



Commentary

Social understanding through direct perception? Yes, by interacting

Hanne De Jaegher

Department of General Psychiatry, University of Heidelberg, Germany
 Centre for Computational Neuroscience and Robotics, University of Sussex, UK

ARTICLE INFO

Article history:

Available online 16 December 2008

Keywords:

Social cognition
 Theory of mind
 Simulation theory
 Direct perception
 Interaction process
 Coordination
 Participatory sense-making
 Intersubjectivity
 Phenomenology
 Cognitive science

ABSTRACT

This paper comments on Gallagher's recently published direct perception proposal about social cognition [Gallagher, S. (2008a). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17(2), 535–543]. I show that direct perception is in danger of being appropriated by the very cognitivist accounts criticised by Gallagher (theory theory and simulation theory). Then I argue that the experiential directness of perception in social situations can be understood only in the context of the role of the *interaction process* in social cognition. I elaborate on the role of social interaction with a discussion of *participatory sense-making* to show that direct perception, rather than being a perception enriched by mainly individual capacities, can be best understood as an interactional phenomenon.

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Is there no problem of other minds? In a recent paper, Shaun Gallagher proposes that we understand each other partly through *direct perception*. Direct perception is the phenomenological notion of a rich, 'smart' perception. When we encounter an other we "have a direct perceptual grasp of the other person's intentions, feelings, etc." (Gallagher, 2008a, p. 535). Gallagher contrasts this notion with traditional, cognitivist approaches to social cognition such as theory theory (TT) and simulation theory (ST), which start from a 'dumb' perception that needs to be supplemented with interpretation sustained by inferential or simulation processes before an individual can figure out an other's emotions and intentions.

In this paper, I intend to show first that Gallagher's direct perception proposal may not be able to withstand a potential cognitivist hijack into an inferencing/simulating account in which social perception builds on fast and behind-the-scenes cognitive processing. Yes, we have the experience of a direct grasp of others' intentions and feelings, the cognitivist can say, and this is because of the very fast processing going on inside the head. Gallagher's claim that mirror neurons are not a mechanism of simulation but one of perception does not help because it does not dispel the idea that social cognition is something done in the individual head. In order to safeguard direct perception, I propose that the experiential directness of social perception can be understood only in the context of the role of *social interaction* in social cognition.

In contrast to the majority of current perspectives on social understanding, taking interaction seriously involves focusing on how participants engage with one another, not on how each figures out the other. Whereas traditional approaches centre on how individual cognisers interpret an other's behaviour, the proposal here is to focus on the process of interacting, which can be analysed as a dynamical system spanning the individuals involved. Such an analysis may involve investigating patterns of movement or of language use across individuals during an interaction – i.e. not patterns of movement in one person and in the other, but patterns characteristic of the system as a whole as it unfolds in time. Recent experimental and modeling

E-mail addresses: h.de.jaegher@gmail.com, h.de.jaegher@uni-heidelberg.de

work has shown how collective patterns of coordination account for the performance of sophisticated social tasks, such as distinguishing otherwise identical contingent stimuli from non-contingent ones. In such cases, it can be shown that individuals are not capable of performing such tasks by themselves, but can only do it through the interaction (Di Paolo, Rohde, & Iizuka, 2008; Auvray, Lenay, & Stewart, in press). Other examples of guiding questions for this research are: How do two people construct a sentence together in conversation (Goodwin, 1979)? What is the role of addressed and non-addressed recipients when telling a story in a group (Goodwin, 1984)? How does the meaning of a specific word evolve with its use in a certain conversation? Interactional movement can be studied in just such a way. For instance, how do people establish a rhythm, intensity and duration for making eye-contact that is comfortable for each participant in the course of an encounter, and what role do specific body movements and interpersonal rhythms play in this? What is the process by which we become locked in a certain interaction pattern despite the fact that each of us wants to stop it or change it (Granic, 2000)?

I have argued before that understanding the social interaction process should be the central task of any account of intersubjectivity (De Jaegher & Di Paolo, 2007). The present paper is dedicated to raising the concrete possibility that an interactive approach to social cognition can account for direct perception – one of many intersubjective phenomena – and how such an account avoids giving in to potential cognitivist temptations. The overall point of the criticism is (1) to extend and complement Gallagher's proposal and (2) to move the debate into a properly social domain that involves understanding participants in the act of interacting.

