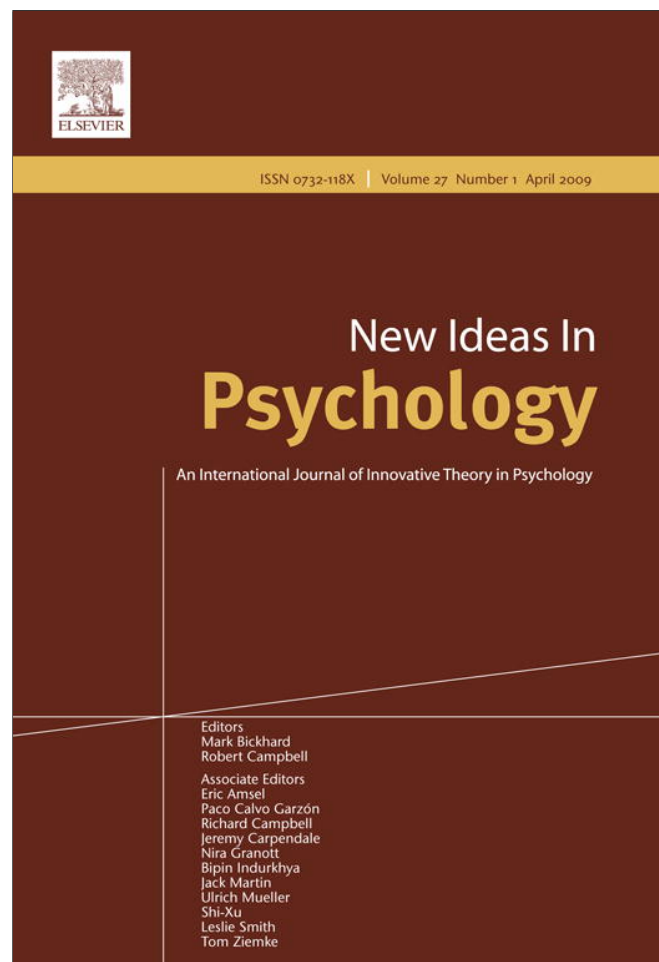




This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Perceptual interactions in a minimalist virtual environment

Malika Auvray*, Charles Lenay, John Stewart

COSTECH., Université de Technologie de Compiègne, Compiègne, France

Available online 3 March 2008

Abstract

How in real-life or through the use of technical devices can we recognize the presence of other persons and under what conditions can we differentiate them from objects? In order to approach this question, in the study reported here we explored the most basic conditions necessary for participants to recognize the presence of another person during a perceptual interaction. We created a mini-network of two minimalist devices and investigated whether participants were able to differentiate the perception of another person from the perception of a fixed and a mobile object even when the pattern of sensory stimulation was reduced to a bare minimum. We show that participants can recognize when the all-or-none tactile stimulation they experienced was attributable to an encounter with the other participant's avatar or the mobile object rather than with a fixed object. Participants were also able to establish different strategies in order to favor the situations of mutual perception. Thus, in the minimalist conditions of our experiment, the perception of another intentional subject was not based purely on any particular shape or objective trajectories of displacement; it was also based on properties that are intrinsic to the joint perceptual activity itself.

© 2008 Elsevier Ltd. All rights reserved.

Keywords: Interaction; Intentionality; Sensory substitution; Tactile perception; Intersubjectivity; Perceptual crossing

1. Introduction

Our natural tendency to ascribe mental states to others has been termed “naive psychology” (e.g., Clark, 1987; Hayes, 1979). More generally, the “theory of mind” refers to our ability to

* Corresponding author. COSTECH, Centre Pierre Guillaumat, Université de Technologie de Compiègne, BP 60319, 60206 Compiègne Cedex, France. Tel.: +33 660481444.

E-mail address: malika@malika-auvray.com (M. Auvray).

explain and predict others' actions by attributing causal intentional mental states to them, such as beliefs, desires, and intentions (e.g., see [Dennett, 1987](#); [Meltzoff, 1999](#)). The majority of studies conducted to date have investigated the mechanisms involved in the attribution of intentionality in situations of unilateral perception. Several factors have been proposed in order to account for the recognition of another intentional system under conditions where we do not interact with the perceived element. For example, the recognition of others has been explained by the adoption of a “teleological stance” which involves representing the actions of the other entity in terms of the principle of rational action. This principle assumes that intentional subjects act in order to achieve goals by the most efficient means available (see [Csibra, B  r  , Ko  s, & Gergely, 2003](#); [Csibra & Gergely, 1998](#); [Gergely & Csibra, 2003](#); [Gergely, N  dasdy, Csibra, & B  r  , 1995](#)). Other criteria are directly based on perceptual mechanisms such as the type of movement of the perceived object. If the movement can be explained simply by interactions with other objects, it will be attributed to a module of “naive physics” and considered as an object. On the other hand, if the perceived element seems to be able to change its direction in an autonomous way (self-propelled object) it will be attributed to a module of “naive psychology” and considered as an intentional subject (see [Baron-Cohen, 1994](#); [Baron-Cohen & Cross, 1992](#); [Premack, 1990](#); [Premack & Premack, 1997](#)).

In spite of their diversity, these criteria involve common underlying principles. In each case, recognizing the intentionality of another person is supposed to be the result of an internal judgment: the perceiver observes a behavior and then, on the basis of general criteria (such as its movements), she decides whether or not to interpret this behavior as being animated by intentional motives. However, it should be mentioned that early studies by [Heider and Simmel \(1944\)](#) suggested that there are no characteristics of form or behavior which are sufficiently specific to recognize with certainty the intentionality of an observed agent. The participants in Heider and Simmel's study easily attributed intentionality to a simple point in movement when its behavior was sufficiently rich. Thus, even though the movements of this point were causally determined, the observers, nevertheless, interpreted them as being motivated by intentional goals.

