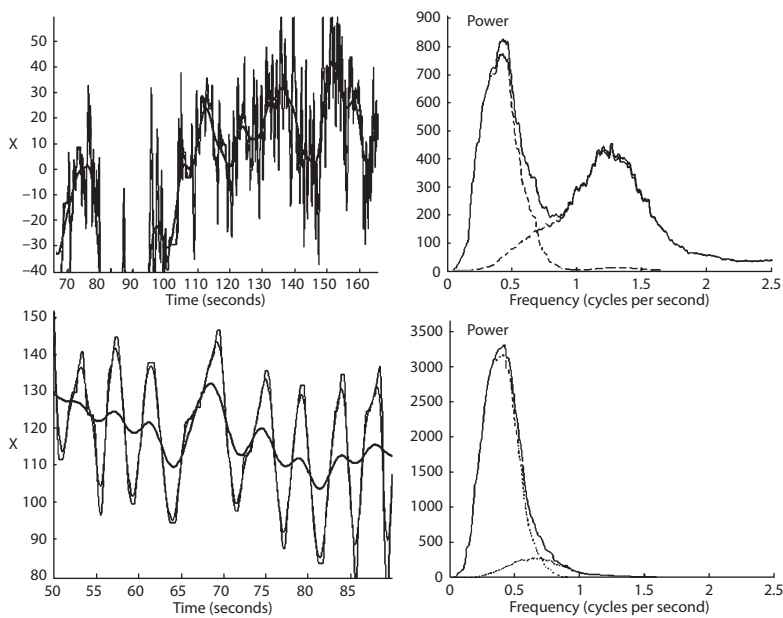


Figure 6. Two examples of time-series (on the left) and the corresponding Fourier analysis (on the right). For the upper figures, there are clearly two fairly distinct components, one with a frequency of around 1.1 cps, and the other with a frequency of 0.4 cps. In the lower figures, the high-frequency component is practically absent, and the pure medium-frequency oscillations are directly apparent

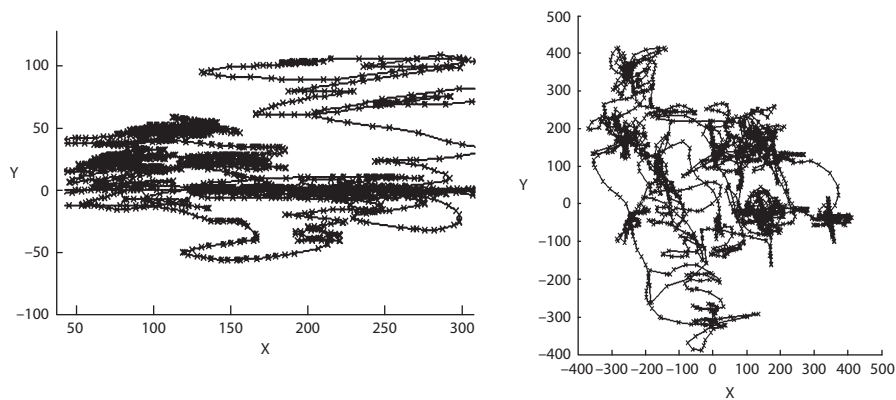


Figure 7. In the trajectory shown on the left (note difference in scale on both axes), the oscillations are restricted to the horizontal X-axis; on the Y-axis this subject only drifts slowly. By contrast, in the trajectory shown on the right, the oscillatory movements occur in both the X- and Y-axes, so that there is no preferred direction of the movements which are sometimes circular

The major characteristics of the oscillations can be described by 3 parameters derived from these Fourier analyses:

- a. The total power, corresponding to the amplitude of the oscillations (TP).
- b. The proportion of this power in the Y-dimension (PY). This parameter corresponds to the distribution of orientations of the oscillations; the value of PY is low for the generally horizontal oscillations on the left of Figure 7, whereas it is high for the trajectory on the right.
- c. The proportion of power in the high-frequency oscillations (P-Hfr). This parameter is relatively high for the trajectory shown in the upper part of Figure 6, whereas it is low for the lower part.

