



# Minimalist approach to perceptual interactions

Charles Lenay\* and John Stewart

EA 2223 COSTECH, Département Technologie et Sciences de l'Homme, Université de Technologie de Compiègne, Compiègne, France

## Edited by:

Ulrich Pfeiffer, University Hospital  
Cologne, Germany

## Reviewed by:

Ezequiel Alejandro Di Paolo,  
Ikerbasque – Basque Foundation for  
Science, Spain  
Hanne De Jaegher, University of the  
Basque Country, Spain

## \*Correspondence:

Charles Lenay, EA 2223 COSTECH,  
Département Technologie et Sciences  
de l'Homme, Université de  
Technologie de Compiègne, Centre  
Pierre Guillaumat, BP 60319, Rue du  
Docteur Schweitzer, 60203  
Compiègne Cedex, France.  
e-mail: charles.lenay@utc.fr

Work aimed at studying social cognition in an interactionist perspective often encounters substantial theoretical and methodological difficulties: identifying the significant behavioral variables; recording them without disturbing the interaction; and distinguishing between: (a) the necessary and sufficient contributions of each individual partner for a collective dynamics to emerge; (b) features which derive from this collective dynamics and escape from the control of the individual partners; and (c) the phenomena arising from this collective dynamics which are subsequently appropriated and used by the partners. We propose a minimalist experimental paradigm as a basis for this conceptual discussion: by reducing the sensory inputs to a strict minimum, we force a spatial and temporal deployment of the perceptual activities, which makes it possible to obtain a complete recording and control of the dynamics of interaction. After presenting the principles of this minimalist approach to perception, we describe a series of experiments on two major questions in social cognition: recognizing the presence of another intentional subject; and phenomena of imitation. In both cases, we propose explanatory schema which render an interactionist approach to social cognition clear and explicit. Starting from our earlier work on perceptual crossing we present a new experiment on the mechanisms of reciprocal recognition of the perceptual intentionality of the other subject: the emergent collective dynamics of the perceptual crossing can be appropriated by each subject. We then present an experimental study of opaque imitation (when the subjects cannot see what they themselves are doing). This study makes it possible to characterize what a properly interactionist approach to imitation might be. In conclusion, we draw on these results, to show how an interactionist approach can contribute to a fully social approach to social cognition.

**Keywords:** perceptual crossing, imitation, recognition of intentionality, interaction, minimalism

## INTRODUCTION

Research on social cognition often finds itself caught in an uneasy, paradoxical tension. On the one hand, understanding social phenomena requires giving an essential place to collective dynamics. On the other hand, from the viewpoint of cognitive science, one adopts a more or less explicit version of methodological individualism according to which social interactions must be explained on the basis of individual capacities. In this case, it is indeed difficult to maintain that there exist components which are proper to the collective domain, especially if one admits that such components only take on meaning if they are taken up in individual experience, and moreover that the explanation of the emergence of such components is based on individual competencies. It seems obvious that if there is a social interaction, the interacting subjects must possess the individual competences that are necessary for this interaction to occur. It would therefore seem a legitimate method to isolate the subjects in order to identify these competences, before going on to study how the individuals interact. The classical approaches, prolonging such a procedure, consider that these individual competences correspond to all the social know-how and knowledge involved: a capacity to recognize other subjects, to imitate, and finally common-sense psychology (theory of mind). But in this case, how is it possible to account for the inter-individual coordination of actions in space and time in order to attain shared

goals and to organize in society according to a particular culture? The classical answer is to suppose that individual capacities for coordination and joint attention make it possible to share representations of current objects and events (Sebanz et al., 2006; Knoblich et al., 2011); and it is thought that the neurological basis for these capacities are to be found in the “mirror-neuron” systems which associate representations of actions performed with representations of observed movements (Blakemore and Decety, 2001; Gallese et al., 2004; Rizzolatti and Craighero, 2004; Gallese, 2007). However, if a “representation” is a state or process that is strictly internal to each subject, the collective remains in fact internalized in each subject (Tomasello et al., 2005; Thomas et al., 2006) and it is quite difficult to understand how the “social” as such could play a constitutive role (see for example the synthetic dossier in *Topics in Cognitive science*; Galantucci and Sebanz, 2009).

By contrast, interactionist approaches postulate that interactions play a role right from the start in the constitution of social phenomena. These are ancient traditions in social psychology and philosophy, from the work of George Herbert Mead (Gillespie, 2005) and Erving Goffman (Goffman and Best, 2005) to the pragmatist approach of John Dewey (Petras, 1968) or the social psychology of Theodore Newcomb (Newcomb et al., 1966), from the phenomenology of Alfred Schütz (Schütz, 1970) to the work of Thomas Luckmann and Peter Berger (Berger and Luckmann,

1966). This sort of attention to interactions is found currently in the framework of theories of development (Trevorthen, 1993; Reddy, 2008), in dynamical systems theory (Coleman and Watson, 2000; Fogel, 2006), in the current of ecological psychology (Gibson, 1986), in enactive cognitive science (De Jaegher et al., 2010), and recently in certain areas of neuroscience (Schilbach et al., 2006; Dumas, 2011). However, all these studies are still affected by a tension between the collective aspects and individual cognition.

For example, studies in ecological psychology bear essentially on phenomena of the coordination of action (Schmidt and Richardson, 2008). These phenomena turn out to be very general, and can be explained by mechanisms related to the physics of embodiment and perceptual systems. However, it seems to us that this approach lacks an explanation of the passage from coordinations of this sort, and social interactions that are meaningful for the actors themselves. Similarly, in the framework of externalist conceptions of situated cognition and the extended mind (Hutchins, 1995; Clark and Chalmers, 1998) one can develop models of social activity based on a shared material and technical environment. And in the perspective of *embodied embedded cognition* (van Dijk et al., 2008), one can seek to understand social behaviors in terms of the relations of the organism with its social ecological niche (Marsh et al., 2009), or with social norms (Steiner and Stewart, 2009). But in all these cases, the social structures, or traces left by collective activity, must at some point be mobilized by individuals; and this requires understanding how the environment comes to have such a meaning for the actors themselves – for example, how they come to recognize other organisms as intentional subjects, how they recognize the behaviors as action following some norms.

In the enactive approach, explanations start from the viewpoint of the living organism (Varela et al., 1999). Rather than evoking internal representations, this approach is based on the coupling between an organism and its environment which results in the enaction of a meaningful world. In this framework, De Jaegher proposes to articulate the collective dynamics of social interactions and individual autonomy. The interaction between organisms is regulated by the organisms themselves through an activity of collective construction of meaning, a *participatory sense-making* (De Jaegher and Di Paolo, 2007). In this version, the interactionist position can postulate that social capacities and knowledge are actually constituted in the very course of social interactions; the essential aspects of individual social capacities are the result, and not the cause, of social interactions (McGann and De Jaegher, 2009; De Jaegher et al., 2010).

Now even this strong interactionist position poses a certain number of theoretical and methodological difficulties. First of all, methodologically, it is necessary to identify the relevant behavioral variables, and to record them without disturbing the interaction. Even more important, in order to account for the emergence of social phenomena in the course of the interaction, it is necessary to clearly identify and distinguish between (a) what is contributed by the individual subjects and (b) what is contributed by the actual dynamics of the interaction, in order to show that the dynamics of inter-individual interaction leads to *more* than what the individual organisms in interaction bring to the situation. At the same time, it is also necessary to explain (c) how some of these emergent

phenomena can be subsequently appropriated by the individuals (Froese and Di Paolo, 2011b). As a basis for this methodological and conceptual discussion, we will propose here a *minimalist* experimental paradigm for the study of perceptual interactions in which it will be possible to clearly distinguish these three components. This will allow us to propose some explanatory schema concerning two major questions in social cognition: the recognition of the presence of another intentional subject; and the realm of imitation. In order to present this experimental paradigm and to argue for its heuristic value, we start by describing a minimalist method for the analysis of individual perception.

## A MINIMALIST EXPERIMENTAL METHOD FOR THE STUDY OF PERCEPTION

Many empirical studies of inter-individual interactions proceed by setting up well-controlled conditions of observation (Marsh et al., 2009). However, even if the sorts of actions that can be performed are simplified as much as possible, the subjects interact using their natural perceptual systems. It is therefore very difficult to observe the internal mechanisms for controlling the actions, and it is difficult to understand the link between these actions and the social meaning that the individuals attribute to the interactions. The minimalist method that we propose here aims at providing the means to answer these questions, by precisely controlling the perceptual activities of the subjects in interaction. To do this, we base ourselves on studies of sensory substitution.

The general principle of such studies consists of transforming the stimuli proper to one sensory modality (for example ocular vision) to stimuli of a different sensory modality (for example touch; Collins and Bach-y-Rita, 1973; Schiff and Foulke, 1982; Visell, 2009). On condition that the user is active (manipulating the camera by lateral movement, rotation and zoom), (s)he is able to develop spectacular perceptual capabilities, in particular for the spatial localization and recognition of shapes (Guarniero, 1974; Bach-y-Rita, 2004). The use of a technical mediation for perception has four major advantages:

- (1) These devices enable the study of the genesis of a new kind of perceptual modality, in an experimental situation that can be closely controlled (Collins and Bach-y-Rita, 1973; Lenay et al., 2003; Auvray and Myin, 2009). Of course, this sort of perception is quite particular, and has to occur on the background of perceptual know-how already present for the user. Nevertheless, we are here in the presence of a genuine genesis of novel perceptual capabilities which were clearly absent before the learning process.
- (2) The perceptual learning involved in this experiment is evidence of an impressive plasticity of the central nervous system. The tactile sensory input has nothing to do with that of ocular vision, just as the control of the camera with the hands has nothing to do with the commands to ocular muscles and the head. Nevertheless, the technical device defines a space of coupling, a specific set of sensorimotor regularities. In conditions suitable for progressive learning, the use of this device leads to vast functional reorganization (Bach-y-Rita, 1990; De Volder et al., 1999), which results in robust, general know-how (Bach-y-Rita and Kerckel, 2003) and a perceptual world where

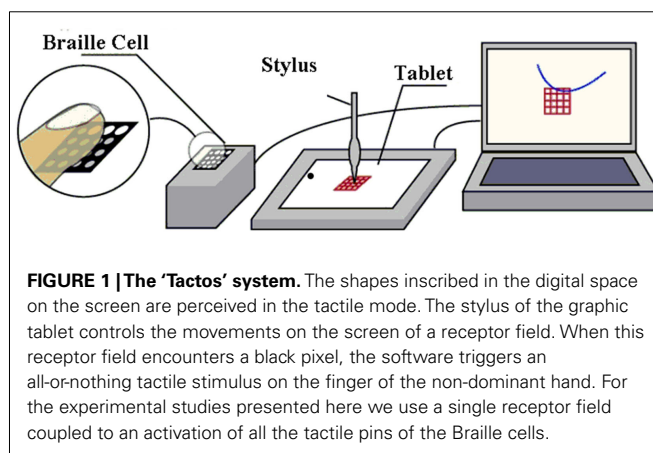
the shapes and events are highly analogous to those involved in visual perception.

- (3) The opportunity of working with adult subjects makes it possible to combine a psycho-physiological study of this perceptual genesis with a phenomenological description of alterations in lived experience (Lenay and Steiner, 2010; Ward and Meijer, 2010).
- (4) The fourth advantage of using such technical mediation of perception is that it can simplify the repertoires of actions and sensory feedback which are available to the subject, and render them amenable to precise observation. We can then study, in each case, what objects can be constituted and what are the operations involved in this constitution.

In order to determine the minimal technical conditions which are necessary to enable the perception of an externalized object in a space where it can be localized, we have simplified the system of Bach-y-Rita to a single photoelectric cell connected to a single tactile stimulator. At each moment, the blindfolded subject thus receives only a minimal amount of information (1 bit) corresponding to the presence or absence of the tactile stimulation. We have been able to show that even with such a simple device, the spatial localization of luminous sources remains possible (Lenay et al., 1997). Here, it is manifestly clear that the perception cannot be based on an internal analysis of the sensory information, because this information has no intrinsic spatiality whatsoever. It is thus only through the movements of his/her exploratory activity that the subject can succeed in performing the perceptual task. By reducing the sensory input to a strict minimum, we force the subject to deploy a perceptual activity in the form of trajectories that can easily be observed and recorded. By construction, so to say, we adopt the theoretical framework of active perception (Varela, 1979; Gibson, 1986; Brooks, 1991; O'Regan and Noë, 2001). The spatial characteristics of an object are defined by the "Laws of Sensorimotor Contingency"; i.e., the laws which govern the sensory feedback as a function of the actions performed.