The alternative approach to perceptual interactions, the interactionist view, makes the hypothesis that the recognition of the intentionality of another person is intrinsic to a shared perceptual activity (e.g., [Fogel, 1993](#); [Stern, 2002](#)). The view that social interactions are created dynamically, as an emergent outcome of the interaction itself, is particularly well illustrated in the case of reciprocal perception between two intentional subjects (e.g., [Butterworth & Jarrett, 1991](#)). In particular, situations in which two perceptual activities of the same nature interact with each other (as in the case of mutual touch or catching one another's eye) make it possible to recognize the presence of another intentional subject. These particular situations seem to be immediately recognizable. For instance, when two people look at each other, what they perceive is not so much particular and determined movements of each others' eyes, rather, each person perceives a look, an intentional presence oriented toward themselves ([Argyle & Cook, 1976](#)). To illustrate this point, it has been shown that 9-month-old infants are sensitive to the direction of gaze of others and seem able to guess what is being looked at ([Scaife & Bruner, 1975](#)). As underlined by [Baron-Cohen and Cross \(1992\)](#) these results suggest that the understanding of the direction of someone else's gaze is fundamental for the understanding of her visual perceptual experience. However, it should be stressed that the existence of a perceptual crossing (i.e., mutual gaze) between an infant and an adult does not necessarily imply that the infant recognizes this particular situation. The perceptual crossing might only be a case of what [Tomasello, Carpenter, Call, Behne, and Moll \(2005\)](#) call a “protoconversation” or a “dyadic

engagement”. According to this view, the shared behaviors as well as the emotions that can arise from them can occur without the need for the infant to understand the internal structure of adult’s intentional actions.

The subsequent question arises: does the recognition of a perceptual crossing necessarily require the concept of an intentional subject who possesses internal goals that are a pre-requisite for her action? Or, alternatively, would it be possible to consider the perceptual crossing as occurring more directly, i.e., prior to the elaboration of a complete theory of intentionality? In the first perspective, that of the individualist approach, the recognition of a perceptual crossing is the result of a cognitive inference. This process can be conceived of as occurring in three stages. First, the observer perceives the behavior of another organism (its form and movements). Then, among the other’s movements, the observer differentiates between those which are merely the result of a purely mechanical causality, and those which can be attributed to an intentional subject who is perceptually directing her actions. Finally, the observer discriminates the particular cases when the goal of the other’s perception is precisely her own look. In this perspective, perceptions of social interactions, such as perceptual crossing, are first performed at an individual level (i.e., social interactions occur between cognitive systems initially isolated from each other), for instance by an innate module that detects the interactions (e.g., Dupuy, 1989; Gärdenfors, 2005; Gergely & Watson, 1996; Lewis, 1969).

By contrast, according to the second perspective, the interactionist approach, the recognition of a perceptual crossing occurs in a more direct way. According to this view, there are shared processes which serve as a basis for the discrimination between another perceptual intentionality and objects whose movements are independent of my own perception (Fogel, 1993; Schütz, 1962). In this case, the observers first have a relatively primitive capacity to recognize perceptual crossing directly; this ability then serves as a basis for subsequently understanding the gaze of others when they are oriented toward other objects. In this perspective, the social dimension is constituted collectively by the dynamics of the perceivers’ interactions. The importance of the processes of interaction for the recognition of others has been illustrated by Murray and Trevarthen (1985). In their studies, 2-month-old infants interacted with their mothers via a double-video projection. The video could display to the infants either their mother interacting with them in real time or a video pre-recorded from a previous interaction. The infants engaged in coordination with the video only in the first case, whereas in the second case they showed signs of distress. The fact that the children were able to distinguish a live interaction with their mother from a pre-recorded one suggests that the recognition of another subject does not only consist of the simple recognition of a particular shape or pattern of movements, but also involves the perception of how the movements of others are related to our own.

The aim of the study reported here is to further investigate whether, in situations of perceptual interactions, some of the mechanisms underlying the recognition of others are intrinsic to the shared perceptual activity itself (i.e., intrinsic to the interdependence between the two perceptual activities). To do so, pairs of participants were placed in a common virtual perceptual space via a network of two minimalist devices. Each participant moved an avatar, i.e., a representation of her body in the virtual environment which is used to perceive the objects of the environment and which, at the same time, can be perceived by the other participant. Each participant could encounter three types of objects within this simplified virtual space: the other participant’s avatar, a fixed object, and a mobile object (that we will refer to as the “mobile lure”). The fixed object and the mobile lure constitute representations of inanimate objects, whereas the avatars constitute representations of persons. In addition, the avatar and mobile lure of each person were designed in order to have the same shape and similar objective

trajectories of displacement. This was done in order to make sure that the only difference between the avatar and its corresponding mobile lure is that the avatar can at the same time perceive and be perceived. In other words, the only difference between the movements of the partner's avatar and those of the mobile lure is that the former are contingent on the active exploration by the participant. In this way, we made a simplified parallel with Murray and Trevarthen's (1985) experiment in which the infants could interact either in real time with a video of their mother or with a pre-recorded video, that is, a video that is consequently not contingent on the infants' actions. In the virtual environment we designed, the perception of the other participant's avatar as well as the perception of the fixed object and mobile lure generated a succession of all-or-none tactile stimuli. The task for participants was to recognize when the tactile stimuli they received were attributable to the encounter with the other participant's avatar. Our principal hypothesis was that the interdependence of the two perceptual activities might be a sufficient factor to enable the participants to click more often after having met the partner's avatar (i.e., respond correctly) than after having met the mobile lure.

2. Methods

2.1. Participants

Twenty participants (10 females and 10 males) took part in this experiment. Their ages ranged from 20 to 42 years (mean age of 29.4 years). All of the participants reported normal tactile perception. The experiment lasted 25 min and was performed in accordance with the ethical standards laid down in the 1991 Declaration of Helsinki.

2.2. Apparatus: the device *Tactos*

For this experiment, we used an adapted version of the minimalist sensory substitution system *Tactos* (see Lenay, Hanneton, Gapenne, Marque, & Vanhoutte, 2003). The device consisted of a computer mouse connected to a matrix of Braille cells, via a computer software written in C++. The participants explored graphic information by means of the computer mouse and received tactile stimulation on the index pad of their other hand (see Fig. 1A). The graphic information consisted of black pixels (giving rise to a stimulation) or white pixels (no stimulation) only. The displacement of the mouse determined the displacement of a cursor on the space of the computer screen (i.e., the relation between the cursor and the mouse was not absolute in terms of position but it was constrained by the relative displacement of the mouse). This cursor (that we refer to as the avatar) corresponded to a four-pixel receptor field. Each time the avatar covered a black pixel, an all-or-none tactile stimulation consisting of the simultaneous activation of eight tactors on the matrix of electronic Braille cells was delivered on the participant's index pad. There was no limitation on the duration of the tactile feedback: it lasted as long as the avatar overlapped with at least one black pixel. In this way, the sensory data available to each participant are reduced to a strict minimum, i.e., one bit of information corresponding to the activation or not of the tactile stimulation. This minimalist experimental paradigm makes it possible to record and to analyze the participants' perceptual trajectories, that is, the precise sequence of movements they executed and the corresponding sensory feedback they received.