Table 2. The mean, range, and standard deviation expressed as a percentage of the mean, for the three parameters “total power” (TP), “proportion of power in the Y-dimension” (PY) and “proportion of power in high-frequency oscillations” (P-Hfr). The fourth line shows the F-statistic for the ratio of between-individual to within-individual variance, based on the 3 separate sessions for each subject. All three values are statistically highly significant at less than 0.1%

	Parameter		
	TP	PY	P-Hfr
Mean	23.6	0.36	0.33
Range	12.5–37.0	0.05–0.67	0.11–0.56
SD%	30%	44%	42%
F19,40	6.43	14.9	13.4

As is shown in Table 2, individual subjects vary widely for each of these parameters. These results show that the characteristic nature of the oscillations is consistent within each subject, thus confirming that the oscillations are real phenomena and are not just due to random movements.

In the artificial agent simulations, the frequency of oscillation is related to sensorimotor latencies which do not vary substantially between humans. Since the values for P-Hfr in humans do vary considerably, this indicates that here the oscillations are deliberate and not simply the mechanical result of sensorimotor latencies. The simulations also predict that oscillations should be one-dimensional and perpendicular to the lower arm. This prediction is difficult to verify directly, since postural information was not recorded. However, the fact that humans remain rather constant in the orientation of their oscillatory interaction (low variance for PY within subjects), and a general trend towards horizontal oscillations, is in accordance with this prediction.

Another prediction derived from the model was that these oscillations are used functionally in order to maintain contact with sources of stimulation. This prediction is confirmed by the following analysis.

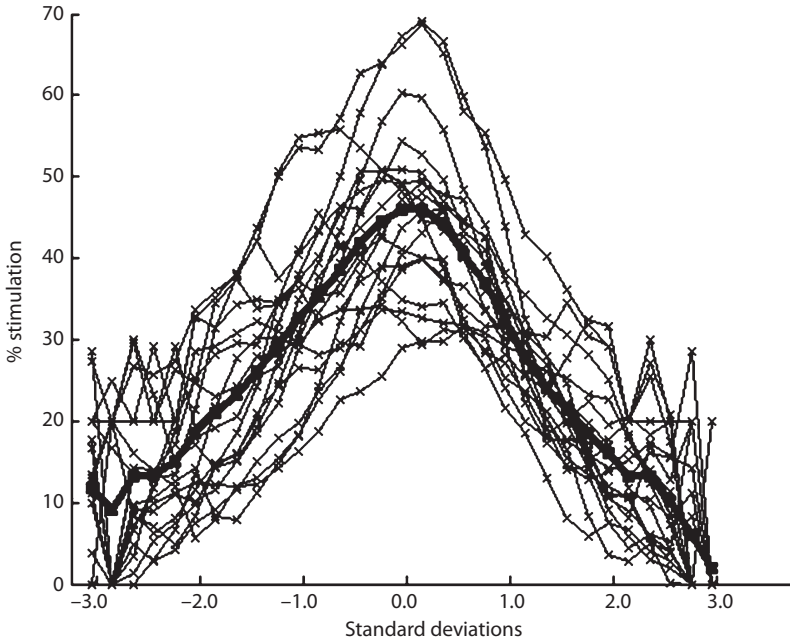


Figure 8. The rate of stimulation, plotted as a function of the distance from the centre of the oscillations. Black crosses and thin lines: curves for individual subjects. Heavy black line: mean of 20 subjects

Figure 8 shows that overall, the oscillations are centred on sources of stimulation (although to varying degrees according to the subjects, a point to which we shall return). At the tails of the distribution shown in Figure 8, the rate of stimulation drops off, suggesting that when contact is lost the subjects use this functionally as an indication to turn back again towards the source of stimulation. The existence of such behaviour is confirmed by the pattern shown in Figure 9, which shows that some subjects do indeed inverse their trajectory when contact with the source of stimulation is lost.