It is useful to go even further in the simplification by reducing the dimensions of action to two or even just one single dimension. On this principle, we have developed the "Tactos system" which will be used in the experiments involving social interactions that we shall present below (see **Figure 1**). This system consists essentially of a device for controlling tactile stimulators (Braille cells which electronically generate the movements of small pins) as a function of the movements of the cursor on a computer screen; the receptor field is guided by a pointing device (mouse, touchpad, graphic tablet, or tactile screen), and when it passes over a colored pixel it commands the activation of an all-or-nothing tactile stimulator placed under the finger (see Lenay et al., 2003 for details).

In these highly restrictive experimental conditions, it has been shown that the users (blind persons or blindfolded adults) can learn to recognize simple shapes. As explained above, the spatial perception of a shape is necessarily active because there is no intrinsic spatiality in the sensory input. This perception is thus realized essentially through a perceptual trajectory that can easily be recorded, analyzed, and modeled (Stewart and Gapenne, 2004). Of course, the space of all the motor commands that produce



**FIGURE 1 | The 'Tactos' system.** The shapes inscribed in the digital space on the screen are perceived in the tactile mode. The stylus of the graphic tablet controls the movements on the screen of a receptor field. When this receptor field encounters a black pixel, the software triggers an all-or-nothing tactile stimulus on the finger of the non-dominant hand. For the experimental studies presented here we use a single receptor field coupled to an activation of all the tactile pins of the Braille cells.

movements of the hand and arm is vast; but the relevant space of significant actions is defined by the interface, and boils down to translations of the receptor field in the space where the shapes are situated. Besides, one observes that during the course of learning the attention of the subjects, which was initially focused on the tactile stimulation, turns toward the space of two-dimensional action. It is in this space that the subjects situate themselves and act. An interesting consequence of this radical minimalism is that the perceptual trajectories for localizing or recognizing shapes, as well as the perceptual strategies that make it possible to carry them out, seem to be the same whether the sensory feedback is tactile, auditory, or visual (Gapenne et al., 2005). It therefore seems that the simplicity of the device makes it possible to elucidate some of the fundamental properties of the perception, independently of the sensory modality. Moreover, to the extent that the space that is explored is defined by a computer, the Tactos system makes it possible to set up a virtual space which can be shared by several users, even physically situated at a distance. It is this experimental setup that forms the basis for the experiments on social interaction that we shall now describe. In each case, this setup allows to clearly distinguish (a) the necessary and sufficient contributions of each individual partner; (b) that which emerges from the collective dynamics; and (c) that which can be subsequently appropriated by the individual partners.

## AN EXPERIMENTAL PARADIGM FOR THE STUDY OF THE RECOGNITION OF ANOTHER SUBJECT

Classically, in the framework of the philosophy of mind and the representationalist paradigm in cognitive science, one considers that the problem of the recognition of another subject comes down to the question of the adoption of an "intentional stance" with respect to the object in question (Dennett, 1971; Heider, 1982; Tremoulet and Feldman, 2000). In this framework, the question is thus to determine the criteria and mechanisms used by the subjects in deciding to treat the perceived objects, either as simple "things" which obey a mechanical causality, or else as "intentional agents" who act as a function of internal representations and goals. Various approaches are in competition, from the "simulation theory" (Meltzoff, 1995) which of late incorporates internal structures such as "mirror neurons" (Gallese et al., 2004) to the "theory theory" (Gergely et al., 1995; Csibra et al., 2003), passing

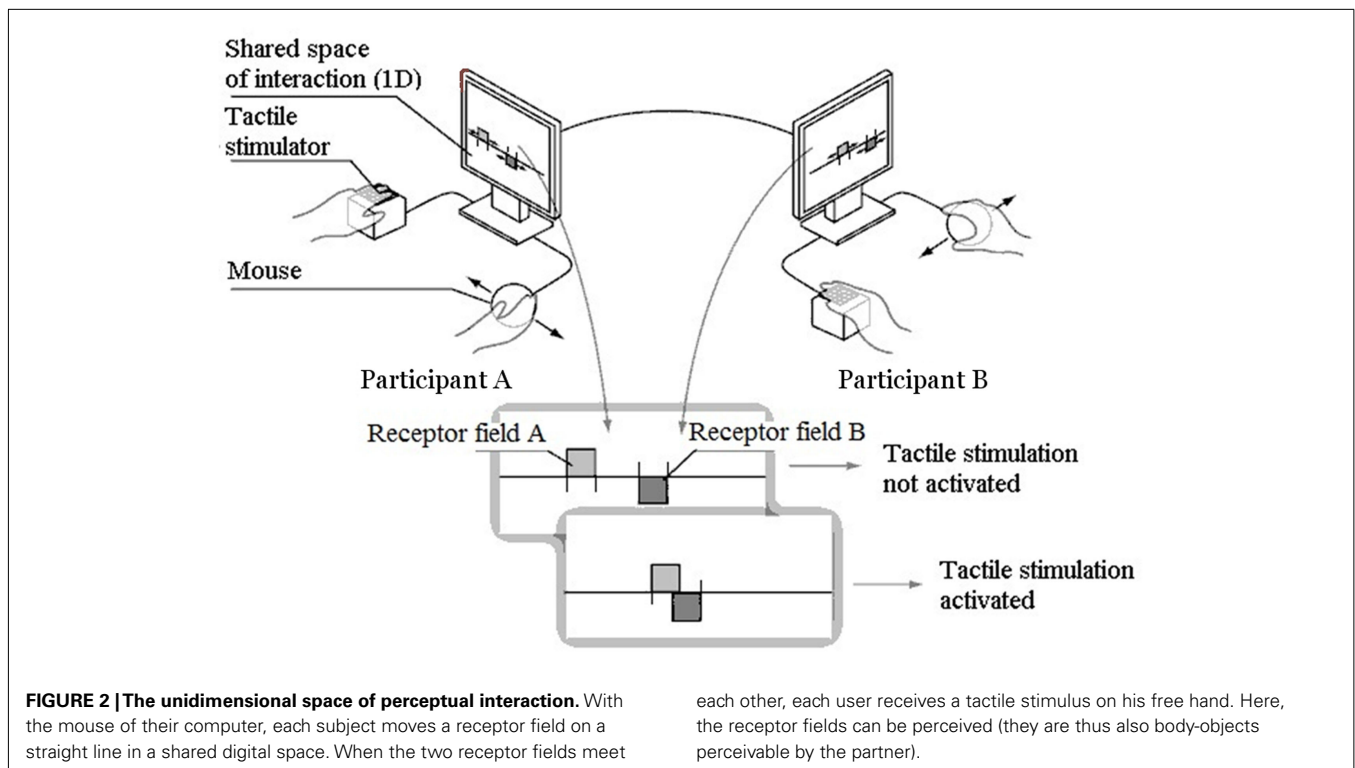
by the hypothesis of low-level perceptual modules (Leslie, 1987; Premack, 1990; Baron-Cohen and Cross, 1992; Povinelli et al., 2000; Tomasello et al., 2005). However, in spite of their diversity and above and beyond their oppositions, all these theories are based on the same type of experimental method. In all cases one establishes a strict separation between the observing subject and the scene that is observed.

By contrast, in an interactionist approach to this question the recognition of an intentional subject ought to take place during an interaction where the perceived subject can reciprocally recognize the observer himself as an intentional subject. We designate by the term “perceptual crossing” all those situations where two perceptual activities meet, as for example in mutual touching, looks where both subjects “catch each other’s eyes,” or a proto-conversation between mother and infant (Butterworth and Jarrett, 1991). The feelings of intimacy and the importance of the emotional values attached to this sort of inter-individual interaction are well known (Argyle and Dean, 1965). It is commonly reported that there is a feeling of immediate reciprocal recognition of the presence of another perceptual intentionality (Farroni et al., 2002). There is a question, however: when two subjects catch each other’s eyes, for example, is it because the subjects recognize each other as intentional subjects that they look at each other; or is it the other way round, because the looks are fixed on each other that there is reciprocal recognition as intentional subjects? (Baron-Cohen and Cross, 1992; Baron-Cohen et al., 2001) In the first case, each subject starts by unilaterally judging the presence of another subject on the basis of his behavior *before* entering into interaction, unless of course this interaction can supervene independently of any attribution of intentionality (Tomasello et al., 2005). In the second case, it is the perceptual interaction itself which produces

the mutual recognition. In this case, the situation of perceptual crossing makes it possible to discriminate the specificity of a perceptual activity directed at oneself. In order to give an empirical content to this intuition we have used the minimalist experimental paradigm described above, in a form which gives rise to an elementary sort of perceptual crossing. As explained above, this situation allows for a precise and detailed observation of the joint perceptual dynamics. An initial experimental study of this sort, which has already been presented elsewhere (Lenay et al., 2006; Di Paolo et al., 2008; Auvray et al., 2009), must nevertheless be presented in some detail here because it will serve as the basis for the following experiments.

**EXPERIMENT 1: STUDY OF PERCEPTUAL CROSSING**

In order to purify the notion of “perceptual crossing,” and to make it possible to proceed to a precise analysis of the mutual dynamics, we have reduced the space of action of the participants to a single dimension, and reduced also the repertoire of sensory input to a single all-or-nothing stimulation (just 1 bit of information at each time-point). Two blindfolded participants are placed in different rooms, and can only interact via the device. They each explore a computer screen with a mouse, and receive tactile stimulation on the index finger of their free hand. The movements of the mouse control the movements of a receptor field of 4 pixels in a one-dimensional space. Only the horizontal movements of the mouse are taken into account. The space of action consists of a straight line 600 pixels long, which loops round to form a continuous circle so as to avoid edge-effects. Various objects, consisting of black pixels, are placed on this line. Each time the receptor field encounters a black pixel, the participant receives an all-or-nothing tactile stimulation on the Braille cell (see Figure 2).



**FIGURE 2 | The unidimensional space of perceptual interaction.** With the mouse of their computer, each subject moves a receptor field on a straight line in a shared digital space. When the two receptor fields meet

each other, each user receives a tactile stimulus on his free hand. Here, the receptor fields can be perceived (they are thus also body-objects perceivable by the partner).

Two systems of this sort are connected in a network, so that the two participants share the same one-dimensional space. There are three sorts of objects that each participant can encounter:

- (1) The body-object of the other participant (his perceived body) which exactly matches his receptor field (4 pixels wide). When the two participants are in the same position, each receives an all-or-nothing tactile stimulation. We call this situation “perceptual crossing.”
- (2) A fixed object that we call the “fixed lure”: this is a segment 4 pixels wide. The fixed lure for the participant 1 is invisible for the participant 2, and is placed in a different position than the fixed lure for participant 2 (see **Figure 3**).
- (3) A moving object (4 pixels wide) that we call the “mobile lure.” In order to ensure that the mobile lure would have the same richness of movement as the body-object of the other participant, but without being responsive to perceptual crossings, we attached it by a rigid virtual link to the receptor field/body-object of the partner. The mobile lure thus follows exactly, but at a constant distance, all the movements performed by the partner. The lure was placed 50 pixels to the right of the receptor field (see **Figure 3**). In all that follows, distances between two objects are measured in pixels from the left-most pixel of one object to the left-most pixel of the other.

This experimental configuration makes it possible to test a theoretical hypothesis: even though the mobile lure and the body-object of the partner (which corresponds to his/her receptor field) have objectively exactly the same movements, will the participants be able to distinguish them on the sole basis that the receptor field of the partner is sensitive and animated by a perceptual activity turned toward their own movements?

Ten pairs of participants took part in this experiment. The participants were blindfolded and placed in different rooms. It is explained to them that the left/right movements of the mouse allow them to move in a shared one-dimensional space. In this space they can encounter three sorts of objects: a fixed object; a mobile object; and the body-object of their partner. The relation between the mobile object and the body-object of the partner is not explained to them. The instruction was to click on the left button of the mouse when they judged that they had met their partner. This experimental setup has a number of advantages:

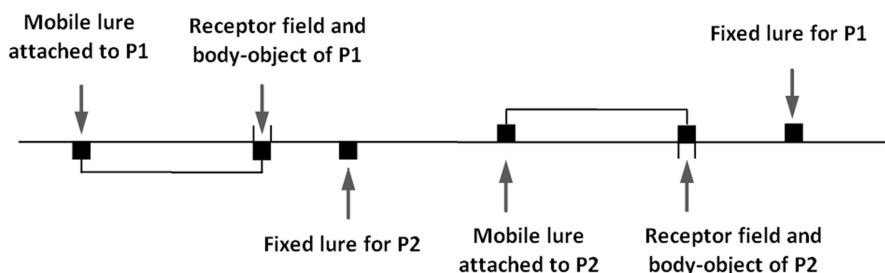
- (1) The perceptual situation is radically novel for the subjects. We thus avoid the direct importation of knowledge already elaborated. On the contrary, a learning period is necessary, and this makes it possible to observe the genesis of the phenomena.
- (2) The reduction of the sensory input forces a spatial and temporal deployment of the perceptual activities, and this makes it possible to record them and to analyze them in detail.
- (3) The simplicity of the setup makes it possible to elucidate the sufficient conditions for a detailed explanatory scheme of the collective dynamics, which we may hope has some generality.