The environment explored by the participants consisted of a one-dimensional space (i.e., a line), 600 pixels long, with the ends joined to form a torus in order to avoid singularities due to edges. There was a one-to-one mapping between the horizontal displacement of the

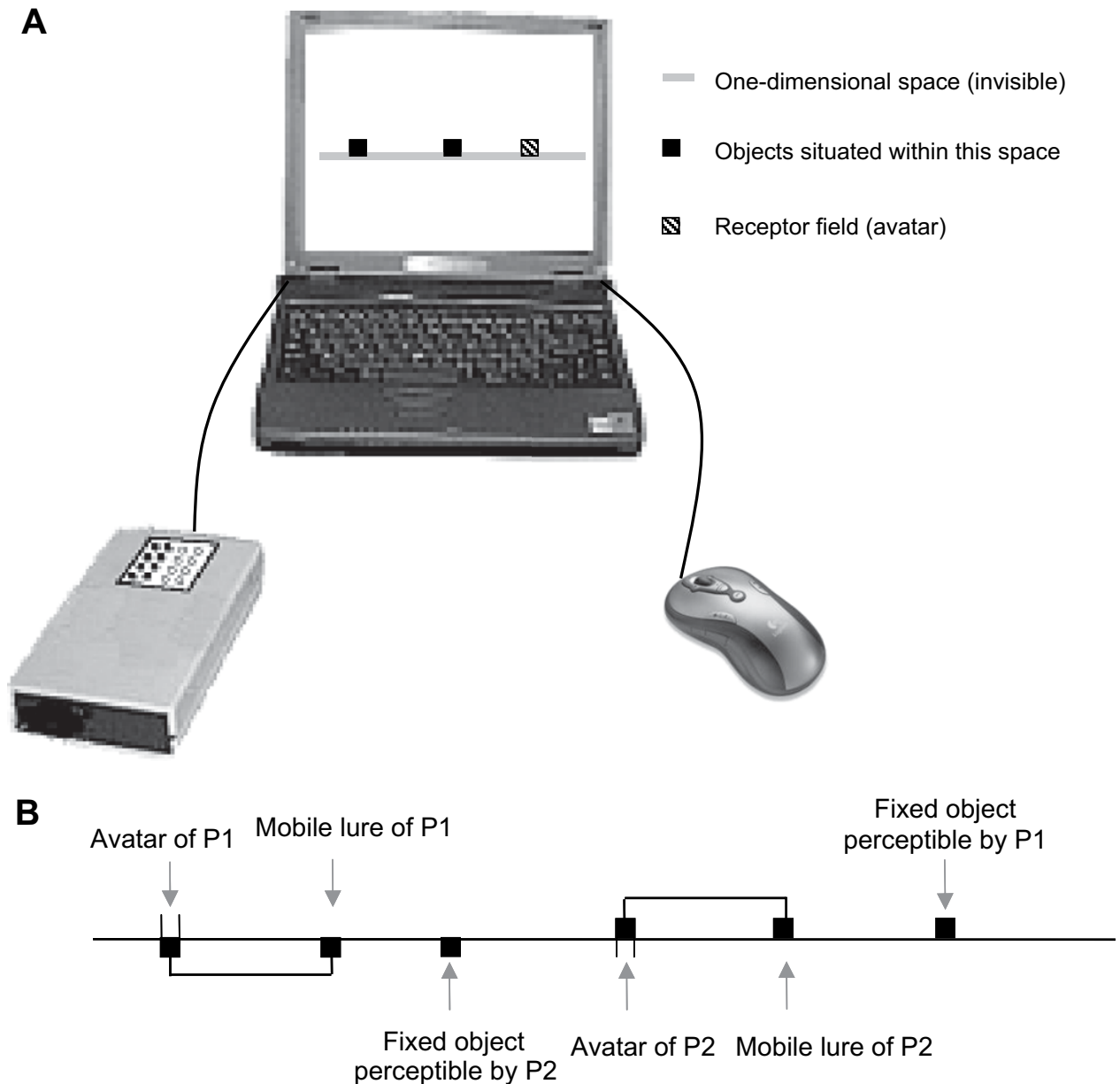


Fig. 1. (A) Schematic illustration of the experimental device Tactos developed by the Suppléance Perceptive team of Compiègne. The position of the cursor of the mouse determined the position of a four-pixel receptor field (i.e., avatar). Whenever the avatar covered at least one black pixel, it activated all eight tactile stimulators on the matrix of electronic Braille cells. (B) Schematic illustration of the one-dimensional space explored by the participants. It should be mentioned that, for clarity, we have represented in this figure, the three possible objects that can appear on the computer screen. However, the participants were blindfolded and did not have any access to these visual representations; they only accessed the resultant all-or-none tactile stimulations.

mouse and the horizontal displacement of the avatar. There was no constraint on the displacement of the mouse (i.e., it could be moved along a bi-dimensional surface), however, only the horizontal displacement of the mouse was taken into account by the software.

2.3. The shared virtual environment

Two Tactos devices were combined in a network so that two participants shared the common one-dimensional space. Each participant could encounter three types of object:

- (1) The four-pixel receptor field of the other participant, henceforth the “avatar”. Thus, when the two participants’ avatars overlapped (i.e., when there was an intersection of at least one pixel), both of them received an all-or-none tactile stimulation. We refer to this situation as the “perceptual interaction”. The fact that the two participants received tactile stimuli in case of an overlap between the two avatars was chosen in order to model other shared perceptions (such as mutual gaze or touch) where a perceiver cannot observe someone else without, at the same time, being perceived.
- (2) A fixed four-pixel wide object. The fixed object perceived by participant 1 (i.e., giving rise to tactile stimulation to participant 1 when her avatar overlapped with at least one pixel of this fixed object) was placed between 148 and 152 pixels and cannot be perceived by participant 2. The fixed object perceived by participant 2 was placed between 448 and 452 pixels and cannot be perceived by participant 1. These fixed positions were diametrically opposite on the one-dimensional torus. We chose not to place the two fixed object in the same position, in front of each other in order to favor the avatars’ displacement within the whole one-dimensional space.
- (3) A four-pixel wide object that we referred to as the “mobile lure”. In order to ensure that the mobile lure and the avatar have similar objective trajectories of displacement, the mobile lure was attached by a constant virtual rigid link at a distance 48–52 pixels from the center of the avatar (see Fig. 1B). The mobile lure thus reproduces exactly the same movements as its corresponding avatar. It should be noted that when participant 1 explored participant 2’s mobile lure, participant 2 did not receive any tactile feedback (and conversely, participant 1 did not receive any tactile feedback when her mobile lure was explored by participant 2). In contrast, when one of the participants explored the other participant’s avatar, both received tactile stimulation. It should be emphasized that these encounters constituted the only sensory information that the participants received. In particular, participants did not receive any information when their partner clicked on the mouse button. There are thus two objects — the avatar and its corresponding mobile lure — that have exactly the same behavior. The difference between the two is that only the overlap of the avatar (and not the mobile lure) with another object (fixed or mobile lure) gives rise to tactile stimuli.