We have noted above (Section 3.1) that the overall average rate of stimulation is 37.9%. However there are individual variations: the average per subject ranges from 27.3% to 55.0%, with a standard deviation of 6.9%. As shown in Figure 10, these variations correlate strongly both with the rate of stimulation at the centre of oscillations (cf Figure 8, $r = 0.75$); and with the rate of stimulation recovered 1 second after loss of contact (cf Figure 9, $r = 0.82$). The functional significance of

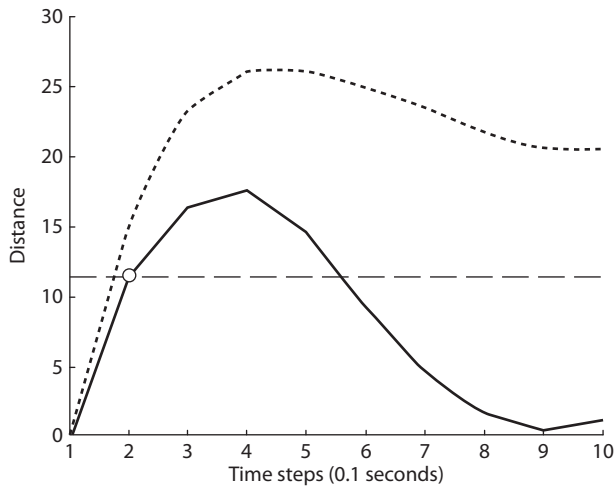


Figure 9. Behaviour after losing contact (mean of all values for one particular subject). Heavy black line: at time step 1 there is a stimulation; at time step 2, marked by a small circle, this contact is lost. The thin dashed horizontal line shows the distance at which contact was lost. The subsequent time steps show that this subject reverses direction, and returns to the position at which contact was last obtained. Heavy dashed line: for comparison, here there is no stimulation at time-step 1, but stimulation occurs at time-step 2. In this case, the subject does not return, but (on the average) remains in the zone where contact was found

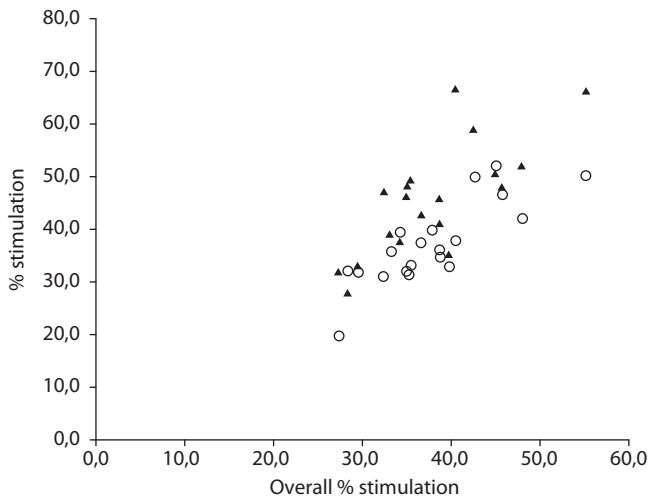


Figure 10. Triangles: rate of stimulation at the centre of oscillations. Circles: rate of stimulation recovered 1 second after loss of contact (see Figures 8 and 9). Both of these rates are plotted as a function of the overall rate of stimulation, for each of the 20 subjects

the oscillations, as discussed here, does seem to be a rather generic feature of the “perceptual crossing” situation, occurring in both the robotics simulations, and the previous study (Auvray et al. 2009).

In order to confirm that the simple mechanism of turning back when contact is lost can account both for the oscillations and for their effectiveness in maintaining contact, we have calculated the sort of trajectory that would be produced. The basis of these calculations is the following: as long as there is no change in the stimulus (i.e. either no contact, or constant contact), the speed remains constant in both X and Y dimensions. When contact is lost (cf. Figure 9), there is an appropriate acceleration in the direction opposite to the current velocity, which continues either until contact is re-established or until a maximum return velocity is reached. Finally, a certain level of noise is added to the current velocity. As illustrated in Figure 11, this simple mechanism is indeed sufficient to produce oscillations similar to those described above.