The results for all the participants and all the sessions showed that the majority of clicks (62%) occurred when the two partners were indeed in front of each other, i.e., in a situation of perceptual crossing (see **Figure 4**).

We then analyzed the distribution of clicks as a function of the cause of the stimulations received by the participant during the preceding 2 s. The results over all the participants show that 66% ( $\pm 4$ ) of the clicks follow stimulations from perceptual crossing; 23% ( $\pm 10$ ) of the clicks follow stimulations due to the mobile lure; and only 11% ( $\pm 9$ ) follow stimulations due to the fixed lure. These results show that the participants are able to distinguish between the three categories of object that they encounter in the one-dimensional space. They distinguish between the receptor field of the partner and an object, be it fixed or mobile. This overall success may seem surprising since, by construction, the mobile lure has exactly the same movement as the receptor field of their partner. It seems that what is recognized is indeed the *activity* of a perceptual subject directed toward themselves, and not just the objective structure of the movements (Wilkerson, 1999). However, further analysis shows that this apparent success at the overall level masks what was actually a revealing failure at the individual level.

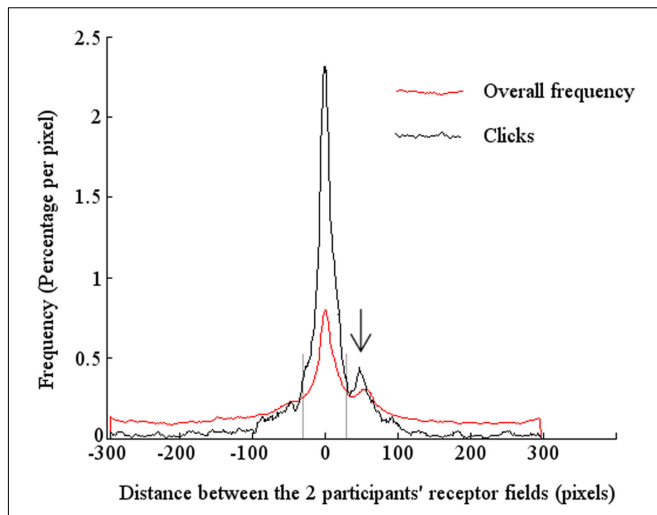
We first carried out a comparison between the distribution of the clicks and the distribution of the tactile stimulations received. The overall results for all the participants show that 52% ( $\pm 12$ ) of the stimulations come from a perceptual crossing, 33% ( $\pm 12$ ) come from the fixed lure, and only 15% ( $\pm 6$ ) from the mobile lure (see **Table 1**; **Figure 4**).

When we calculate the *ratio* of clicks/stimulations, we find 0.33 for the fixed lure, 1.26 for the perceptual crossing, and 1.51 for the mobile lure. These results show a major difference between the fixed lure on one hand (0.33) and the mobile entities on the other



**FIGURE 3 | Schematic illustration of the one-dimensional space explored by the subjects.** Subject P1 receives a tactile stimulus whenever (s)he encounters either his fixed object, or the receptor field of subject P2, or the mobile object attached to the receptor field of P2.





**FIGURE 4 | Distribution of frequencies as a function of the distance between the receptor fields of the two participants.** The fine line represents the total frequency of clicks made: 62% of the distribution lies between  $\pm 30$  pixels. The thick line represents the total frequency of stimulations received by the subjects: 28% of the distribution lies between  $\pm 30$  pixels. In both cases, there is a clear peak around the distance 0 pixels, i.e., the situation of perceptual crossing, which shows that there is an attractor at this point, at least in the weak descriptive sense that once the subjects have attained the situation of perceptual crossing they tend to remain in this stable dynamic configuration. A minor peak at the distance of 50 pixels (marked by an arrow) corresponds to the mobile lure.

**Table 1 | Distribution of the clicks and of the tactile stimulations – experiment 1.**

	Receptor field	Mobile lure	Fixed lure
Percentage of clicks	66 $\pm$ 4	23 $\pm$ 10	11 $\pm$ 9
Percentage of stimulations	52 $\pm$ 12	15 $\pm$ 6	33 $\pm$ 12
Ratio clicks/stimulations	1.26	1.51	0.33

(1.26 and 1.51). The participants have a probability of clicking that is four times greater if the stimulation comes from a mobile entity than it is due to the fixed lure. Thus, the ratio between clicks and stimulations shows that overall each participant does not seem to distinguish between stimulations due to perceptual crossing and stimulations due to the mobile lure (1.26 vs. 1.51). The difference in clicks on the mobile lure and on the receptor field of their partner (23 vs. 66%) seems to be due only to strategies of movement which are such that encounters with the mobile lure are much less frequent than encounters due to perceptual crossing (15 vs. 52%). If the participants succeed in the task, it is essentially because they succeed in situating themselves face-to-face with their partner, and not because they recognize in the pattern of stimulation any clues which discriminate the receptor field of their partner from that of the mobile lure. The only difference resides in the interaction itself. In order to account for these results, there are two things to be explained. On the one hand, the capacity of the participants to privilege the situation of being face-to-face; on the other hand, the reasons that leads them to click.

**Attractor in the collective dynamics**

We may note that all the observations made with these minimalist setups show that the perception of an object in a particular position is realized by its active, reversible exploration: the subjects come and go around the singularity that provokes a sensory return (Sribunruangrit et al., 2004). Thus, there is a general strategy which consists of reversing the movement of the receptor field following a sensory event. To the extent that the perceptual strategy of each participant consists of inverting their movement following an alteration in sensory input, if a participant meets their partner (s)he will invert his movement while the latter will do the same. The two receptor fields will thus enter into a sort of dance. This can be described as constituting an attractor in the collective dynamics; an attractor which is not a spatially fixed point, but a region which may itself be displaced. Even though the participants do not have a specifically collaborative aim, their simultaneous efforts to discriminate the presence of their partner produces an attractor in the collective dynamics of their perceptual activities (Froese and Di Paolo, 2010, 2011a).

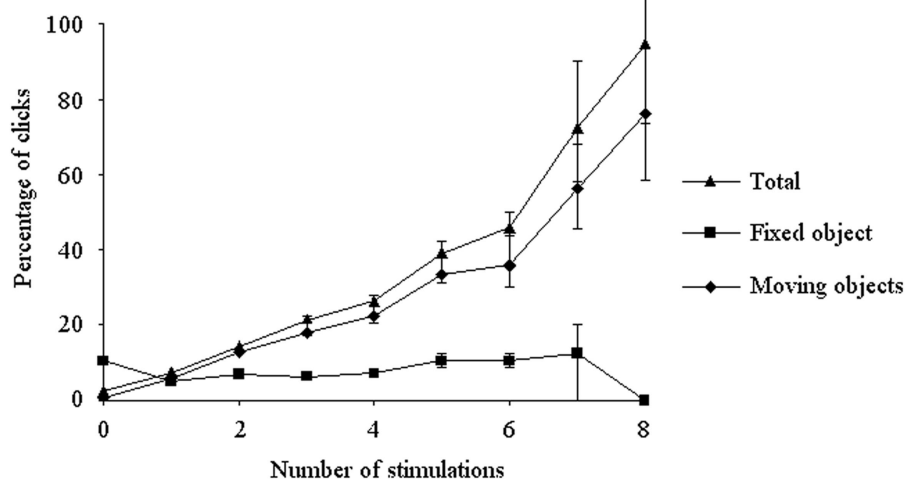
**The reasons that lead the subjects to click**

If we study the events which precede each click, we observe that if over the last 2 s of his perceptual activity a subject meets:

- (1) few stimulations, no perception is constituted and the probability of clicking is low;
- (2) many stimulations, but for an object that is recognized as fixed (sensorimotor stability), the probability of clicking is again low;
- (3) but if there are many stimulations, for an object that remains undetermined spatially, the probability of clicking is high (see Figure 5).

In the latter case, the participant is probably in the presence of the other participant, but it is also possible that it is the mobile lure. Thus the clicks of the participants can be largely explained by the conjunction of two criteria, one negative and one positive:

- (1) “Another subject” is something which resists precise spatial determination: it is neither a fixed object, nor an object with movements determined by a simple rule.
- (2) However, at the same time, “another subject” is something which maintains its presence. This is indeed a characteristic of the body-object of another participant, but not of the mobile lure, because it is only this body-object which has a receptor field sensitive in its turn to the presence of objects, i.e., likely to change its behavior according to the sensory input it receives. The (only) difference between the receptor field of the other participant and the mobile lure attached to it is that only the former is sensitive to my presence; and as we have seen, this sensitivity is linked to a perceptual intentionality which constantly aims at remaining in the vicinity of a singularity. This is precisely a sufficient condition for the formation of an attractor in the joint dynamics which tends to augment the probability that the partner will be present. Thus, the criterion which seems to be employed by the participants for clicking is not arbitrary, but ensues logically from the meeting



**FIGURE 5 | Probability of clicking.** The probability of participants' clicks, plotted as a function of the number of distinct stimulations received during the preceding 2 s. Lozenges: total stimulations (body-object, 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.

of two perceptual intentionalities. This criterion is coherent with the very content of that which is to be recognized. The other subject is recognized just as something that resists its precise determination and yet which persists in being present. An analysis of the phenomenological descriptions given by the subjects themselves would exceed the limits of this article. It is sufficient to note here that the modifications of the lived experience of the subjects can only be built on the basis of the objective elements that we present.

However, even if these criteria seem judicious, they are not sufficient here to guarantee against a failure to distinguish between the receptor field of the partner and the mobile lure. If by a stroke of bad luck it is the mobile lure which remains present, the participants have the same probability of clicking as for the receptor field. For example, if the receptor field of my partner is engaged in oscillating around an object situated at 50 pixels from my own position, and thus causing a movement of the attached lure around my own position, I will be induced to click on the attached lure.

### Conclusion to experiment 1

In this first experiment where the aim is to discriminate the presence of another subject, the individuals fail whereas the collective action succeeds. Thus, the collective success cannot be explained by an individual capacity to recognize another subject by means of a particular sensation (Michael and Overgaard, 2012). The collective success is explained principally by a collective dynamics which results from the engagement of each subject in his perceptual activity searching for a partner. The clicks result from a decision rule which appears to be judicious, but which is insufficient at the individual level to distinguish between specific sensations.

The question which arises now for an interactionist approach is whether it is possible for the individuals to appropriate the collective success. This seems feasible if we relax a particularly

unrealistic condition of our experimental situation in which there was no intrinsic difference between the various objects. If the subjects are able to recognize different intrinsic properties for the three objects, they may be able to use these properties to categorize the different situations of interaction in which they are engaged. In order to show that this is the case, we have carried out a new experiment with a protocol that is very similar, but which this time consists of categorizing the objects (fixed lure, mobile lure, and receptor field) that can be easily discriminated in their own right.

### EXPERIMENT 2: RECOGNITION OF PERCEPTUAL CROSSING

The experimental setup is the same as that of the first experiment, except that this time the sensory feedback is no longer a tactile stimulus, but a sound which is different for each of the objects which can be encountered.

#### Method

Twenty participants took part in this experiment. Their ages ranged from 20 to 32 years (mean age of 22.4 years). All of the participants reported normal tactile perception. The experiment took approximately 25 min to complete and was performed in accordance with the ethical standards laid down in the 1991 Declaration of Helsinki.