2.4. Procedure

Before the evaluation session, the functioning of the device — the relation between the avatar, encounters with objects in the environment, and the tactile stimulation — was explained to the participants. The participants were then trained on the device during three phases of 1 min each: exploration of a fixed four-pixels wide object (announced as such); exploration of an object four-pixels wide moving at a constant speed of 15 pixels/s; and exploration of the same object moving at a constant speed of 30 pixels/s. The participants were thus trained with two different displacement speeds of the mobile object in order to prepare them to future possibly different displacement speeds of the other’s avatar. The task during the learning period was simply to maintain contact with the target.

The participants then started the evaluation session. The participants were placed by pairs and performed the experiment once. The two participants of each pair were blindfolded and seated in different rooms in front of the Tactos device. The two participants can only interact via the network, i.e., the one-dimensional virtual space. The experimental task was explained to the participants; they were told that they could freely explore the one-dimensional space

containing three types of tactile object: the partner's avatar, fixed objects, and mobile objects. However, the precise relation of the mobile lure yoked to the avatar was not explained. The instruction was to click on the left button of the mouse when the participants judged that their tactile sensations resulted from having met the other participant's avatar, and only in this case. The participants were instructed to click as many times as they judged that they met the other participants' avatar. Thus, the task for the participants involved neither a goal of mutual perception, nor a competition, but simply the recognition of the situations in which the tactile stimuli they received are due to having met the other participants' avatar. The evaluation session lasted 15 min, with short breaks after each 5 min. Thus, the duration of the evaluation session was constrained (15 min), but not the number of times the two participants' avatar met each other.

3. Results

3.1. General results

This section details the analysis of the participants' raw results, prior to providing an analysis of the categories of events that might have triggered their answers as well as their possible strategies.

3.1.1. Frequency of stimulation

We first classified each stimulation as being due to the avatar, to the fixed object, or to the mobile lure. To do so, we have classified across all the tactile stimuli *received* by the participants which object generated the tactile stimuli. An ANOVA performed on the Stimulation (avatar, mobile lure, and fixed object) highlighted a significant main effect of this factor [$F(2,38) = 38.92$; $p < 0.001$]. A Duncan post hoc test revealed a significant difference among all of the three conditions of Stimulation (all $ps < 0.01$). The majority of stimulation experienced by the participants was due to the encounter with the other participant's avatar ($52.2\% \pm \text{S.D. of } 11.8$) followed by the fixed object ($32.7\% \pm 11.8$), and then the mobile lure ($15.2\% \pm 6.2$).

3.1.2. Cause of clicks

The next stage was to attribute a cause to each of the clicks. To do this, we counted for each participant the number of stimulations in each category during the 2 s preceding each click.¹ The cause of the click was then attributed to the category with the largest number of stimulations. An ANOVA performed on the Cause of click (avatar, mobile lure, and fixed object) highlighted a significant effect of this factor [$F(2,38) = 74.32$; $p < 0.001$]. A Duncan post hoc test revealed a significant difference among all of the three conditions of Cause of click (all $ps < 0.05$). The great majority of clicks followed stimulations attributable to the other participant's avatar ($65.9\% \pm \text{S.D. of } 13.9$). The clicks following stimulation due to the mobile lure ($23.0\% \pm 10.4$) and due to the fixed object ($11.0\% \pm 8.9$) were less frequent.

3.1.3. Probability of clicks

Finally, for each participant, we divided the percentage of clicks by the percentage of stimulation, which gives a relative probability of clicks for each type of stimulation. An ANOVA

¹ The period of 2 s was chosen so that a minimal proportion of clicks (less than 0.3%) had no stimulation at all in this period.

performed on the Probability of clicks (avatar, mobile lure, and fixed object) showed a significant main effect of this factor [$F(2,38) = 20.27$; $p < 0.001$]. A Duncan post hoc test revealed a significant difference between fixed object (0.33) and the other two probabilities of clicks: avatar (1.26) and mobile lure (1.51) (both $p < 0.01$), but no significant difference between the latter two conditions.

3.1.4. Summary of the general results

The results on the cause of clicks showed that participants clicked significantly more often on the other participant's avatar than on the fixed object and mobile lure. An analysis of the probabilities of clicks gave a first indication of the strategies used by participants. The relatively small proportion of clicks on the fixed object is presumably due to the participants' capacity to differentiate between the fixed object and the two moving objects (avatar and mobile lure). Indeed, the probability of a click was four times lower when the source of the stimulation was a fixed object than when it was a moving object. By contrast, the participants did not seem able to discriminate between stimuli due to the avatar and those due to the mobile lure: the probability of a click for these two types of stimuli was not significantly different. Thus, the large difference observed between clicks on the avatar and clicks on the mobile lure (65.9% vs. 23.0%) must be attributed to the joint strategies of displacement, which are such that the frequency of stimulation associated with the avatar was higher than the frequency of stimulation associated with the mobile lure (52.2% vs. 15.2%). If the participants succeeded in the perceptual task, it is essentially because they succeeded in situating their avatars in front of each other. There are thus two observations to account for: first, the participants' ability to distinguish between the fixed object on one hand and the two moving objects (avatar and mobile lure) on the other hand; and second, their ability to favor the situations of mutual perception.