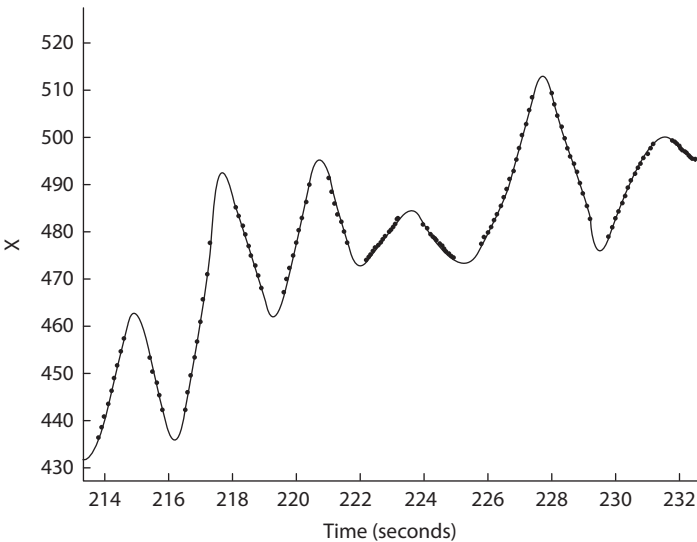


Figure 11. Positions on the X-axis plotted as a function of time. Dots: stimulation received. Oscillations very similar to those in Figure 6 are clearly apparent

In these calculations, the proportions of the 3 types of stimulation – Perceptual Crossing, Fixed Lure and Mobile Lure – are 53.6%, 27.3% and 19.1%. These values are very close indeed to those observed empirically for the human subjects: 50.3%, 31.2% and 18.5% (cf. Table 1). An interesting feature of these calculations is that even without any specific mechanism for leaving contact with the Fixed Lure, there are more stimulations due to contact with the other partner than with

the Fixed Lure. Of course, the actual mechanisms employed by human subjects are likely to be somewhat different; but this result shows that the situation of Perceptual Crossing constitutes a strong dynamic attractor, in a rather generic sense.

4.2 Clicks on mobile stimuli

We come now to the second point mentioned at the end of Section 3.1, i.e. the ability to distinguish the Fixed Lure from both the mobile stimuli, as evidenced by a lower click-rate (Table 1).

We will start by noting a very simple strategy, which consists of just remaining completely stationary, in both the X and Y dimensions, for periods ranging from a few time-steps to several seconds. This behaviour appears to be fairly general, since it has also been observed in the robotics simulations. The usefulness of this tactic (particularly when employed in combination with behaviours aiming at maintaining contact, as above) is fairly obvious: if a change of stimulation (in either direction, 0–1 or 1–0) occurs while the subject is completely stationary, this is a clear indication that the contact occurred with a mobile object.

The proportion of all time-steps for which the subject is completely stationary varies according to the individual subject, ranging from 16% to 57% with a mean value of 37%. However, before developing further statistics on this behaviour, we shall first generalize it to the notion of “surprise”.

We propose the more general hypothesis that the subjects form anticipations, based on what would be expected if stimulation were due to contact with a fixed object. If a pattern of stimulation occurs which is not in accordance with such anticipations, we denote the event as a “surprise”. Operationally, the procedure employed to test for “surprises” is illustrated in Figure 12. Consider all the points which gave rise to a stimulation. Two vertical lines are drawn, through the points with the highest and lowest values of X; two horizontal lines are drawn, through the points with the highest and lowest values of Y; these lines form a rectangle. If all the points within this rectangle are crosses, i.e. did effectively give rise to a stimulation, this is compatible with the source of stimulation being stationary, so there are no « surprises » (as in Figure 12 left). If however the rectangle also includes some points which did not give rise to a stimulation (as in Figure 12 right), this indicates that the source of stimulation must have moved, and constitutes a “surprise”.