We used an adapted version of the minimalist sensory substitution system Tactos. Blindfolded participants explored graphic information by means of a computer mouse and received auditory information via headphones. The displacement of the computer mouse produced the displacement of a 4-pixel receptor field in a one-dimensional space (a line 600 pixels long, with the ends joined to form a torus). Only the horizontal displacement of the mouse was taken into account. Several objects consisting of black pixels were situated on this line. Each time the receptor field covered a black pixel, a sound is emitted which varies according to the nature

of the object. There are three possible sounds: (1) the horn of a car, (2) the horn of a big lorry, and (3) the tinkling of a bicycle bell. These sounds were chosen to be easily differentiated and named. Two Tactos devices were combined in a network so that each pair of participants shared a common one-dimensional space. As in the previous experiment, each participant could encounter three types of object:

- (1) The 4-pixel receptor field of the other participant.
- (2) A fixed 4-pixel wide object. The fixed object perceived by participant 1 was placed between 148 and 152 pixels and was invisible for participant 2; the fixed object perceived by participant 2 was placed between 448 and 452 pixels and was invisible for participant 1.
- (3) A mobile 4-pixel wide object. In order to ensure that the movements of this object have exactly the same dynamic structure as the movements of a receptor field, two conditions were tested:
  - (i) *Condition C1*: the mobile object was attached by a virtual rigid link at a distance 100 pixels from the center of the receptor field (see **Figure 3**). It should be noted that when participant 1 explored participant 2's mobile object, participant 2 did not receive any auditory feedback (and conversely, participant 1 did not receive any auditory feedback when his mobile object was explored by participant 2). In contrast, when one of the participants explored the other participant's receptor field, both received auditory feedback.
  - (ii) *Condition C2*: just like C1 except that the mobile object is animated by the trajectory of the partner recorded from the previous session.

### Procedure

Each pair of participants performed the experiment once. They were blindfolded and seated in different rooms in front of the Tactos device. There was no means of communicating between the participants other than the experimental setup. The functioning of the device – the relation between the receptor field, the objects in the environment, and the auditory feedback – was explained to the participants. The participants were then trained on the device during three phases of 1 min each: exploration of three fixed objects to which the three sounds were attributed (we verified that the sounds were clearly differentiated); exploration of an object 4 pixels wide moving at a constant speed of 15 pixels/s; then at 30 pixels/s. Then, the experimental task was explained to the participants. They were told that they could freely explore the one-dimensional space containing three types of auditory object: (1) the receptor field of the

partner, (2) fixed objects, (3) and mobile objects. However, the nature of the dynamics of the mobile object was not explained. The instruction was to associate a sound to each of the three types of objects.

There were four sessions of 2 min each. In sessions S1 and S2 the condition was C1 (attached lure), in session S3 and S4 the condition was C2 (the lure follows the recorded movements of the partner in session S2). A sound was associated with each object; the sounds are reattributed differently and randomly for each session. At the end of the four sessions the strategies and impressions of each participant are noted and recorded.

### Results

First of all, we looked at the frequency of correct responses for each of the three objects (see **Table 2**).

For the set of all 80 sessions (considering each participant independently), 60 were perfect. There was an improvement between sessions 1 and 4. This may be explained by an effect of learning with respect to the setup, but also with respect to the behavior of the partner which stays the same. This learning effect masks any possible difference between conditions C1 and C2. The ease of identifying the fixed object is confirmed by a success rate of almost 94%. If we consider that this recognition is generally achieved, the results for the categorization of the other two objects remain largely significant, especially in the fourth session where the success rate for the 20 participants is 85% compared to the chance rate of 50%.

We then examined the relation between the success of one participant and the concomitant success of the partner. Of the 60 sessions which were perfect for one participant, 46 (77%) were also perfect for the other participant, whereas by chance there would have been 35 (59%;  $\chi^2 = 3.46, p < 0.07$ ).

### Analysis and conclusion for experiment 2

After self-learning, the results of this complementary experiment are significantly in favor of a good categorization of the sound corresponding to the presence of the partner. There is no significant difference between the conditions C1 and C2 (attached lure or recorded lure). The analysis of the behavioral trajectories, and the questions posed at the end of the experiment, allow us to elucidate the strategies of the participants.

Almost all the participants adopt a “sweeping” technique when they encounter an object, i.e., they oscillate around the position where they perceived an object. A large majority of the participants first sought to identify the fixed object ( $19 \pm 18\%$  of the positions are concentrated between +30 and -30 pixels from the fixed lure, out of the 800 pixels of the total space), and then they

**Table 2 | Frequency of correct responses – experiment 2.**

	Partner (%)	Mobile object (%)	Fixed object (%)	Perfect 3/3 (%)	$\chi^2$ test (%)
Success over all four sessions	77.50	75.00	93.75	75.00	2.53 (significant)
Success session S1 (C1)	65.00	65.00	95.00	65.00	17.97 (not significant)
Success session S2 (C1)	75.00	75.00	95.00	75.00	2.53 (significant)
Success session S3 (C2)	85.00	75.00	90.00	75.00	2.53 (significant)
Success session S4 (C2)	85.00	85.00	95.00	85.00	0.17 (highly significant)



sought to remain in contact with a mobile object. When two participants meet, they “sweep” on each other and seek to remain in contact. In this way, we find again a dynamic attractor around the position of perceptual crossing, although it is somewhat less marked than in the previous experiment ( $21 \pm 8\%$  of the positions are concentrated between  $+30$  and  $-30$  pixels from the position of the receptor field of their partner). When that succeeds sufficiently, they decide that it is their partner. When they fail to follow a mobile object, they decide that it is a lure which is indifferent to their presence ( $16 \pm 7\%$  between  $+30$  and  $-30$  pixels around the mobile lure).

As in the first experiment, the collective dynamics results from the common engagement of both participants in the perceptual activity. We may note in addition that the success of one participant has an influence on the success of his partner. This can be easily understood, since the two participants are engaged in the same perceptual task. If for example one of the participants does not move his receptor field, his partner will have no means of distinguishing it from the fixed lure.

What explains the individual success here is the ability of the participants to distinguish the dynamics of perceptual crossing from the dynamics of interacting with other objects. Since the participants have access to different intrinsic properties for the three objects, they can recognize *different occurrences of the same object*. Thus, the different sounds make it possible to disambiguate the situations of interaction: faced with a lure which presents the same criteria which led the participants to click in the first experiment (frequent sensory stimulation combined with an indeterminate position), the participants no longer make the mistake. The difference in the intrinsic properties allows them to recognize that this situation is not the same as the encounter with the other participant, which is more frequent because it corresponds to a stronger attractor in the dynamics of interaction. Thus, the situation of perceptual crossing is now recognized as a property of an object already identified by other means. We shall return to this point in the conclusions. Now that we have been able to define the collective dynamics of perceptual crossing (see Attractor in the Collective Dynamics), we may enquire whether this form of synchronization may make it possible to propose an original approach to the phenomena of imitation of facial expressions.

## PROCESSES OF IMITATION

It is indeed difficult to understand how, just after birth, an infant can establish a relation between the movements observed in a conspecific subject and the proprioceptive data concerning her own movements, in particular her *opaque actions* such as her own facial movements that she cannot see herself (Meltzoff and Moore, 1997). Even though some authors are skeptical of the new-born imitation data (Ray and Heyes, 2011), the phenomenon remains interesting; and the importance of this *correspondence problem* (Brass and Heyes, 2005) is not limited to imitation but applies also to action understanding (Rizzolatti and Craighero, 2004), the use of tools (Ferrari et al., 2005), empathy (Gallese et al., 2004), learning a language (Rizzolatti and Arbib, 1998), and the “theory of mind” (Gallese, 2007). In the literature there are two main positions concerning the underlying mechanisms.

The first position consists of postulating an innate “Active Intermodal Matching” system (AIM) which performs a supra-modal representation of bodily actions which are seen or performed (Meltzoff and Moore, 1999; Nagy, 2006). In this case, an “innate mirror-neuron system” participates in the neuronal cabling between perceived facial expressions and the expressions that are produced (Iacoboni et al., 1999; Rizzolatti et al., 2002). However this solution, as indeed all “hereditarian” solutions in general, does not really explain anything at all, but consists of merely giving oneself a phenomenon whose genesis remains to be explained. If the question is that of explaining imitation, we have to show that it can occur without any prior knowledge of it.

The second position consists of postulating a “learning” of this matching between action and perception (Butterworth, 1999). Sensorimotor training is supposed to configure these internal structures by setting up an association between representations of the actions and representations of their sensory consequences, in particular visual consequences (Catmur et al., 2007, 2009; Cook et al., 2010). If the question is that of explaining imitation, we have to show that it can occur without any consciousness of imitating.

In both cases, imitation is postulated as being effected by structures that are internal to each individual, principally “mirror neurons.” However, if one was able to account for a phenomenon that appears as “imitation” without appealing to such internal structures, one would start to have the means to *account* for the setting up of such structures, whether it be through individual learning or by an evolutionary process. The path we propose to explore here consists of seeking the conditions for the appearance of “mimetic phenomena” in the very dynamics of the perceptual interactions – and this in the absence of any previous internal knowledge of the subjects concerning their own facial expressions. It is not the imitation which accounts for the interactions, but the dynamics of interaction which produces the imitation. Here again we propose a particular experimental study which makes it possible to elaborate a conceptual scheme, whose generality will of course have to be examined subsequently.

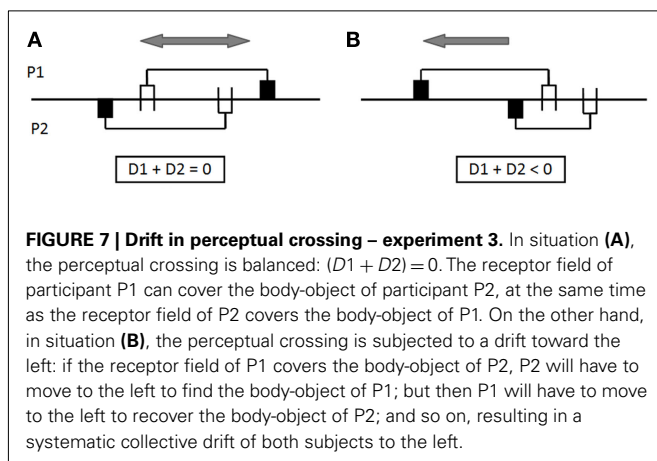
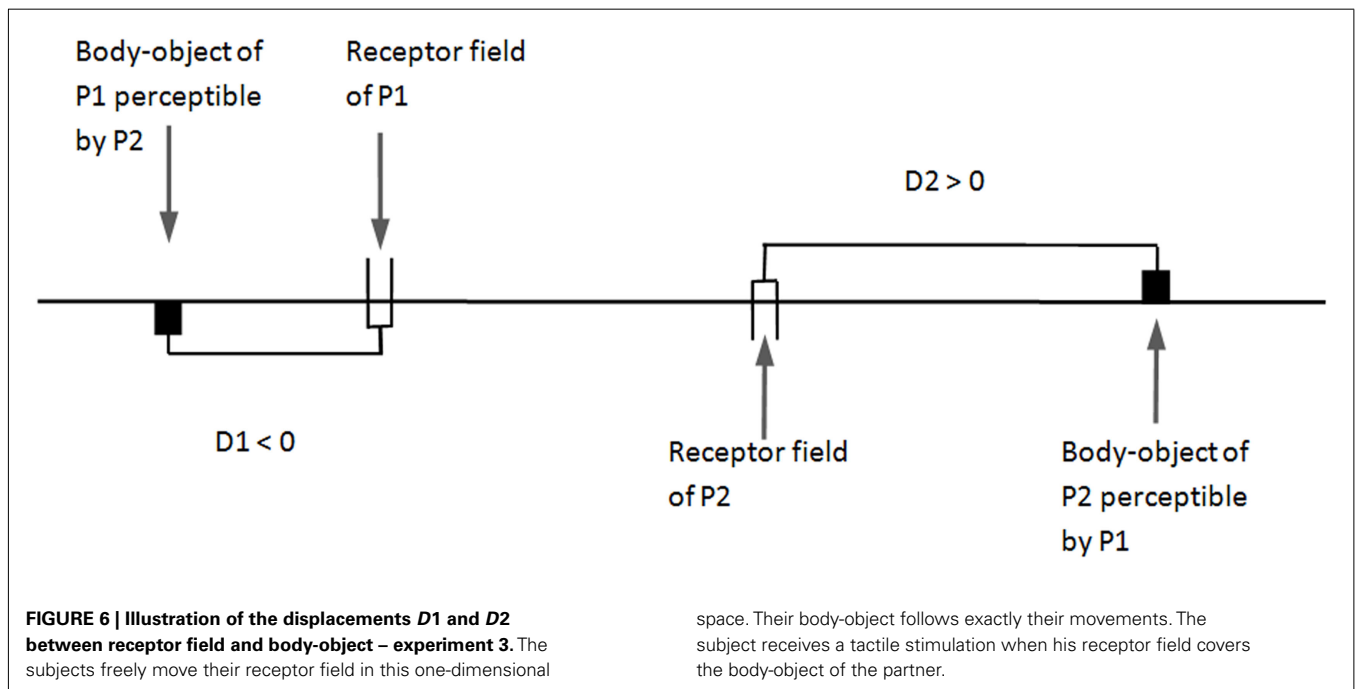
## EXPERIMENT 3: MIMETIC DYNAMICS IN THE PERCEPTUAL CROSSING

We have thus taken up our experiment of minimalist perceptual crossing; but this time, the participants can modify what is presented to their partner. In accordance with our minimalist approach, we have chosen as a minimal modification of the body the relative distance between the body-object and the receptor field. The objective external description of “imitation” will be a similarity in the relative distances of the body-objects of the two subjects, relative distances that the subjects themselves do not perceive. If the subjects do succeed in matching these distances ( $D1$  and  $D2$ , see below), this will illustrate our contention that “imitation” as such is largely in the eye of the beholder.