3.2. Discrimination between the fixed object and the two moving objects (avatar and mobile lure)

The first analysis aims to account for the participants' ability to differentiate between the stimulations due to a fixed object and the stimulations due to a moving object be it avatar or mobile lure. Previous studies with the Tactos device on purely perceptual tasks of exploration of static objects revealed a very general strategy of the participants (Sribunruangrit, Marque, Lenay, Gapenne, & Vanhoutte, 2004). This strategy, apparently also used in the present experiment, consists of generating oscillations around the point of stimulation. This result leads us to propose a general *a priori* hypothesis: an object will be identified as being a moving object if the expectations appropriate to the perception of a fixed object are flouted, in one way or another. On this basis, we identified seven types of event that can only occur in the presence of a moving object. These events are as follows: there is a change in stimulation even though the participants did not move (E1). There are two distinct consecutive stimuli even though the participants have been moving monotonically in a constant direction (E2). The next three types of event occur when the participants leave a source of stimulation and then reverse direction to relocate this source of stimulation. If the stimulation is due to the fixed object, the location of the second stimulation should be identical to that of the first. If the stimulation is due to a moving object, the second stimulation can be encountered sooner than expected (E3), later than expected (E4), or not encountered at all (E5). Finally, if the participants explore a fixed object, the distance over which continuous stimulation is encountered is seven pixels (four-pixels width and three pixels over which the avatar overlap partly with the fixed object). The

Table 1

Frequency with which each type of event allowing for the recognition of a moving object was followed by a click

Types of event		C1	C2	C3	C4
		Overall frequencies	% Of clicks	Event efficiency	% Of success
E1	Changes in the stimulation without moving	23.9	54.9	20.9	98.6
E2	Two consecutive stimuli experienced	19.4	32.3	15.1	93.5
E3	Stimulation encountered sooner than expected	10.5	14.4	12.5	84.3
E4	Stimulation encountered later than expected	10.2	14.7	13.1	90.6
E5	Stimulation not encountered at all	10.6	11.6	10.0	95.5
E6	Distance lower than seven pixels	16.8	31.3	17.0	98.2
E7	Distance bigger than seven pixels	8.6	9.1	9.6	79.6
Total/mean		100.0	168.4	14.0	91.5

C1: Relative frequencies of each type of event in the overall data.

C2: Percentage of clicks that were preceded by an event of each type. The sum is more than 100 because a click can be preceded by several events.

C3: The probabilities that an event of each type would be followed by an effective click. These probabilities are generally well below 100%.

C4: For clicks preceded by each type of event, the probability of “success”, i.e., of a click on a moving object (avatar or mobile lure) and not on the fixed object. The mean value of 91.5% as reconstructed from these categories of “events” is close to the observed results of 89.0 ± 2.0 showing that these events constitute a reasonable explanation of participants’ ability to distinguish between fixed and moving objects.

participants can thus identify a moving object if the distance is lower than seven pixels, which occurs when the two avatars cross each other with opposite movements (E6) or if it is bigger than seven pixels, which occurs when the two participants cross each other when moving toward the same direction with two different displacement speeds (E7).

The frequency with which each type of event was followed by a click was calculated across all the participants’ results. The results regarding these events are shown in Table 1. In particular, an important percentage of participants’ clicks was preceded by three types of events: the fact that there was change in stimulation even though the participant did not move (E1, 54.9%), the experience of two distinct consecutive stimuli even though the participant has been moving monotonically in a constant direction (E2, 32.3%), and the fact that the participant experiences a distance lower than seven pixels when crossing an object (E6, 31.3%).

3.3. Convergence toward the situations of perceptual interaction

The second result to account for is the participants’ ability to favor the situations of mutual perception. As was previously mentioned, studies with the minimalist device Tactos have revealed that the perception of an object is based on active and reversible exploration. Participants oscillate around the singularity which results in sensory feedback (Sribunruangrit et al., 2004; Stewart & Gapenne, 2003). More precisely, a basic strategy consists of turning back each time some form of sensory stimulation is encountered. When such a strategy succeeds, the temporal succession of stimuli gives rise to the perception of a spatially localized object (Lenay, Gapenne, Hanneton, Genouëlle, & Marque, 2003). In order to investigate whether participants used this strategy in the present study, we computed the mean acceleration during 1 s after losing contact with a source of stimulation as a function of the mean velocity during 1 s preceding this event. The strong negative correlation of -0.72 indicates that when participants

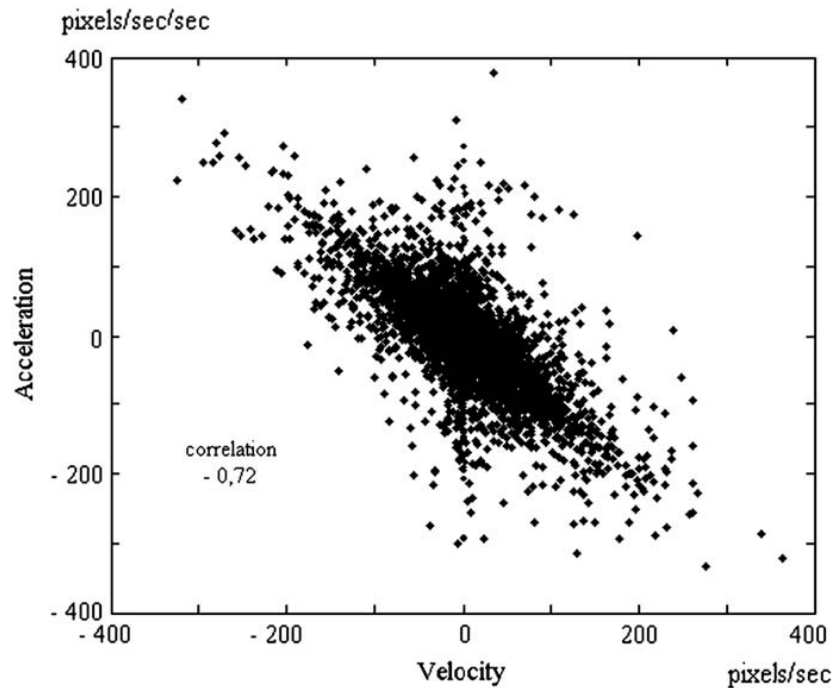


Fig. 2. Mean acceleration during 1 s after losing contact with a source of stimulation as a function of the mean velocity during 1 s preceding this event. The analysis, performed on the grouped data from all participants, shows a negative correlation of -0.72 . This negative correlation reveals that when participants crossed an object in the environment, they strongly tended to reverse their direction of movement.

crossed a source of stimulation, they tended to reverse their direction of movement (see Fig. 2). An example of the joint trajectories of the two participants' avatars is shown in Fig. 3.