This procedure generalises the tactic of remaining stationary, identified above. If the same point, without movement, gives rise both to a stimulus and to an absence of stimulus, then the rectangle around this point will also include a point without stimulus. On this basis, we may now examine what would happen if subjects clicked each time a “surprise” occurred. The results are the following: PC 58,3%; FL 13,5%; ML 28,1%. These figures are gratifyingly close to those

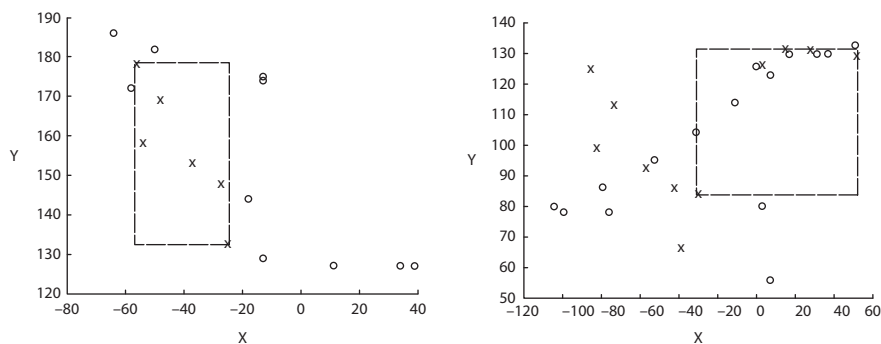


Figure 12. Time period: 20 time-steps, i.e. 2 seconds. Crosses: points that gave rise to a stimulation. Circles: points that did not give rise to a stimulation. A rectangle is drawn around all the crosses. If all the points within this rectangle are crosses – as is the case in the left figure – then there are no “surprises”. On the right, however, the square enclosing all the crosses also encloses a certain number of circles which did not give rise to a stimulation. This situation constitutes a “surprise”

observed for clicks by the human subjects: PC 63,4% ; FL 13,2% ; ML 23,5% (cf. Table 1). It may nevertheless be asked why clicks based on “surprises” are not perfect, in the sense of avoiding any clicks at all on the Fixed Lure. The main reason is that the attribution of a click to the three categories FL, PC and ML is based on the category of the majority of stimuli over the last 2 seconds. In some cases, the majority of stimuli were Sc1 (FL), but there were a few stimuli Sc2 or Sc3 which were sufficient to produce a “surprise” as defined here. It would of course have been possible to redefine the attribution of a category for clicks, so as to count only those clicks where 100% of the stimuli were indeed due to an encounter with the FL. However in this case, if a hypothetical subject were to click quite randomly the % of clicks on the FL would be much less than the overall percentage of stimuli. This did not seem appropriate. We may therefore conclude, at this stage, that clicks by human subjects are related to the occurrence of “surprises” as defined here.

In spite of this close agreement in the proportions of the three types of click (PC, FL and ML), there is nevertheless a discrepancy which requires comment. In terms of total numbers of clicks, rather than simply the proportions, the actual clicks produced by human subjects are 6 times less frequent than if every “surprise” were to trigger a click. A possible explanation for this is shown in Figure 13, which shows the probability of a click by human subjects as a function of the number of distinct stimuli during the preceding 2 seconds, plotted separately for the cases where the stimuli are due to a mobile object (PC or ML), and to a fixed object (FL).

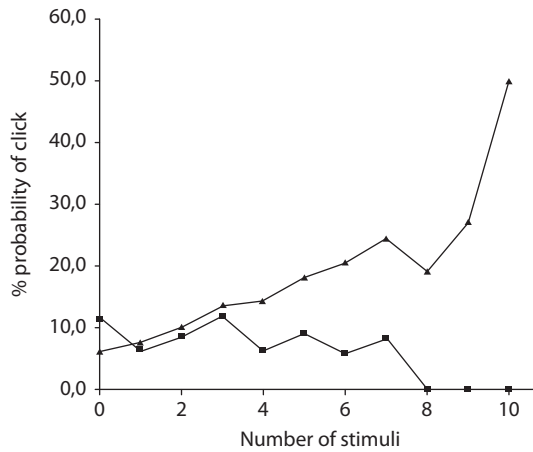


Figure 13. The probability of a click, expressed as a function of the number of distinct stimuli in the preceding 2 seconds. Triangles: stimuli due to Perceptual Crossing or Mobile Lure. Squares: stimuli due to Fixed Lure