### Methods

The experimental setup is the same as that of the first experiment, except that this time, there is no fixed lure, and the receptor field is no longer directly perceivable by the partner. All that is perceivable is the body-object that is attached to the receptor field.

We call  $D1$  the position of the body-object of participant 1 with respect to his receptor field, and  $D2$  for participant 2; the



algebraic values of  $D1$  or  $D2$  are positive if the body-object is to the right of the receptor field, negative if it is to the left (see Figure 6). If  $D1 + D2 = 0$ , when the receptor field of P1 is exactly in front of the body-object of P2, the receptor field of P2 is also directly in front of the body-object of P1. In this configuration, there should be no problem for achieving perceptual crossing since the two partners perceive each other mutually at the same time. However, if  $D1 + D2 < 0$ , the perceptual crossing is unbalanced, each participant moving to the left to find his partner. Similarly, if  $D1 + D2 > 0$ , the perceptual crossing should drift to the right (see Figure 7).

The distance between the receptor field and his own body-object can be actively modified by the participant. By clicking on the right or left button of the mouse, the participants can displace their body-object to the right or to the left relatively to their receptor field, 2 pixels at a time for each click. However, they do not know the initial position of their body-object, and they cannot perceive the receptor field of their partner.

The body-objects and receptor fields all have a width of 8 pixels. The displacement of the computer mouse produces the simultaneous displacement of the 8-pixel receptor field and the 8-pixel body-object in a one-dimensional space. Only the horizontal displacement of the mouse was taken into account. The one-dimensional space consisted of a line 800 pixels long, with the ends joined to form a torus in order to avoid singularities due to edges. Each time the receptor field encounters a black pixel of the body-object of his partner, the participant receives an all-or-nothing tactile stimulation on the Braille cell.

**Procedure**

Twelve participants took part in this experiment. Their ages ranged from 18 to 32 years (mean age of 20.4 years). All of the participants reported normal tactile perception. The experiment took approximately 35 min to complete and was performed in accordance with the ethical standards laid down in the 1991 Declaration of Helsinki.

The participants are blindfolded, placed in different rooms, and able to interact only via the device. Each pair of participants performed the experiment once. The functioning of the device – the relation between the receptor field, the objects in the environment, and the tactile feedback – was explained to the participants. During a learning period (with  $D1 + D2 = 0$ ), the participants learned to maintain the situation of perceptual crossing. The explicit instructions are the following: the participants must be attentive to the possible drift of their perceptual crossing, and that by clicking they can restore the balance. One informs them that if they feel that the drift occurs toward the right they should click on the left button, and vice versa. The experiment was performed over four sessions of 3 min each, with different starting conditions:

- S1:  $D1 = +30, D2 = +30$  thus  $D1 + D2 = 60$
- S2:  $D1 = +16, D2 = +30$  thus  $D1 + D2 = 46$

S3:  $D1 = +30, D2 = -30$  thus  $D1 + D2 = 0$   
 S4:  $D1 = -16, D2 = +30$  thus  $D1 + D2 = 14$

The participants are instructed to maintain the perceptual crossing as long as possible. They do not know the position of their own body-object, nor that of their partner. Each participant clicks left or right according to his own feeling concerning the drift of the perceptual crossing.

**Results**

Overall, there was clearly a convergence toward the situation where  $D1 + D2 = 0$ , that we may identify as a situation of imitation. Even though the participants do not know the position of their body-object, either at the beginning or at the end, their joint search for a situation of balanced perceptual crossing rapidly leads to a similarity in these positions. In 3 min, the disequilibria in  $D1 + D2$  are reduced to less than 30% of their initial values (see Table 3; Figure 8).

At the same time the diversity of the actual values for  $D1$  or  $D2$  increases over time [the standard deviation of  $(D1 - D2)$  between pairs of participants passes from 0 to 16 pixels]; this is understandable, since the positions of equilibrium that are sought belong to an infinite class of situations where  $D1 + D2 = 0$ . Even in S3,

where the initial position was already perfectly balanced, there is a differentiation of the situations of equilibrium.

The condition S4 is also interesting because, given the width of the receptor fields and body-objects (8 pixels), the participants could have satisfied themselves with a state of equilibrium where both participants remained immobile while receiving stimulation. Nevertheless, what is observed is a continuation of the process of convergence toward a better imitation (decrease in  $D1 + D2$ ). This initial experiment thus enabled us to test our hypothesis: at least in these experimental conditions, the collective dynamics leads to a stabilization of a phenomenon of imitation.

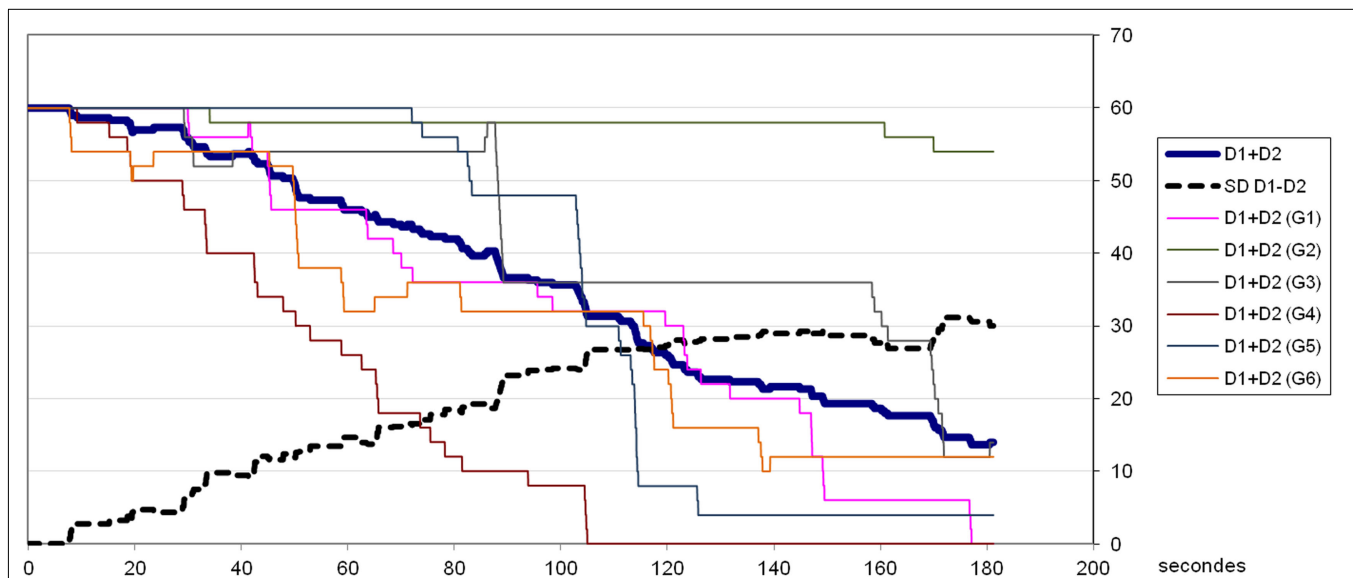
**Analysis**

In order to understand how the participants manage to succeed in this task, we can come back to the analysis of the perceptual trajectories and sensory feedbacks, which represent all that the participants have access to. We will then attempt to determine the strategies adopted by the participants, whereby they link variations in their sensory input to their subsequent actions. In Figure 9, we illustrate the existence of an attractor in the relative positions,  $X1-X2$ .

Even when  $(D1 + D2)$  is relatively large, so that the participants cannot both perceive each other simultaneously, the dynamics of

**Table 3 | The initial and final values of  $(D1 + D2)$ , measured in pixels – experiment 3.**

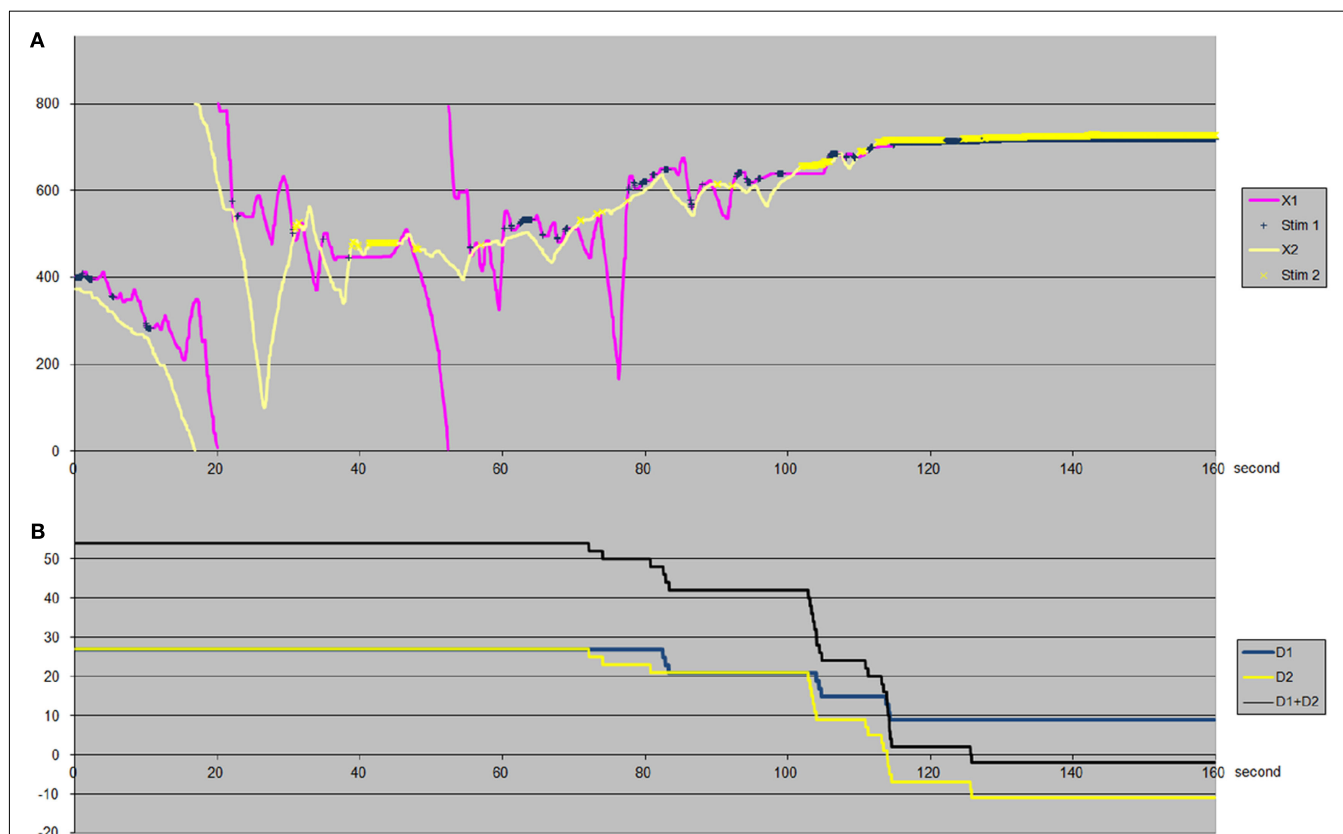
Experiment	Initial state		Average final state	
	$D1 + D2$	$D1 + D2$	Standard deviation $D1 + D2$	Standard deviation $D1 - D2$
S1	60	13	$\pm 20$	$\pm 30.6$
S2	46	10	$\pm 14$	$\pm 23.3$
S3	0	-1.6	$\pm 4.8$	$\pm 1.8$
S4	14	5	$\pm 1.7$	$\pm 8.8$



**FIGURE 8 | Results of the first experimental session – experiment 3.**

Legend. The fine lines represent the evolution of  $(D1 + D2)$ , i.e., the distance between the body-objects of the six pairs of subjects, over the course of the

3 min of the session. The thick line represents the evolution of the mean of  $(D1 + D2)$ . The dotted line indicates the dispersion of the values  $D1$  and  $D2$  (the standard deviation of the difference  $D1 - D2$ ).