2. From dumb to smart perception?

According to Gallagher, direct perception is an aspect of social understanding that gives us an insight into what is going on with another person. In explaining what direct perception is, he seems to want to go two ways. In some places, he approaches direct perception in a Gibsonian manner:

“Practically speaking, direct perception, etc. delivers what I need to interact with others most of the time. In the broad range of normal circumstances there is already so much available in the person's movements, gestures, facial expressions, and so on, as well as in the pragmatic or social context, that I can grasp everything I need for understanding in what is perceptually available” (Gallagher, 2008a, p. 540).

Here, all the information we need to understand an other is out there, in her comportment. Gallagher proposes that rather than being opaque, for purposes of ongoing interaction we are transparent to each other most of the time. We see others' emotions and intentions – mental states are not hidden, they are in plain view for all who want to see. On this take, there is little explaining to do with regard to how direct perception works – we just do it. But this is not satisfactory, and Gallagher also provides an account of how it works.

He contrasts direct perception with the kind of perception ostensibly at play in the traditional theory theory (TT) and simulation theory (ST) accounts. This is done by conceiving of perception as ranging from dumb to clever. The kind of perception at play in TT and ST, according to Gallagher, is dumb; it – crucially – needs to be supplemented by cognitive efforts such as mindreading or simulation in order for there to be full interpersonal understanding. Even though these approaches are generally not clear about what notion of perception they use in the first place, it is evident that if it wasn't supplemented by whatever “cognitive machinery” advocated, our mere perception of an other would not make us any the wiser about their intentions and feelings (Gallagher, 2008a, p. 536). Perception alone is supposedly not enough. We must infer and/or simulate in order to understand each other. This reveals the problem of other minds that these approaches start from: we are opaque to each other, and so we need to figure each other out – literally, by exercising a cognitive faculty.

This is what Gallagher reacts to, based on the phenomenology of everyday social understanding. He wants to avoid postulating high-level cognitive mechanisms at the core of our intersubjective skills. He offers a kind of perception, *direct perception*, which is “smart”. Unlike the perception of ST and TT, direct perception does not need cognitive supplementation. Gallagher illustrates this with the example of seeing his car. He does not first see a red blob and then discern that this is his car – he immediately sees his car. This kind of perception already ‘understands’ the stimuli before it.

This is an important point to make, but when Gallagher supports it here with arguments inspired by cognitive science, some of the basics of the point start to become obscured. Gallagher's overall project of bringing phenomenology and cognitive science together and letting the fields speak to each other – so evident in many of his works (for a broad view, see e.g. Gallagher, 2005; and Gallagher & Zahavi, 2007) – is an admirable and worthwhile endeavour. But the way in which he builds his argument here makes us ask the question *How does perception get so clever?* and this leads to trouble.

In answering this question, Gallagher lists a series of ‘enrichers’ of perception, and it is here that the argumentation starts to move away from the richness of the phenomenon of intersubjectivity. One source of perception's intelligence, says Gallagher, is experience, which “can tune our sensory-motor neuronal systems” (Gallagher, 2008a, p. 538). Not all perception is enriched by experience however, for even creatures who do not have (much) experience – infants for example – show skills that reveal a smart perception according to Gallagher. Infants show a preference for certain stimuli such as the human voice or their caregiver's mother tongue, they can imitate gestures very soon after birth, and they have a different kind of interest in biological movement than they have in non-animate movement. These findings lead Gallagher to propose an inborn drive to communicate (Gallagher, 2005), in accordance with for instance Trevarthen's ideas (see e.g. Trevarthen & Aitken, 2001). Later in development, perception is informed by the language and concepts we learn. Another enricher of perception is context: according to Gallagher the context of what we perceive informs our understanding of others. Perception is also clever because of the work done by the brain: “[t]he visual cortex is processing information about shape and color” and

“more processing in the inferior temporal cortex, and ‘top-down’ processes may focus and integrate the sensory information” (Gallagher, 2008a, p. 537).

Each of these suggestions about how perception is informed or enriched needs to be expanded to see how they can work, and they can each be critically discussed in their own right. For instance with regard to infant imitation it is not quite clear how infants achieve it. Is it really a matter of internally mapping visual and proprioceptive elements onto each other, or could another, perhaps more parsimonious explanation be given? Is declaring the innateness of a certain skill or trait not often a way of saying ‘we do not really know where it comes from’? We could also ask how context itself is perceived, and so on.