As shown in Figure 13, the probability of clicking increases with the number of distinct stimuli received during the preceding 2-second period – but only on condition that the stimuli are indeed due to a mobile object. If the stimuli are due to contact with the Fixed Lure, the probability of clicking actually decreases – confirming again that the subjects do have the capacity to distinguish between contacts with a mobile object and those with the Fixed Lure. Given that this distinction is made, and since it is rather unlikely that multiple successive stimulations would be due to the mobile lure, it is indeed a good strategy to click only when a sufficiently large number of distinct stimulations are produced in the preceding period.

This completes our analysis of the second component mentioned in 3.1., i.e. the capacity of subjects to distinguish between fixed objects and mobile objects.

5. Conclusions

In this experimental study of perceptual crossing, and in the robotics simulations, the sensory data available to each subject at a given point in time are reduced to the strict minimum, i.e. one bit of information: all or none. This sensory input was delivered in the form of a tactile stimulus, but does not depend on any particular sensory modality. By limiting the sensory input to the extreme limit, the perception of an entity can only be achieved through active exploration; the minimalism of the device forces the subjects to externalize their perceptual activity in the form of a trajectory in space and time which can be easily observed and recorded.

These experiments show that subjects succeed in clicking with substantially greater frequency following encounters with the other subject, rather than with a mobile lure or a fixed lure. What is recognized does indeed seem to be the activity of a perceiving subject, and not a simple objective pattern in the movements of the object, because the mobile lure has exactly the same movements as the receptor field of the subject (Lenay 2006; Auvray 2009). This capacity seems to have two components: the capacity to discriminate against the fixed object; and the emergent property that the two perceptual activities attract each other. If the subjects succeed in the perceptual task, it is essentially because they succeed in situating themselves “face-to-face”. The perceptual activities attract each other – just as, in the visual domain, we “catch someone’s eye” i.e. the gazes attract each other. Although the participants do not have any explicit goal of collaborating, their simultaneous efforts to try and discriminate the presence of their partner end up producing a collective dynamic attractor. A similar sort of collective dynamic behaviour was observed in the robotics simulations, which shows that in order for an attractor in the joint dynamics to be established it is not necessary that there be a previous recognition of the other, nor for any complex form of internal representation. Thus, here recognition of the “Other” does not necessarily require complex mental representations such as those involved in the “Theory of Mind”, nor indeed any purely individual cognitive capacities; rather, it is an emergent result of the embodied interaction process itself. At least under the conditions of our experiments, perception of the intentionality of the other is intrinsic to the joint perceptual activity, and is not a later deduction based on analysis of the shape and the movements of the perceived object.

It is not because the subjects recognize each other that they establish perceptual crossing; on the contrary, it is because they establish perceptual crossing that they recognize each other. The criteria employed by the subjects in order to discriminate the presence of the other correspond to a resistance towards any definite spatial determination of the position of the object: another subject distinguishes himself from physical objects by his capacity to continually surprise us (Figure 12). The human subjects seek for regularities in the objects they encounter, notably by reversing their trajectory after they have crossed an object. It is when they fail to establish regularities, in an object which nevertheless maintains its presence, that they conclude they are in the presence of the other subject. Thus, the Other is characterized by the feature that there is sufficient regularity in the interactions to encourage the formation of anticipations; but that once they are formed, these anticipations are regularly flouted. The Other is something which resists any precise spatial determination: it is not a fixed object. However, the Other is also something which maintains its presence. This is indeed a

characteristic of the receptor field of the partner, but not of the mobile lure, since only the receptor field is sensitive to my presence. Now this sensitivity is related to a perceptual intentionality which ceaselessly aims at stabilizing around a singularity. And this is precisely a sufficient condition for the formation of an attractor in the collective dynamics which tends to augment the probability that the partner will be present. Thus, the criteria employed by the human subjects for clicking (a stimulation is encountered which maintains its presence without being spatially determined) are not arbitrary, but follow logically from the encounter between two intentionalities which is characteristic of perceptual crossing. The Other is recognized as that which resists its precise constitution, and which nevertheless persists in being present.