**FIGURE 9 | An example of interaction trajectories.** Time is indicated in seconds on the abscissa. **(A)** The trajectory of participant 1 (X1) is in blue, that of participant 2 (X2) in yellow (breaks in the trajectories correspond simply to passages in the torus). Stimulations received are marked by crosses on the trajectory. **(B)** The displacement between receptor field and body-object for participant 1,  $D1$ , is in thick yellow, and that of participant 2,  $D2$ , in thick blue. The sum of the displacements ( $D1 + D2$ ) is

indicated by a black line. It can be seen that at the start the participants drift toward the right (the bottom of the graph), then from  $t = 55$  s they start to drift toward the left (the top of the graph) but slower and slower as they progressively stabilize. At the start the displacement ( $D1 + D2$ ) is 60 pixels ( $D1 = D2 = 30$ ). The participants start clicking from  $t = 72$  s. From  $t = 125$  s onward, the two subjects both receive a continuous stimulation and they stop moving: ( $D1 + D2$ ) = 4 pixels.

interaction still exhibits a sort of perceptual crossing in the form of a mutual oscillation of the participants around each other. This “attractor” can be characterized by the standard deviation of the distribution of distances between the participants<sup>1</sup>. As shown in **Figure 10**, this attractor becomes narrower as the values of ( $D1 + D2$ ) decrease; the participants are more and more often in front of each other and their movements are more and more reduced; the correlation coefficient of 0.52 is highly significant ( $p < 0.1$ ). When ( $D1 + D2$ ) is less than 16, so that the two participants could stop moving in a situation where they both receive a stimulation, it is striking to note that most often their activity continues and, on the average, the attractor of the perceptual crossing narrows still further. In the fourth session, ( $D1 + D2$ ) decreases to  $5 \pm 1.7$  pixels. Here, it is clear that it is only with respect to the dynamics of their interaction that the participants can grasp whether or not there is a drift in their perceptual crossing, and seek a situation with a well-balanced face-to-face.

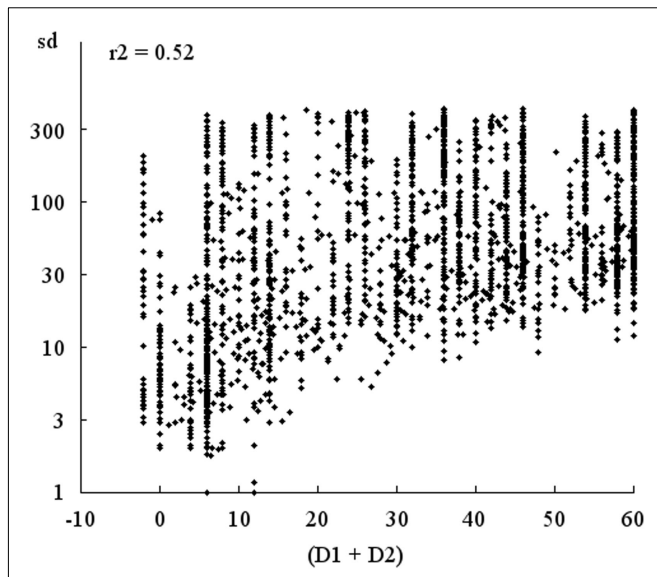
<sup>1</sup>In the experiments illustrated in **Figures 10, 11, and 12**, the data making up the distributions were collected at regular 1-s intervals over the course of the experiments.

Viewed from the outside, the actions of the participants produce a tightening of the attractor in their dynamics of interaction. The question arises as to the clues that the participants may use to guide their adjustments of  $D1$  and  $D2$ . As we have already seen in the first experiment, the participants seem to be sensitive to the frequency of stimulations received whilst they seek to establish a perceptual crossing. As shown in **Figure 11**, a decrease in ( $D1 + D2$ ) is accompanied by an increase in the frequency of stimulation: the correlation coefficient of  $(-0.62)$  is highly significant.

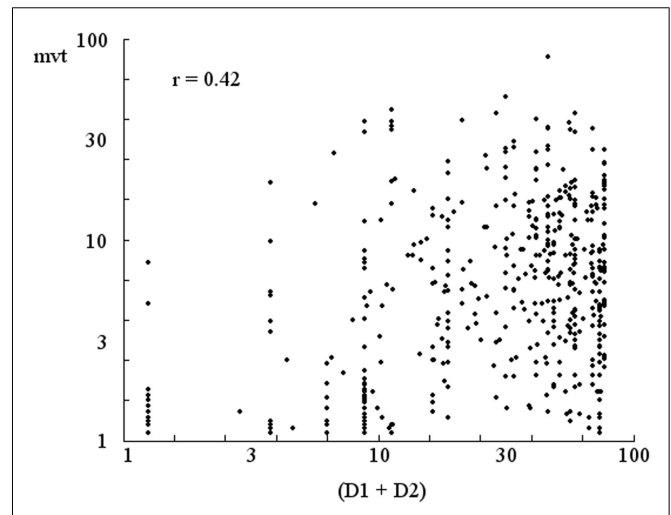
Moreover, as shown in **Figure 12**, the participants may also be sensitive to the systematic drift in their average positions over a 5-s period. The correlation coefficient of 0.342 is highly significant.

From the point of view of each participant, the value of ( $D1 + D2$ ) defines a situation of interaction which leads to a certain speed of the drift of the perceptual crossing, and to a certain frequency of sensory stimulations. Conversely, this speed of the drift and changes in the frequency of stimulations can serve as a clue to click and so to modify the value of ( $D1 + D2$ ).

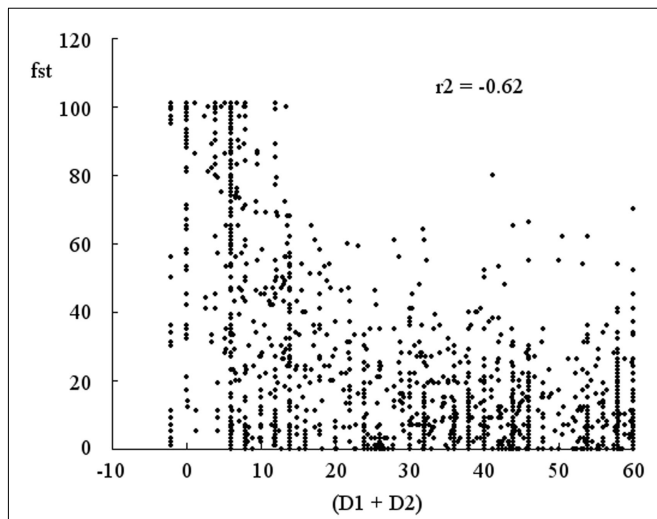
In all the sessions, both participants are necessarily active in moving to obtain sensory stimulations. However it happens quite



**FIGURE 10 | Correlation between the width of the attractor “SD” and (D1 + D2).** “SD” is the standard deviation of the distance of the perceptual field from the center of the point of stimulation, i.e., the body-object of the other subject. As the disequilibrium (D1 + D2) decreases over time, “SD” also decreases, i.e., the attractor shown in **Figure 4** becomes narrower. The correlation coefficient of 0.52 is highly significant.



**FIGURE 12 | Correlation between the speed of the drift of the perceptual crossing and (D1 + D2).** The rate of drift of the mean position over a 5-s interval, “mvt,” decreases as the disequilibrium (D1 + D2) decreases. The correlation coefficient of 0.342 is highly significant.



**FIGURE 11 | Correlation between the frequency of stimulation and (D1 + D2).** The frequency of stimulation, “fst,” increases as the disequilibrium (D1 + D2) decreases. The correlation coefficient of -0.62 is highly significant.

**CONCLUSION ON EXPERIMENT 3**

In this experiment, two dynamics are coupled: a rapid perceptual dynamics of the movements of the receptor fields controlled by movements of the computer mouse; and a slower dynamics, corresponding to modifications of the distance “D” (between receptor field and body-object) which is controlled by left and right clicks on the mouse button. We see that this second, slower dynamics is controlled by the results of the first, rapid dynamics. In his rapid perceptual dynamics, each participant makes an effort to find and to sustain a good perceptual crossing (sweeping movements around the body-object of the partner). The participants reveal that they are able to perceive the orientation of a drift that they are *subjected to*. Even when the perceptual crossing is perfectly balanced (D1 + D2 = 0), the participants can move together in one overall direction or the other; but here, the participants seem to perceive that this drift has a “force,” a “systematicity,” that they can correct by clicking. The clicks of both participants act on a common spatial variable, the relative distance (D1 + D2), which determines the balance of the perceptual interaction. By bringing this common variable to 0, they produce a stabilization of the perceptual crossing which, from the point of view of an external observer, corresponds to a mirror-resemblance of the images that are presented to the partner (D1 = -D2).

Of course it is a long way from this radically simplified situation, to those of natural multimodal encounters. We shall come back to this point in the final discussion. The point that seems to us important here, and that the experimental setup aimed at showing, is that the adjustment between the two participants occurs even though they do not know what image they present to their partner, nor what is the exact effect of their actions (the mouse clicks). They only have access to the collective dynamics, and it is through this that they guide their actions. Here, imitation does *not* result from learning the relations between what is perceived of another subject (visual perception of facial expressions) and what

often, in one-third of the sessions (8 out of the 24), that only one of the participants clicks (thus changing only D1 or D2 as the case may be); the other participant is active only in maintaining the perceptual crossing. Such a differentiation in the roles is possible because the functionally significant variable is actually the sum (D1 + D2), and each participant can act alone on this variable.



is perceived of one's own actions (proprioceptive perception of one's own expressions).

## GENERAL CONCLUSION AND DISCUSSION

### MINIMALISM AND TECHNICAL MEDIATION OF PERCEPTUAL ACTIVITY

The object of the experiments presented here was to create the empirical conditions for a theoretical discussion by reducing conceptual ambiguities to a minimum. The various individual and collective components of the observed phenomena can be clearly distinguished, and sufficiently complete explanations can be proposed. To achieve this, the technical mediation serves as a prism which makes it possible to separate out and to analyze the components of complex interactive processes. By reducing the sensory information to 1 bit of all-or-nothing information, and by reducing the actions to movements on a one-dimensional line, the perceptual activities were externalized in the form of perceptual trajectories which can be easily recorded, permitting a complete analysis of the sensorimotor relations.

A large number of other studies are currently under way, using the same sort of deliberate minimalism. For example, we have verified that the dynamics of perceptual crossing remains essentially the same if the space of actions is two-dimensional (rather than the one-dimensional space used here; Lenay et al., 2011). In this framework, one can also study differentiation of the roles of the two partners following variations in the relative sizes of their perceiving bodies (receptor fields) and their perceived bodies (body-objects; Maillet et al., 2008). A similar experimental situation is also being used to carry out "Turing-test" experiments, where the participants have to discriminate between a human partner and automatic robots of increasing complexity (Deschamps et al., 2012). Perceptual crossings in one- and two-dimensional spaces have also been studied using the methods of evolutionary robotics, which makes it possible to explore the field of possible solutions (Di Paolo and Iizuka, 2008; Froese and Di Paolo, 2008; Rohde and Di Paolo, 2008). A study of the same type has also been carried out for the minimalist imitation experiment presented here (Froese et al., submitted).

Now it may be objected that this minimalism only accounts for an artificial perception, widely removed from natural perceptual functions, and so it does not teach us much about the natural situations. Our reply is that the constraints of minimalism make it possible to clearly control what was absent at the start, and so was constituted during the course of the experiment. Even if it is limited, there is nevertheless a genuine genesis of social cognitive capacities. The explanatory scheme that we propose for this particular situation may then serve as a model, as a tool. In fact, we consider that the boot is on the other foot: if other authors wish to maintain that other mechanisms are necessary to account for imitation in natural situations, it is up to them to demonstrate clearly the existence of such mechanisms – preferably in suitably minimalist experimental conditions.

### DYNAMICS OF INTERACTION AND INDIVIDUAL APPROPRIATION

Contrary to the methodological individualism which poses as a matter of principle that all social phenomena must be explained on the basis of purely individual skills and abilities, we propose an alternative approach where certain social abilities that can be

recognized in individuals are not the cause, but rather the consequence of interactions where an irreducibly collective component intervenes (De Jaegher et al., 2010). To do this, we have to show how these collective components can emerge, and how they can play a role in the activity of individuals. The experiments we have presented here attempt to fill this requirement, since they make it possible to precisely define:

- (a) the initial individual abilities, and quite explicitly those that were initially absent;
- (b) the emergent phenomena resulting from the collective dynamics; and
- (c) the appropriation by individuals of the collective phenomena which are constituted in this way.