We then calculated the frequency distribution of the distances between the two participants' avatars. As illustrated in Fig. 4, there was an attraction of the two avatars for the situations of

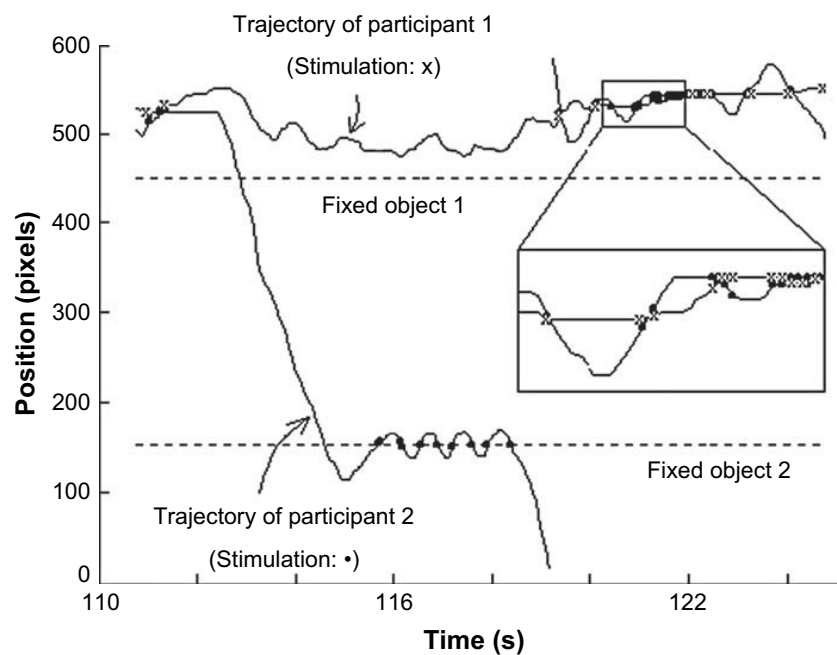


Fig. 3. An illustrative example of joint trajectories. After an initial contact, participant 1 explored the whole one-dimensional space, oscillated for about 3 s around the fixed object (without clicking), then continued (past the junction of the torus) and re-encountered participant two's avatar. At this point the two participants engaged in mutual oscillations of decreasing amplitude (shown enlarged in the inset) following which, both participants did click.

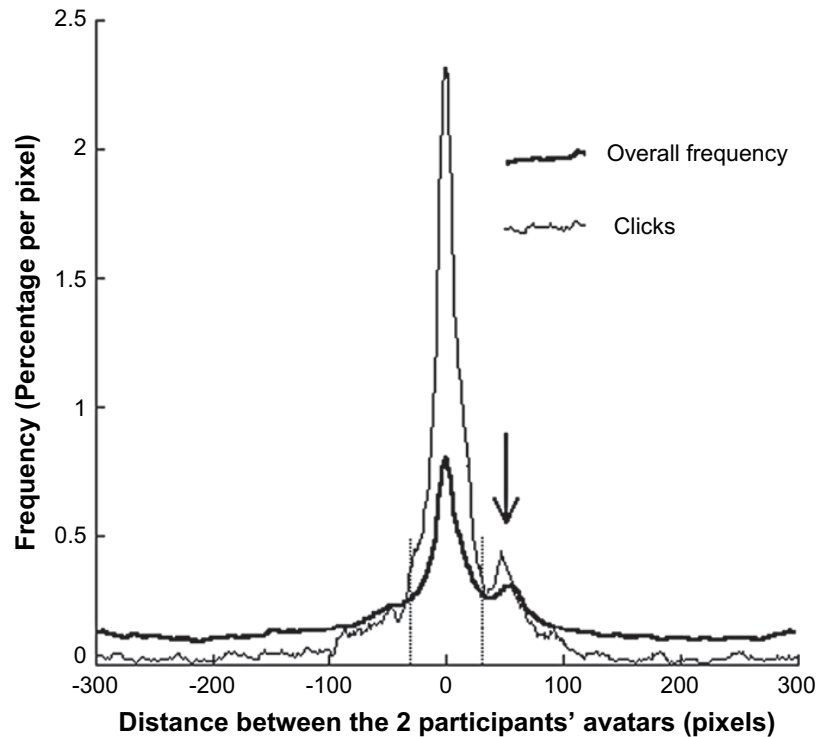


Fig. 4. Frequency distributions as a function of the distance between the two participants' avatar. The thick line represents the overall unconditional frequency distribution; 28% of the distribution laid between ± 30 pixels (10% of the total space), as indicated by the dotted lines. The thin line represents the distribution of distances when the participants clicked; 62% of the distribution laid between ± 30 pixels. In both cases there is a clear peak at a distance of zero pixel i.e., in situations of mutual perception; showing the existence of an attractor around this point. The slight subsidiary peak at a distance of 50 pixels, marked by an arrow, corresponds to the mobile lure.

perceptual interaction. We will refer to this mutual attraction by using the dynamic system concept of “attractor”. There was also a slight subsidiary peak on the mobile lure. One main hypothesis can be put forward in order to explain the strongest attraction around the site of the perceptual interaction. When the trajectories of the avatars cross, *both* participants receive a stimulation; if, as explained above, each participant then turns back, then they will meet again, and this pattern forms a relatively stable dynamic attractor. This co-dependence of the two perceptual activities contributed to favor the situations of mutual perception. By contrast, participant 2 did not receive any tactile feedback when participant 1 explored her mobile lure (and vice-versa). Thus there could not be any dependence between the movements of the two participants, and the mobile lure easily “escaped” from the attractor. We also calculated the frequency distribution of clicks as a function of the distance between the two participants' avatars at the time of the click. The peak in the distribution of the clicks when the participants' avatars were opposite one another in a situation of perceptual interaction was even sharper than for mere stimulations.

In order to explain further the distribution of participants' clicks, we measured the probability of clicks as a function of the number of distinct stimulations received during 2 s. As shown in Fig. 5, the probability of clicking on a moving object (avatar or mobile lure) increased with the number of stimuli: from 6% for one stimulus to 76% when eight stimuli were experienced during 2 s. It should also be noted that an increase in stimuli due to the fixed object did not favor the probability of clicks, which confirms participants' ability to distinguish between moving and fixed objects. In summary, the distribution of participants' clicks may be due on one hand to the