Acknowledgements

The experimental work performed, as well as the data analysis, have been carried out in the Costech laboratory at the Technological University of Compiègne with the financial support of the Regional Council of Picardie. The robotics modelling was supported by the School of Science and Technology at the University of Sussex and by the HFSP grant (2006) on “Mechanisms of associative learning in human perception”.

Note

1. Tongue in cheek, we termed this strategy “autistic” in Rohde & Di Paolo 2008.

References

- Argyle, M. & Dean, J. (1965). Eye-contact, distance and affiliation, *Sociometry*, 28, 289–304.
- Argyle, M. & Cook M. (1976). *Gaze and Mutual Gaze*. London: Cambridge University Press.
- Auvray, M., Lenay, C. & Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, 27, 32–47.
- Bailenson, J.N., Blascovich, J., Beall, A.C. & Loomis, J.M. (2001). Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments, *Presence*, 10, 6, 583–598.
- Beer, R. (1996). Toward the evolution of dynamical neural networks for minimally cognitive behavior. In P. Maes., M. Mataric, J. Meyer, J. Pollack, S. Wilson, (Eds.), *From Animals to Animats 4* (pp. 421–429). Cambridge MA.: MIT Press.
- Csibra, G., Gergely, G., Biro, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of ‘pure reason’ in infancy. *Cognition*, 72, 237–267.
- Dennett, D. (1987). *The Intentional Stance*. Cambridge MA.: MIT press.

- Di Paolo, E., Rohde, M. & Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas in Psychology* 26, 278–294.
- Dinstein I., Thomas C., Behrmann M., & Heeger D.J. (2008). A mirror up to nature. *Current Biology* 18 (1): R13–8.
- Fogel, A. (2006). Dynamic systems research on interindividual communication: The transformation of meaning-making. *The Journal of Developmental Processes*, 1, 7–30.
- Fogel, A. & Garvey, A. (2007). Alive communication, *Infant Behavior and Development*, 30, 231–257.
- Hanneton, S., Gapenne, O., Genouel, C., Lenay, C., & Marque, C. (1999). Dynamics of shape recognition through a minimal visuo-tactile sensory substitution interface, *Third Int. Conf. On Cognitive and Neural Systems*, Boston, 26–29.
- Harvey, I., Di Paolo, E., Wood, R., Quinn, M. & Tuci, E. (2005). Evolutionary Robotics: A new scientific tool for studying cognition. *Artificial Life*, 11, 1–2, 79–98.
- Keyesers, C. (2010). Mirror Neurons. *Current Biology* 19 (21): R971–97.
- Lenay, C., Auvray, M., Sebbah, F.-D., Stewart, J. (2006). Perception of an intentional subject: An enactive approach. In *Proceedings of the Third International Conference on Enactive Interface*, Montpellier, France, 37–38.
- Lewis, D.K. (1969). *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- Meltzoff, A.N. & Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience, *Phil.Trans.R.Soc. Lond. B*, 358, 491–500.
- Murray, L. & Trevarthen, C. (1985). Emotional regulations of interactions between two-month-olds and their mothers. In T.M.F. Fox (Ed.), *Social perception in Infants*, Norwood, N.J.: Ablex., 177–197.
- Nadel, J. & Muir, D. (Eds.) (2005). *Emotional Development*, Oxford: Oxford University Press.
- Nolfi, S. & Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence and Technology of Self-Organizing Machines*. Cambridge MA: MIT Press.
- Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, 36, 1–16.
- Reddy, V. (2003). On being the object of attention: implications for self-other consciousness. *Trends in Cognitive Sciences*, 7, 9, 297–402.
- Rizzolatti G. & Craighero L. (2004). "The mirror-neuron system". *Annual Review of Neuroscience* 27: 169–192.
- Rohde, M. & Di Paolo, E. (2008). Embodiment and Perceptual Crossing in 2D: A Comparative Evolutionary Robotics Study. In *Proceedings of the 10th International Conference on the Simulation of Adaptive Behavior (SAB 2008)*, Osaka, Japan, 83–92.
- Rohde, M. (2010). Enaction, Embodiment, Evolutionary Robotics. Simulation Models for a Post-Cognitive Science of Mind. *Thinking Machines*, 1. Amsterdam, Paris: Atlantis Press.
- Sribunruangrit, N., Marque, C., Lenay, C., Gapenne, O. & Vanhoutte, C. (2002). Braille Box: Analysis of the Parallelism Concept to Access Graphic Information for Blind People. In *Proceedings of the Second Joint EMBS/BMES Conference*. Houston, Texas, 2424–2425.
- Thomas, R., Press, C., Haggard, P. (2006). Shared representations in body perception, *Acta Psychologica*, 121, 317–330.
- Tomasello, M., Carpenter, M., Call, J., Behne T. & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition, *Behavioral and Brain Sciences*, 28, 675–735.