However, a more pertinent question is this: to what extent does this account present us with something new and to what extent do some of the underlying assumptions of cognitivism remain part of this proposal? More precisely, does Gallagher move away from the individualist and modular view of cognition that characterises cognitivism? Conceiving of direct perception as clever or informed inevitably raises the issue of how perception gets like this. Gallagher’s listing of several ‘enrichers’ does not avoid that which he holds against the notion of perception employed in TT and ST: it being a perception+, perception+cognitive efforts – except that in his proposal what comes after the + does not happen *alongside* or *after* the perceiving, it happens *within it*. Even if what is added into perception is not mindreading or simulation and it takes place somehow within it rather than alongside or at the tail end, what Gallagher does here is to bring to the table more and more ingredients for enriching perception, as if it were a sauce in need of flavouring. The ingredients may be sourced from better pastures (more local and organic perhaps), there may be more of them, but the recipe has not changed much. This does not mean that the recipe is one to stick with. In whatever way perception is enriched, whether by inference, simulation, enactive exploration, context, or innate drives, it remains an ‘enriched’ perception. The notion of direct perception is therefore not radically different from that of supplemented dumb perception. The fact that it makes us ask how perception gets *informed* points to an underlying assumption: that this perception was also dumb to start with. Or to put it differently, that perception and cognition are separate things that need glueing together. Compare it to the very critical evaluation of the notion of ‘embodiment’ that Sheets-Johnstone provides: first we carve nature up at artificial joints – we split mind and body apart – and then we need to fasten the two together again, a task for which the notion of embodiment is, according to Sheets-Johnstone’s assessment, used as a kind of glue (1999a, 1999b). But glueing the two back together does not bring back the original “integrity and nature of the whole” (Sheets-Johnstone, *in press*) – on the contrary, the crack remains visible. In a similar way, by relying on arguments in cognitive science here (even if in other places he does this masterly, see for instance several of the discussions in his book, 2005), Gallagher risks losing sight of precisely that which is at the origin of his proposal for direct perception: perception is an activity imbued with meaning from the start. In his argumentation, though not in his starting point, Gallagher accepts the split between perception and understanding that characterises cognitivist approaches. His choice of argumentative tools here therefore does not move the debate away from the idea that perception is something that needs to be filled in, added to or complemented in some way in order to allow interpersonal understanding, it merely relocates the sources of enrichment of perception.

Gallagher is arguing against TT and ST and this is needed indeed. But by staying on their turf, he risks facing problems very similar to the ones they face, as I have shown here. Gallagher’s commitment certainly lies with moving on the debate, but the very basis on which this is done needs to be re-thought. The criticisms of TT and ST that he and others formulate point to anomalies that cannot be solved within the existing paradigm. Instead, they ask for the development of a new paradigm. This is also evident in the other, bigger problem: in traditional approaches to social cognition, and in Gallagher’s too, the interaction process is all but neglected.

3. Where is the interaction?

The discussion of where direct perception gets its cleverness from has almost made us forget that the main question we are dealing with here is one of *intersubjectivity*, of *interpersonal* understanding – it is telling in this regard that most of the discussion of the ‘enrichers’ of perception could have been applied to Gallagher’s own example of perceiving his car. Direct social perception is not the kind of process in which some form of *social skilfulness* already inheres and does not need to be added. Or is it? In his article, Gallagher points to the importance of the interaction process: “what we call social cognition is often nothing more than that social interaction” (p. 540). This is a clear claim to the effect that social interaction often is social understanding. Taken literally it could be the aim of an interesting research project: work out *how* social interaction is social understanding. But such elaboration is a much neglected task. An in-depth investigation of interacting and its role in interpersonal understanding is mostly lacking in social cognition research. Some researchers, including Gallagher, hold the role of the interaction in high regard and emphasise its importance (Gallagher, 2004; Gallagher & Hutto, 2008; Ratcliffe, 2007). Although this is very valuable, behind acknowledging the importance of something sometimes lurks the danger of taking that something for granted. This is exactly what happens with regard to the interaction process. Nobody denies its importance and nobody explains its role. Gallagher himself has proposed an alternative to TT and ST in the form of Interaction Theory (Gallagher, 2004, 2008b; Gallagher & Hutto, 2008) and this forms an excellent starting point for further investigation. What needs to be done now is to go beyond stating *that* the interaction is important and to start treating this as a real shift in the focus of social cognition research and to work out its implications.

We can understand social perception as the rich process Gallagher rightly says it is – if we confront the more general question of what processes ground meaning and get a grip on the role of interacting in social cognition.