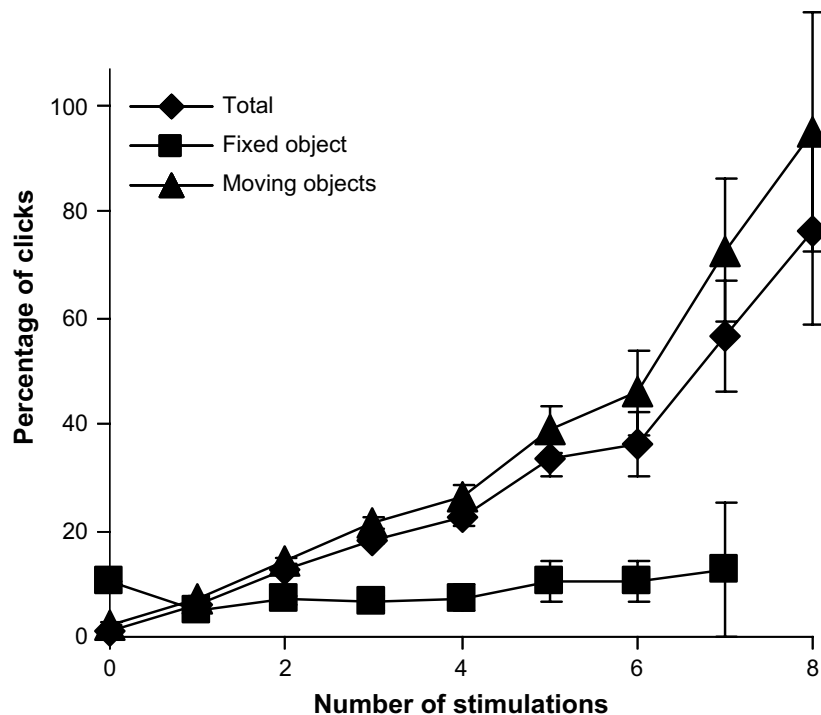


Fig. 5. The probability of participants' clicks, plotted as a function of the number of distinct stimulations received during the preceding 2 s. Lozenges: total stimulations (avatar, fixed object, and mobile lure); squares: stimulations due to encounters with the fixed object; triangles: stimulations due to encounters with a moving object (avatar or mobile lure). Error bars represent the standard errors of the means.

exploratory trajectories which favored the situations of perceptual interaction; and on the other hand to the capacity of participants both to disregard stimuli due to the fixed object and to respond more often when the number of stimuli received during a short period of time increased.

4. Discussion

The experiment reported here revealed that participants interacting in a minimalist environment clicked significantly more often on the other participant's avatar (i.e., correctly) than on the fixed object and the mobile lure. Subsequent analysis revealed that participants were able to distinguish when the patterns of stimuli they received were due to the fixed object rather than to the two moving objects (avatar and mobile lure). However, they were not able to differentiate the stimulation due to the other participant's avatar from the one due to the mobile lure (the probability of a click was similar). The high proportion of correct responses was explained by the participants' exploratory trajectories which converged toward joint strategies of mutual exploration. Thus, if the recognition of an intentional subject was to be defined as the consequence of an ability to categorize the stimuli, this recognition does not occur in our experiment: for a given stimulus, the probability of clicks was similar for the avatar and the mobile lure. This result is not surprising given that both have identical objective trajectories of displacement. However, the higher proportion of clicks on the partner's avatar than on the mobile lure can be accounted for by the sensorimotor dynamics of the interaction which favored an attraction for the situations in which the two avatars are located in front of each other (that is, the situations of mutual perception).

In the experiment reported here, the criteria that favored the attribution of intentions to others thus resulted from two different processes: first, a set of perceptual criteria based

on the characteristics of displacement of the object (which allowed for a distinction between the fixed object and the two moving objects: avatar and mobile lure); and second, a process intrinsic to the dynamics of the interaction (which favored the situations of perceptual interaction). With regard to the first set of criteria, the distinction between the fixed object and the two moving objects can be understood as a set of spatial characteristics of the objects linked to the minimalist design that was deliberately adopted. Given the all-or-none nature of the sensory feedback, the perception of an object becomes possible only by means of a dynamic exploration. More precisely, the spatial characteristics of an object can be defined by specific “laws of sensorimotor contingencies” which make it possible to anticipate the sensory consequences of one’s actions during the course of an active exploration (Noë, 2005; O’Regan & Noë, 2001). To illustrate this idea, we have observed a rather general strategy adopted by the participants which consisted of inverting the direction of the movement of their avatar after each encounter giving rise to tactile feedback (as shown in Fig. 2). This strategy gives rise to predictions with regard to the encountered objects. For example, a regular symmetrical oscillation of the avatar around a sensory event constitutes the perception of a fixed object. An asymmetrical oscillation around a source of stimulation which is constantly displaced constitutes the perception of an object which is moving. However, in the cases when the partner is also engaged in an active perceptual activity, her avatar and mobile lure also have oscillatory displacement and the superposition of these two oscillatory movements prevents an accurate determination of their location. This criterion favoring the attribution of intentions can be summarized as “something that resists being spatially determined”: it is neither a fixed object, nor an object with a simple displacement. This criterion is in line with the idea developed by Sartre (1943, p. 299) that the other, as an intentional subject, cannot be constituted as an objectively determined entity: “one meets the other, one does not constitute him”.

With regard to the second process, the fact that the situations of perceptual interactions were favored can be accounted for by the dynamics of the interaction. In particular, even though the participants do not have any explicit or conscious intention to collaborate, their simultaneous efforts to search for their partner’s avatar gave rise to an attractor in the conjoint dynamics of their perceptual activities. In other words, the sole difference between the avatar and the mobile lure attached to it is that the former is sensitive to other’s presence, whereas the latter is not. This co-dependence of the two perceptual activities served as a basis for the formation of an attractor in the collective dynamics. In this situation, the active perceptual activities mutually attracted each other, just as in everyday situations there is an attraction to the situations where two people catch each other’s eye. This result is in line with the interactionist approach, according to which there are processes of social interaction that originate within the interaction itself and thereby are not *per se* present in each of the individuals taken separately. In other words, some components of the social interactions are created dynamically, thanks to the interaction itself (e.g., Stern, 2002).

The minimalist experimental paradigm presented in this paper can serve as a basis to further investigate the dynamics of social interactions. For example, future research could investigate situations in which the virtual space consisted of two or three dimensions in order to determine if similar results would be obtained. Future research could also investigate the parameters which enable participants to distinguish perceptual crossing with another participant from an algorithmic agent implementing simple perceptual strategies. Thus, by varying the parameters governing the agent (or the mobile lure) as well as its degree of reactivity when it receives stimulations, it should be possible to determine under which conditions it is

possible to create the illusion of an encounter with an intentional subject. In addition, the paradigm reported here can easily be implemented in evolutionary robotics computer interactions in order to address the issue of the sensitivity to perceptual interactions. It should be mentioned that Di Paolo, Rohde, and Iizuka (2008) have already started a program of simulation research in that direction, based on the paradigm and methods reported in this paper. Their evolutionary robotics simulations showed similar results as the one reported in our study. Interestingly, and contrarily to any *a priori* prediction, Di Paolo et al. found it easier to evolve agents that can distinguish between the avatar and mobile lure than agents that can distinguish between the avatar and fixed object. As a consequence, according to Di Paolo et al., in the case of social interactions, it is simply not necessary to evolve simulated agents with an individual contingency recognition strategy, given that the social process takes care by itself of inducing the individuals to produce the right behavior.