Authors' addresses

Charles Lenay (corresponding author)
 Université de Technologie
 de Compiègne / Costech
 EA 2223 COSTECH (Knowledge,
 Organisation and Technical Systems)
 Centre Pierre Guillaumat
 BP 60319
 F – 60203 Compiègne Cedex
 France
 Email: Charles.Lenay@utc.fr

John Stewart
 Université de Technologie
 de Compiègne / Costech
 EA 2223 COSTECH (Knowledge,
 Organisation and Technical Systems)
 Centre Pierre Guillaumat
 BP 60319
 F – 60203 Compiègne Cedex
 France
 Email: john.stewart@utc.fr

Amal Ali Amar
 Université de Technologie
 de Compiègne / Costech
 EA 2223 COSTECH (Knowledge,
 Organisation and Technical Systems)
 Centre Pierre Guillaumat
 BP 60319
 F – 60203 Compiègne Cedex
 France
 Email: aaliamar@sqli.com

Marieke Rohde
 Multisensory Perception
 and Action Group
 MPI for Biological Cybernetics
 Spemannstrasse
 4172076 Tübingen
 Germany
 Email: marohde@gmail.com

Biographical notes

Charles Lenay is director of the Research Unit EA2223 COSTECH (Knowledge, Organisation and Technical Systems) at UTC, Compiègne University of Technology since 2004. After serving as Vice-President of the French Association for Cognitive Research (ARCo), he created in 1995 the Perceptual Supplementation Group (GSP) in the framework of this Research Unit. As a professor in Philosophy and History of sciences (2005) his research interests concern the question of cognitive technologies: how tools and technical devices participate in cognitive activity in the broad sense – reasoning, memory, perception, social interaction.

From 2003–2004 **John Stewart** was co-ordinator of the European Intensive Programme “Philosophy, Technology, Cognition”. From 1996–2003 he was editor of the journal “Intellectica”. He was director of COSTECH from 1996–1998. His habilitation thesis was entitled “Modelling biological systems” (1996). Currently his active research interests are in Cognitive Science – in particular the development of Enaction as a new paradigm, with emphasis on the specifically social dimension of human cognition, including technological systems; and the question of the origin of a primordial pre-DNA genetic system in living organisms.

Marieke Rohde currently works as a postdoctoral researcher in the Multisensory Perception and Action group at the Max Planck Institute for Biological Cybernetics in Tübingen. She received a B.Sc. in Cognitive Science (Osnabrück 2003), an M.Sc. in Evolutionary and Adaptive

Systems (Sussex 2004) and a D.Phil. in Computer Science and Artificial Intelligence (Sussex 2008). Her research interests include sensorimotor integration and adaptation, time perception, embodied and dynamical approaches in cognitive science, simulated agent modelling and biological constructivism.

Amal Ali Amar currently works in ergonomics (SQLI Agency) and conducts training in the design and evaluation of websites and business applications. She received her Ph.D. in Humanities and Technology of Cognition and Cooperation, at Compiègne in 2006.