4. Sense-making and participatory sense-making

Di Paolo and I have recently sketched an approach to the role of the interaction process in social cognition (De Jaegher & Di Paolo, 2007, 2008). Here, rather than reintroduce the proposal, I will highlight some of our main points, and then elaborate on some issues that will bring the possibility of a direct perception in the social realm into sharper focus. In Gallagher's work there seems to be an assumption that "direct perception... delivers what [we] need to interact with others most of the time" (Gallagher, 2008a, p. 540). What I argue is exactly the opposite: we may experience an other's feelings and intentions directly, but direct perception builds on something, namely on skilful interaction with others. In other words, social interaction is not derivative, but constitutive of the process of social understanding and also of direct social perception. Therefore, working out a detailed account of social interaction's role in interpersonal understanding is the central element of the story of social cognition. It will allow the issue to move away from the terms of the debate set by TT and ST and followed by direct perception (by the way, none of these approaches have defined what they mean by 'social')¹ and towards a story that explicitly connects meaning and social interaction.

Focusing on the interaction is a radical step. It grounds the study of intersubjectivity and social cognition on two new foundations. First, rather than starting from the idea that social cognition happens within an individual – whether brain or whole body – it makes the investigation of the role of the interaction process in social cognition imperative and possible. Second, it permits us to think of how intentions emerge and are transformed in and through social interaction. The radical step lies in moving away from seeing social phenomena as events external to the perceiver that must be appropriately interpreted, and instead seeing interacting as a process in which an interactor is immersed and as a process that in itself may play a cognitive role.²

This is not to say that individuals are subsumed passively within the interaction process. On the contrary, by thematising the interplay between individual autonomy and the autonomy of the interaction process, the subjectivity of the individuals is recognised in a way that is not present in traditional approaches. It is also not to say that so-called 'high-level cognitive processes', such as wondering about why someone did something or thinking about how best to approach asking someone a favour, do not play a role in social cognition. It is to say that both cognition and social cognition need to be re-thought, and that their basis lies in the process of interacting. There are good developmental reasons to suspect that such higher level cognitive processes are strongly affected by the history of interactions (see for instance Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001, and the commentaries by P. Rochat and D. N. Stern in the same volume; Hobson, 2002; Trevarthen, 2005), a suggestion that Gallagher also moves towards in his (2001) paper.

To begin with the question where meaning comes from, we characterise cognition in general as *sense-making*. This means that "[e]xchanges with the world are inherently significant for the cogniser and this is the definitional property of a cognitive system: the creation and appreciation of meaning or *sense-making* in short" (Di Paolo, Rohde, & De Jaegher, in press). Understanding cognition as sense-making means to ground meaning in aliveness. Different thinkers and schools of thought have expressed this in different ways. Enactivists ground cognition or sense-making in the self-producing, self-organising and self-distinguishing organisation of living beings (Di Paolo, 2005; Di Paolo, in press; Maturana & Varela, 1980; Thompson, 2007; Varela, 1991; Varela, Thompson, & Rosch, 1991; Weber & Varela, 2002). A cogniser is a '*centre of activity*' in the world and this entails that she has a perspective, based in her structure, history and needs. Similarly, Sheets-Johnstone (1999b) characterises sense-making as rooted in animateness and self-movement.³

¹ For a definition of 'the social', see De Jaegher and Di Paolo (2007).

² Perhaps what lies behind the prevalence of the individualist position is the idea that such a skill is simpler than fully engaged interaction – both to explain and to achieve. Like several similar issues in embodied cognition research, the case is exactly the opposite: the apparently more complex situation points much more directly to the nature of the problem and to its solution.

³ The notion of 'sense-making' is very different from the everyday expression 'making sense of something'. It does not carry the latter's connotation of very high level cognitive activity involving reasoning and inference. We use 'sense-making' similarly to the way in which Sheets-Johnstone (1999b), Thompson (2007), and Weber and Varela (2002) use it: to refer to the generation and transformation of meaning in the most general sense, *not* a high-level cognitive process. Meaning, on this account, is grounded in the biology of living beings. Sense is made by living beings in interaction with their world. This world is not a naked environment – the physical processes that surround the organism – but a world given value and meaning by the norms imposed by its self-production and self-maintenance. At the most basic level, the processes involved in sustaining life determine which aspects of the world are relevant for the organism. A spider builds a web in order to catch flies to eat, while I go shopping on my bike and a bushman goes hunting with a spear. A cat lies in the sun to rest, while I, even though I'm also tired, still book a train ticket to go and see my family in three weeks before I go and read a bit to relax. Every organism moves around in its world, makes choices and takes certain paths on the basis of what it needs to survive. On top of this basic survival, we also maintain increasingly complicated kinds of identity that require increasingly less life-or-death determined choices and actions, but rather choices based on other aspects of who we are, such as a daughter, a Frenchman, a friend of such and such, a fireman, and so on. Sense-making, in contrast to information-processing-based cognition, is not representing the objective world internally. It is the self- and world-constituting that an organism is continually engaged in. Inherent in this is the organism's individual perspective (Di Paolo, 2005; Thompson, 2007; Weber & Varela, 2002). The distinction with the everyday expression 'to make sense of something' is indicated by the use of the hyphen: sense-making, making-sense.