### Recognition of the other: the first experiment

- (a) In the first experiment, what the participants possess from the start are their perceptual abilities – in particular, the capacity to localize a shape in the one-dimensional space of exploration. However, by construction, the participants have no indication concerning the shape or the movement which might be associated to the other (the shape and movements of the partner and the mobile lure are exactly similar). Moreover, the body-object of a participant (that which can be perceived by the partner) is not perceivable by the participant himself.
- (b) The meeting of the efforts of each partner to constitute objects in his space of perception produces an attractor for the perceptual activities. This attractor does not correspond for either partner to a deterministic sensorimotor law. Indeed, in the minimalist conditions that we have given ourselves, if a subject does discover a stable sensorimotor law, for example the regular and symmetrical oscillation around a point of sensory stimulation, that will constitute for that subject the perception of an immobile object in the one-dimensional space of action. In the same way, an asymmetrical oscillation around a point of stimulation that is continually shifted will constitute the perception of an object in uniform movement. However, given the minimalism of a single receptor field, if the object moves faster than the participant can move to explore it, its spatial constitution becomes impossible. One of the points of interest of our experimental situation resides here: if the other is, like me, engaged in perceptual activity, the movements of his body-object, like those of the perceptual field which is attached to it, are necessarily too fast for me to be able to determine them spatially. Here it is thus impossible, by construction, to recognize in advance a determinate behavior, and then, by perceptual or cognitive inference, to attribute an intentionality to it (Premack, 1990; Csibra et al., 2003). On the contrary, it is this very impossibility to precisely determine the sensory feedbacks by their actions, which seems to be picked up by the participants as the clue leading them to indicate the presence of the other: in spite of the indeterminacy of the sensorimotor contingencies, the participants can relate their actions to sensory stimuli which are persistently present while remaining unpredictable. Indeed, if the participants respond more often to the presence of the body-object of their partner than to that of the mobile lure, it is because the perceptual

activities attract each other – just as in the visual domain, looks can attract each other.

We may note that the impossibility for the participants to perceive the image that they present to their partner is actually a necessary condition for the appearance of the dynamics of interaction of perceptual crossing. If the image that I present to the other subject was an image that I could perceive myself, an object for my perceptual activity, the dynamics of a perceptual crossing would become impossible since this image would no longer be linked to my perceptual activity.

- (c) However, as we have seen, the participants remain incapable of specifically identifying the presence of the other in any particular stimulation. This individual failure shows that the perceptual crossing does not proceed from a specific recognition of the other. The dynamics of the interactions escapes each of the individual partners. This will change in the second experiment.

### **Recognition of the other: the second experiment**

- (a) In the second experiment, the only difference with respect to the first one is that the participants possess from the start the additional capacity to distinguish the three different types of object (different characteristic sounds).
- (b) The emergent dynamics is the same; but
- (c) This time the emergent dynamics can be appropriated by associating the indeterminacy of the position of an object with one or other of two distinct sounds. The different intrinsic properties of the objects can be associated with properties characteristic of the dynamics of the interaction. An individual learning of the association between a given sound and a behavior of perceptual crossing becomes possible.

This opens up a path for explaining, by means of the functional meaning of the interactions, the formation of internal brain structures which may participate in the recognition of clues associated with this situation. The collective dynamics of the perceptual crossing situation brings about a situation of sensorimotor interactions that are sufficiently stable to serve as the basis for associative learning, i.e., the structuring of a neuronal system which associates concomitant multimodal input sensations, whether their origin be exteroceptive, proprioceptive, or resulting from previous actions.

If we apply this explanatory scheme to the development of the new-born infant, we may suppose that the dynamics of perceptual crossing with the caregiver is associated with the visual perception of the intrinsic properties of their face (Lavelli and Fogel, 2005; Itier and Batty, 2009). In a more general way, in the animal kingdom, the perceptual crossings that an organism exchanges with other organisms (according to the species, these organisms will more or less reliably belong to the same species) will make it possible to set up an association between this situation and characteristic which discriminate fellow creatures. We may note that, in its generality, this explanatory scheme does not decide in favor of either a hereditarian or environmentalist conception of human social cognition. It does however militate strongly in favor of an interactionist approach, and thus a fully social approach to social cognition. The important point is that the dynamics of inter-individual interaction constitutes a situation which associate

on one hand a perceptual crossing, and on the other hand a particular perceptual content. The association between this social signifying dynamics and perceptual contents could equally well be the result of individual associative learning, or of the selection of hereditary characters which accomplish this association. The logical point which is crucial here is that the individual neuronal structures which participate in the association can be the *result* and not the primary cause of this dynamics of interaction. If, on the contrary, the inter-individual interactions had to be the effect of prior internal structures – if it were necessary to already have the means of recognizing partners before engaging in an interaction with them – then the process of learning, or the evolutionary scenario, which account for the appearance of these structures would be almost impossible to imagine, because it would be necessary to associate radically heterogeneous elements (things perceived, actions performed) without any prior concrete association.

### **Imitation: the third experiment**

- (a) In our third experiment, the prior capacities that the participants bring to the situation are again those of being able to engage in a dynamics of interaction. By construction, the participants do not have any possible perceptual access, either exteroceptive or proprioceptive, of their own body-object that they present to the other participant. To the extent that the actions of clicking produce only a displacement of this body-object relatively to their receptor field, they cannot acquire any perceptual meaning for an isolated subject. The only access that a subject can have to the meaning of these actions passes by the indirect route of meeting with another entity which is sensitive to variations in this body-object, i.e., passes by interaction with another subject.
- (b) As before, the absence (see “b” in Section “Recognition of the Other: The First Experiment”) of any access of the participants to their own body-object explains the instantiation of a dynamics of perceptual crossing. It is this perceptual crossing which, as a collective dynamics, is sensitive to the relative positions of the body-objects.
- (c) It is the reappropriation by the individuals of the drift or the stability of their perceptual crossing which serves as a reference, and makes it possible for each participant to discover the meaning of their clicks.

One might say that the perceptual crossing functions like a sign which allows the subjects to know if they are in agreement. However, this sign is not arbitrary (contrary to a linguistic signifier which could be linked to any signified content whatsoever). What is signified here by the agreement of a well-balanced perceptual crossing, are the very conditions for the realization of the perceptual crossing in question.

### **GENERAL CONCLUSIONS**

There is a long distance between the deliberately minimalist situations that we have just explored, and natural situations. However, our aim here is to give an existence proof for a certain sort of explanatory scheme: recognition of the other as subject, or a form of imitation, can be genuinely explained in the framework of an interactionist approach, i.e., without appealing to any prior

knowledge (be it innate or acquired) which correlates what is perceived and what is done. To the extent that we have succeeded, this explanatory scheme should make it possible to actually *account* for the formation of neuronal structures, such as the famous “mirror neuron” system, which are activated both when an action is performed, and when it is perceived in another subject.

The “Associative Sequence Learning” model (ASL) proposes to account for the formation of these structures by classical sensorimotor learning based on the association between the observation and the execution of the same action (Heyes, 2001). Setting up a correspondence between internal representations of the actions performed, and the visual perception of these actions, does not seem to present any particular difficulty when the subject can see their own actions at the same time as those of the other subject. However, in the case of opaque actions (the subject does not see what it is that she is doing), it becomes necessary to imagine an association between the actions that the subject produces, and the sensory return corresponding to what she sees on the face of the model. The problem is that there is then no certainty that the action seen is similar to the action being performed. For that, there has to be a social synchronization, as when the caregiver plays the role of a “model” who actually imitates the expression of the infant (unless one uses an artificial mirror). But even then, two problems remain: (1) how does the infant recognize that she is engaged in a session of “imitation”? and (2) how can she select the relevant visual variables on the face that she is perceiving?

The rather particular situation of imitation that we have presented above proposes a different explanatory scheme which could help to provide some answers to these problematic questions. A certain sort of very simple, basic “imitation” could result directly from the dynamics of the interaction, independently of any deliberate internal matching between the actions of producing facial expressions and the perception of these expressions on the partner. The perception of an “agreement” precedes the knowledge of what the agreement is about. In this perspective, games of proto-conversation do not mean that the infant knows that (s)he is imitating (that her facial expressions are more or less correct reproductions of those of the adult), but only that the infant has the capacity to recognize the existence of an agreement in the interaction (Reddy, 2003; Trevarthen and Reddy, 2007). From this point on, if the infant perceives the expression presented by the caregiver at the same time that he recognizes this agreement, a learning process becomes possible. As an attractor, the

perceptual crossing creates conditions that are stable enough for there to arise an association between the actions performed and the concomitant sensory returns. The existence of structures such as “mirror neurons” could be explained by such an association between different synchronized fluxes of multimodal and proprioceptive sensory inputs, sensory data which comes both from the behavior of the other subject and from the subject’s own actions. We must insist on the fact that it is a question here of an association between the face of the partner and the dynamics of interaction which is socially meaningful (the perceptual crossing). The classical logic of “imitation” is inverted. Here, it is the *de facto* “imitation” resulting from the collective dynamics, which then provides the means for linking the perceived image to proprioceptive sensations. It is only later that the child will discover that what he is doing is in fact an imitation. On the basis of an agreement in the perceptual crossing, the subjects may presume that their own facial expression, that they cannot see, actually resembles that of their partner that they do see.

A major interest of the explanatory reversal that we propose here, is to make it possible to engage a dialog between scientific research and phenomenological descriptions (Varela et al., 1999; Gallagher, 2001; Thompson, 2007). For example, the phenomenological description of the encounter with the Other as a radical otherness which refuses any definitive determination (Levinas, 1979) or that of an intersubjective world in which emotions are shared (Merleau-Ponty, 1996), find corresponding elements in interaction dynamics that can be objectively observed, and that can be associated with bodily and neuronal structures. In this way, we hope that a scientific study on social cognition can be coherent with a description of the lived experience of human activity in a society and a culture where it is meaningful.

## ACKNOWLEDGMENTS

This research was supported by a grant from the Regional Council of Picardie. The authors thank all the members of the group CRED (Cognitive Research and Enactive Design) of the research unit COSTECH (Knowledge, Organization and Technical Systems) for their collaboration in this work, and the students Guillaume Doisy and Siffrénie De Bellabre for their help in the organization of experiments. We would also like to thank Tom Froese for his interesting comments, and subsequent discussion, which has helped us to substantially improve the text.

## REFERENCES

- Argyle, M., and Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry* 289–304.
- Auvray, M., Lenay, C., and Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New Ideas Psychol.* 27, 32–47.
- Auvray, M., and Myin, E. (2009). Perception with compensatory devices: from sensory substitution to sensorimotor extension. *Cogn. Sci.* 33, 1036–1058.
- Bach-y-Rita, P. (1990). Brain plasticity as a basis for recovery of function in humans. *Neuropsychologia* 28, 547–554.
- Bach-y-Rita, P. (2004). Tactile sensory substitution studies. *Ann. N. Y. Acad. Sci.* 1013, 83–91.
- Bach-y-Rita, P., and Kercel, S. W. (2003). Sensory substitution and the human-machine interface. *Trends Cogn. Sci. (Regul. Ed.)* 7, 541–546.
- Baron-Cohen, S., and Cross, P. (1992). Reading the eyes: evidence for the role of perception in the development of a theory of mind. *Mind Lang.* 7, 172–186.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The “Reading the Mind in the Eyes” test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251.
- Berger, P. L., and Luckmann, T. (1966). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor Books.
- Blakemore, S. J., and Decety, J. (2001). From the perception of action to the understanding of intention. *Nat. Rev. Neurosci.* 2, 561–567.
- Brass, M., and Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends Cogn. Sci.* 9, 489–495.
- Brooks, R. A. (1991). Intelligence without representation. *Artif. Intell.* 47, 139–159.