In addition, research on the interactive dimension of the recognition of mutual perception has immediate practical consequences in the area of communication technologies. This research shows that rich perceptions become possible, even with a minimal channel of communication, as soon as the perceptual activities of two or more partners can be linked with each other.² Finally, the interactionist approach to social interactions can also shed light on work on developmental psychology investigating the ability to participate in perceptual crossing very early in infant life. Indeed, it has been shown that there is sensitivity to perceptual interactions from the age of 2 months (Nadel, Carchon, Kervella, Marcelli, & Reserbat-Plantey, 1999). However, the infants can barely be said to have a real recognition of the perceptual crossing in the full sense of the term: this would only come later together with the maturation of the rest of the system of common-sense psychology. The hypothesis of a more direct ability for the infants to be involved in a perceptual crossing explains how it can appear early in life. In this case, the infants' ability to be involved in a perceptual crossing could subsequently play an important role in the genesis of a common-sense psychology.

In summary, the paradigm of minimalist perceptual crossing presented here has made it possible to propose an original approach to the question of the recognition of other intentional subjects. The results reported here suggest that there is sensitivity to the situations of perceptual interaction. This sensitivity, instead of being perceived by each of the participants, arises from the dynamics of the interaction itself. Thus, the paradigm and results presented in the present study provide tools to take up the challenge of investigating the interaction process as a whole, rather than as purely individual abilities to behave socially (see also Di Paolo et al., 2008). They also allow taking into account the dialectics between the individual and social levels of description during social interactions.

Acknowledgement

This research was supported by a grant from the European network of excellence ENACTIVE (IST-002114) and by a grant from France Télécom R&D (Paris, France). We wish to thank Alexandre Lang and Fabien Bénéto for their assistance in running the experiment. We are also grateful to Charles Spence for his helpful comments on this manuscript.

² These possibilities are at the origin of the conception of new systems of tactile interaction via telephone and Internet networks, currently under development at the Compiègne University of Technology, in collaboration with France Télécom R&D.

References

- Argyle, M., & Cook, M. (1976). *Gaze and mutual gaze*. Cambridge: Cambridge University Press.
- Baron-Cohen, S. (1994). How to build a baby that can read minds: cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive*, 13, 1–40.
- Baron-Cohen, S., & Cross, P. (1992). Reading the eyes: evidence for the role of perception in the development of a theory of mind. *Mind & Language*, 7, 172–186.
- Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9, 55–72.
- Clark, A. (1987). From folk psychology to naive psychology. *Cognitive Science*, 11, 139–154.
- Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, 27, 111–133.
- Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: a developmental hypothesis. *Developmental Science*, 1, 255–259.
- Dennett, D. C. (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.
- Dupuy, J. P. (1989). Common knowledge, common sense. *Theory and Decision*, 27, 37–62.
- Fogel, A. (1993). *Developing through relationships: Origins of communication, self, and culture*. Chicago: University of Chicago Press.
- Gärdenfors, P. (2005). *The dynamics of thought*. Dordrecht: Springer.
- Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naive theory of rational action. *Trends in Cognitive Sciences*, 7, 287–292.
- Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, 56, 165–193.
- Gergely, G., & Watson, J. L. (1996). The social biofeedback theory of parental affect-mirroring: the development of emotional self-awareness and self-control in infancy. *International Journal of Psycho-Analysis*, 77, 1181–1212.
- Hayes, P. J. (1979). The naive physics manifesto. In D. Michie (Ed.), *Expert systems in the microelectronic age* (pp. 242–270). Edinburgh: Edinburgh University Press.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243–259.
- Lenay, C., Gapenne, O., Hanne-ton, S., Genouëlle, C., & Marque, C. (2003). Sensory substitution: limits and perspectives. In Y. Hatwell, A. Streri, & E. Gentaz (Eds.), *Touching for knowing* (pp. 275–292). Amsterdam: John Benjamins.
- Lenay, C., Hanne-ton, S., Gapenne, O., Marque, C., & Vanhoutte, C. (2003). Procédé permettant à au moins un utilisateur, notamment un utilisateur aveugle, de percevoir une forme et dispositif pour la mise en oeuvre du procédé [Procedure allowing at least one user, notably a blind user, to perceive a form, and device allowing the procedure]. Patent US-2004-0241623-A1.
- Lewis, D. K. (1969). *Convention: A philosophical study*. Cambridge: Harvard University Press.
- Meltzoff, A. N. (1995). Origins of theory of mind, cognition and communication. *Journal of Communication Disorders*, 32, 251–269.
- Murray, L., & Trevarthen, C. (1985). Emotional regulations of interactions between two-month-olds and their mothers. In T. M. Field, & N. A. Fox (Eds.), *Social perception in infants* (pp. 177–197). Norwood: Ablex.
- Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Reserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. *Developmental Science*, 2, 164–174.
- Noë, A. (2005). *Action in perception*. Cambridge, MA: MIT Press.
- O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24, 939–973.
- Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, 36, 1–16.
- Premack, D., & Premack, A. J. (1997). Infants attribute value $\pm$ to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience*, 9, 848–856.
- Sartre, J. P. (1943). *L'être et le néant [Being and nothingness]*. Paris: Gallimard.
- Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, 253, 265–266.
- Schütz, A. (1962). In M. Natanson (Ed.), *The problem of social reality. Collected papers, Tome 1*. La Haye: Editions Martinus Nijhoff.

- Sribunruangrit, N., Marque, C., Lenay, C., Gapenne, O., & Vanhoutte, C. (2004). Speed-accuracy tradeoff during performance of a tracking task without visual feedback. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 12, 131–139.
- Stern, D. (2002). *The first relationship: Infant and mother*. Cambridge: Harvard University Press.
- Stewart, J., & Gapenne, O. (2003). Reciprocal modelling of active perception of 2-D forms in a simple tactile-vision substitution system. *Minds and Machines*, 14, 309–330.
- 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.