- Butterworth, G. (1999). "Neonatal imitation: existence, mechanisms and motives," in *Imitation in Infancy. Cambridge Studies in Cognitive Perceptual Development*, eds D. J. Nadel and G. Butterworth (New York: Cambridge University Press), 63–88.
- Butterworth, G., and Jarrett, N. (1991). What minds have in common is space: spatial mechanisms serving joint visual attention in infancy. *Br. J. Dev. Psychol.* 9, 55–72.
- Catmur, C., Walsh, V., and Heyes, C. (2007). Sensorimotor learning configures the human mirror system. *Curr. Biol.* 17, 1527–1531.
- Catmur, C., Walsh, V., and Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 2369.
- Clark, A., and Chalmers, D. (1998). The extended mind. *Analysis* 58, 7–19.
- Coleman, P., and Watson, A. (2000). Infant attachment as a dynamic system. *Hum. Dev.* 43, 295–313.
- Collins, C. C., and Bach-y-Rita, P. (1973). Transmission of pictorial information through the skin. *Adv. Biol. Med. Phys.* 14, 285.
- Cook, R., Press, C., Dickinson, A., and Heyes, C. (2010). Acquisition of automatic imitation is sensitive to sensorimotor contingency. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 840.
- Csibra, G., Biro, S., Koós, O., and Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cogn. Sci.* 27, 111–133.
- De Jaegher, H., and Di Paolo, E. (2007). Participatory sense-making. *Phenomenol. Cogn. Sci.* 6, 485–507.
- De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends Cogn. Sci.*, 14, 441–447.
- De Volder, A. G., Catalan-Ahumada, M., Robert, A., Bol, A., Labar, D., Coppens, A., Michel, C., and Veraart, C. (1999). Changes in occipital cortex activity in early blind humans using a sensory substitution device. *Brain Res.* 826, 128–134.
- Dennett, D. C. (1971). Intentional systems. *J. Philos.* 68, 87–106.
- Deschamps, L., Le Bihan, G., Lenay, C., Rovira, K., Stewart, J., and Aubert, D. (2012). "Interpersonal recognition through mediated tactile interaction," *Proceedings of IEEE Haptics Symposium 2012*, March 4–7, Vancouver, 239–245.
- Di Paolo, E. A., and Iizuka, H. (2008). How (not) to model autonomous behaviour. *Biosystems* 91, 409–423.
- Di Paolo, E. A., Rohde, M., and Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas Psychol.* 26, 278–294.
- Dumas, G. (2011). Towards a two-body neuroscience. *Commun. Integr. Biol.* 4, 349.
- Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proc. Natl. Acad. Sci. U.S.A* 99, 9602.
- Ferrari, P. F., Rozzi, S., and Fogassi, L. (2005). Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *J. Cogn. Neurosci.* 17, 212–226.
- Fogel, A. (2006). Dynamic systems research on interindividual communication: the transformation of meaning-making. *J. Dev. Process.* 1, 7–30.
- Froese, T., and Di Paolo, E. (2008). "Stability of coordination requires mutuality of interaction in a model of embodied agents," in *From Animals to Animals 10, The Tenth International Conference on the Simulation of Adaptive Behavior*, Osaka, 52–61.
- Froese, T., and Di Paolo, E. (2011a). The enactive approach: theoretical sketches from cell to society. *Pragmat. Cogn.* 19, 1–36.
- Froese, T., and Di Paolo, E. (2011b). Toward minimally social behavior: social psychology meets evolutionary robotics. *Lecture notes in computer science: Advances in Artificial Life. Darwin Meets von Neumann*, 5777, 426–433.
- Froese, T., and Di Paolo, E. A. (2010). Modelling social interaction as perceptual crossing: an investigation into the dynamics of the interaction process. *Conn. Sci.* 22, 43–68.
- Galantucci, B., and Sebanz, N. (2009). Joint action: current perspectives. *Top. Cogn. Sci.* 1, 255–259.
- Gallagher, S. (2001). The practice of mind. Theory, simulation or primary interaction? *J. Conscious. Stud.* 8, 5, 83–108.
- Gallese, V. (2007). Before and below 'theory of mind': embodied simulation and the neural correlates of social cognition. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 659–669.
- Gallese, V., Keysers, C., and Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends Cogn. Sci. (Regul. Ed.)* 8, 396–403.
- Gapenne, O., Rovira, K., Lenay, C., Stewart, J., and Auvray, M. (2005). "Is form perception necessary tied to specific sensory feedback," in *13th International Conference on Perception and Action (ICPA)*, Monterey, CA, 16.
- Gergely, G., Nádasdy, Z., Csibra, G., and Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193.
- Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum.
- Gillespie, A. (2005). GH Mead: theorist of the social act. *J. Theory Soc. Behav.* 35, 19–39.
- Goffman, E., and Best, J. (2005). *Interaction Ritual: Essays in Face-to-Face Behavior*. Chicago: Aldine Transaction.
- Guarniero, G. (1974). Experience of tactile vision. *Perception* 3, 101–104.
- Heider, F. (1982). *The Psychology of Interpersonal Relations*. Hillsdale, NJ: Lawrence Erlbaum.
- Heyes, C. (2001). Causes and consequences of imitation. *Trends Cogn. Sci. (Regul. Ed.)* 5, 253–261.
- Hutchins, E. (1995). *Cognition in the Wild*, Vol. 19. Cambridge, MA: MIT Press.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* 286, 2526.
- Itier, R. J., and Batty, M. (2009). Neural bases of eye and gaze processing: the core of social cognition. *Neurosci. Biobehav. Rev.* 33, 843–863.
- Knoblich, G., Butterfill, S., and Sebanz, N. (2011). Psychological research on joint action: theory and data. *Psychol. Learn. Motiv.* 54, 59–101.
- Lavelli, M., and Fogel, A. (2005). Developmental changes in the relationship between the infant's attention and emotion during early face-to-face communication: the 2-month transition. *Dev. Psychol.* 41, 265.
- Lenay, C., Auvray, M., Sebbah, F. D., and Stewart, J. (2006). "Perception of an intentional subject: an enactive approach," in *Third International Conference on Enactive Interaction*, Montpellier, 37–38.
- Lenay, C., Canu, S., and Villon, P. (1997). "Technology and perception: the contribution of sensory substitution systems," in *International Conference on Cognitive Technology* (Los Alamitos, CA: IEEE Computer Society), 44.
- Lenay, C., Gapenne, O., Hannequin, S., Marque, C., and Genouelle, C. (2003). "Sensory substitution: limits and perspectives," in *Touching for Knowing, Cognitive Psychology of Haptic Manual Perception*, eds D. Y. Hatwell, A. Streri, and E. Gentaz (Amsterdam: John Benjamins Publishing Company), 275–292.
- Lenay, C., and Steiner, P. (2010). Beyond the internalism/externalism debate: the constitution of the space of perception. *Conscious. Cogn.* 19, 938–952.
- Lenay, C., Stewart, J., Rohde, M., and Ali Ammar, A. (2011). "You never fail to surprise me": the hallmark of the other. Experimental study and simulations of perceptual crossing. *Interact. Stud.* 3, 373–396.
- Leslie, A. M. (1987). Pretense and representation: the origins of "theory of mind." *Psychol. Rev.* 94, 412.
- Levinas, E. (1979). *Totality and Infinity: An Essay on Exteriority*, Vol. 1. Berlin: Springer.
- Maillet, B., Lenay, C., and Guenand, A. (2008). "Designing for interpersonal tactile interaction over distance," in *User Experience of ICTs – Proceedings of the 21st International Symposium on Human Factors in Telecommunications*, March 17–20, Kuala Lumpur, 391–398.
- Marsh, K. L., Richardson, M. J., and Schmidt, R. (2009). Social connection through joint action and interpersonal coordination. *Top. Cogn. Sci.* 1, 320–339.
- McGann, M., and De Jaegher, H. (2009). Self-other contingencies: enacting social perception. *Phenomenol. Cogn. Sci.* 8, 417–437.
- Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Dev. Psychol.* 31, 838.
- Meltzoff, A. N., and Moore, M. K. (1997). Explaining facial imitation: a theoretical model. *Early Dev. Parent.* 6, 179–192.
- Meltzoff, A. N., and Moore, M. K. (1999). "Persons and representation: why infant imitation is important for theories of human development," in *Imitation in Infancy. Cambridge Studies in Cognitive Perceptual Development*, eds D. J. Nadel and G. Butterworth (New York: Cambridge University Press), 9–35.
- Merleau-Ponty, M. (1996). *Phenomenology of Perception*. New Delhi: Motilal Banarsidass Publisher.
- Michael, J., and Overgaard, S. (2012). Interaction and social cognition: a comment on Auvray et al.'s perceptual crossing paradigm. *New Ideas Psychol.* 30, 296–299.
- Nagy, E. (2006). From imitation to conversation: the first dialogues with human neonates. *Infant Child Dev.* 15, 223–232.

- Newcomb, T. M., Converse, P. E., and Turner, R. H. (1966). *Social Psychology: The Study of Human Interaction*. London: Routledge & Kegan Paul.
- O'Regan, J. K., and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–972.
- Petras, J. W. (1968). John Dewey and the rise of interactionism in American social theory. *J. Hist. Behav. Sci.* 4, 18–27.
- Povinelli, D. J., Reaux, J. E., Theall, L. A., and Giambone, S. (2000). *Folk Physics for Apes: The Chimpanzee's Theory of How the World Works*. New York: Oxford University Press.
- Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition* 36, 1–16.
- Ray, E., and Heyes, C. (2011). Imitation in infancy: the wealth of the stimulus. *Dev. Sci.* 14, 92–105.
- Reddy, V. (2003). On being the object of attention: implications for self-other consciousness. *Trends Cogn. Sci. (Regul. Ed.)* 7, 397–402.
- Reddy, V. (2008). *How Infants Know Minds*. Cambridge, MA: Harvard University Press.
- Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.* 21, 188–194.
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192.
- Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (2002). "From mirror neurons to imitation: facts and speculations," in *The Imitative Mind: Development, Evolution, and Brain Bases*, Vol. 6, eds A. N. Meltzoff and W. Prinz (Cambridge: Cambridge University Press), 247–265.
- Rohde, M., and Di Paolo, E. (2008). "Embodiment and perceptual crossing in 2D: a comparative evolutionary robotics study," in *From Animals to Animals 10, The Tenth International Conference on the Simulation of Adaptive Behavior*, eds M. Asada, J. Hallam, J.-A. Meyer, and J. Tani (Berlin: Springer), 83–92.
- Schiff, W., and Foulke, E. (1982). *Tactual Perception: A Sourcebook*. New York: Cambridge University Press.
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., and Vogele, K. (2006). Being with virtual others: neural correlates of social interaction. *Neuropsychologia* 44, 718–730.
- Schmidt, R., and Richardson, M. (2008). "Dynamics of interpersonal coordination," in *Coordination: Neural, Behavioral and Social Dynamics*, eds A. Fuchs and V. Jirsa (Heidelberg: Springer), 281–308.
- Schütz, A. (1970). *Alfred Schutz on Phenomenology and Social Relations*. Chicago: University of Chicago Press.
- Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends Cogn. Sci. (Regul. Ed.)* 10, 70–76.
- Sribunruangrit, N., Marque, C. K., Lenay, C., Hannon, S., Gapenne, O., and Vanhoutte, C. (2004). Speed-accuracy tradeoff during performance of a tracking task without visual feedback. *IEEE Trans. Neural Syst. Rehabil. Eng.* 12, 131–139.
- Steiner, P., and Stewart, J. (2009). From autonomy to heteronomy (and back): the enaction of social life. *Phenomenol. Cogn. Sci.* 8, 527–550.
- Stewart, J., and Gapenne, O. (2004). Reciprocal modelling of active perception of 2-D forms in a simple tactile-vision substitution system. *Minds Mach.* 14, 309–330.
- Thomas, R., Press, C., and Haggard, P. (2006). Shared representations in body perception. *Acta Psychol. (Amst)* 121, 317–330.
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Belknap Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H., and others. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–690.
- Tremoulet, P. D., and Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception* 29, 943–952.
- Trevarthen, C. (1993). "The self born in intersubjectivity: the psychology of an infant communicating," in *The Perceived Self: Ecological and Interpersonal Sources of Self-knowledge*, ed. U. Neisser (New York: Cambridge University Press), 121–173.
- Trevarthen, C., and Reddy, V. (2007). Consciousness in infants. *The Blackwell Companion to Consciousness*, eds M. Velmans and S. Schneider (Malden, MA: Blackwell Publishing), 37–57.
- van Dijk, J., Kerkhofs, R., van Rooij, I., and Haselager, P. (2008). Can there be such a thing as embodied embedded cognitive neuroscience? *Theory Psychol.* 18, 297–316.
- Varela, F. J. (1979). *Principles of Biological Autonomy*. New York: North Holland.
- Varela, F. J., Thompson, E., and Rosch, E. (1999). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.
- Visell, Y. (2009). Tactile sensory substitution: models for enaction in HCI. *Interact. Comput.* 21, 38–53.
- Ward, J., and Meijer, P. (2010). Visual experiences in the blind induced by an auditory sensory substitution device. *Conscious. Cogn.* 19, 492–500.
- Wilkerson, W. S. (1999). From bodily motions to bodily intentions: the perception of bodily activity. *Philos. Psychol.* 12, 61–77.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 December 2011; accepted: 05 April 2012; published online: 09 May 2012.

Citation: Lenay C and Stewart J (2012) Minimalist approach to perceptual interactions. *Front. Hum. Neurosci.* 6:98. doi: 10.3389/fnhum.2012.00098

Copyright © 2012 Lenay and Stewart. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.