Sense-making is the active engagement of a cogniser with her environment.⁴ This activeness of sense-making is elaborated wonderfully in Sheets-Johnstone's account of the primacy of movement. According to her, "Our bodies, *through movement*, [are] the very condition of our constituting the world" (Sheets-Johnstone, 1999b, p. 148–149, emphasis in original). If we were not able to move, we would only be able to perceive whatever happened to drift past us: "the world would reduce to random events which, in the absence of active exploration, could hardly give rise to the idea of full-fledged objects, let alone full-fledged subjects" (p. 148).⁵ In sum, "[o]nly through movement can and does...our knowledge of ourselves and the world...take root" (p. 228). Sense-makers enact or constitute their world in and through their movement.⁶ This is in contrast to significance or meaning happening exclusively in certain neural activity. Meaning or sense is made in what one does with one's world. By doing this rather than that I live rather than die. By taking this path rather than that, I will have access to food rather than suffer hunger. By choosing this course of action rather than that, I strengthen my ties with my family, rather than driving them away from me. I do these things out of a drive to survive, out of love for my family, and so on. These actions are based in my perspectives and are in view of my goals – both of which themselves can be subject to continuous change throughout my activities and engagements.

On such a view, meaning is in the engagements in which an organism builds its world. On such a view, to argue that understanding is an ingredient of perception is like carrying water to the sea. To take perception and understanding as part of one and the same process is like swimming.

Starting from this characterisation of the roots of meaning and cognition in aliveness,⁷ we can move into the domain of intersubjectivity. In order to do this, it is necessary first to make a brief stop at the phenomenon of coordination. Coordination is a ubiquitous and measurable reality of physical and biological systems. We characterise it as "the non-accidental correlation between the behaviours of two or more systems that are in sustained coupling or have been coupled in the past or have been coupled to another, common, system" (De Jaegher & Di Paolo, 2007, p. 490). An example of physical coordination is the synchronisation of the pendulums of grandfather clocks when they are placed in each other's vicinity. Here, coordination is achieved by mechanical means. Even in biological or cognitive systems the source of coordination is often mechanical, and it is therefore not necessary to postulate high-level cognitive mechanisms at its basis when it happens in a cognitive system.

In social science, analyses of interactions and conversations show that people interindividually coordinate their movements (Bavelas, Coates, & Johnson, 2002; Condon & Ogston, 1971; Gill, Kawamori, Katagiri, & Shimojima, 2000; Kendon, 1982; Warner, 1996). The coordination found in this domain is variable both in kind and in strength: among other things, people mirror each others' movements, anticipate them, temporally synchronise or desynchronise and so on. A specific feature of social coordination is that patterns of coordination can influence the continuing disposition of the individuals involved in the encounter to sustain, modify or end it (De Jaegher & Di Paolo, 2007). A good example of this is an interaction that carries on even though none of the participants wishes to continue it. Think of the familiar situation where you encounter someone coming from the opposite direction on a narrow footpath. In attempting to walk past each other, you both step towards the same side. This may happen a few times before you are finally able to bypass each other. Here, the coordination of movements (a temporally synchronised mirroring of sideways steps) ensures (for a brief while) that the interaction process is sustained despite the fact that you both want to stop interacting in this way. This example illustrates how the interpersonal coordination of movements endows the interaction process with a form of autonomy. We can even go so far as to say that the interaction process itself influences the interactors. This insight is one of the consequences of taking the interaction process seriously. What does it buy us in the area of social understanding?

If movements and sense-making are foundationally linked, as seen above, and people's movements can be interindividually coordinated to such an extent that patterns of coordination may take on a 'life of their own', we have a way of conceiving of a sense-making that is truly intersubjective. When movements fundamentally intertwined with sense-making can interindividually coordinate, this means that *sense-making activities can coordinate interindividually*. This opens up a range of socially cognitive phenomena. At the least, interpersonal movement coordination can influence individual sense-making. This happens for instance when someone points to something for you to see. Following the pointing gesture makes you attentive to an aspect of your environment that you were not previously aware of. This also goes for the more metaphorical meaning of 'pointing something out to someone'.

More intricate possibilities are also opened up: we can now think about how intentions can be transformed – even generated – in social interaction. Let us continue with the pointing example. If it is not immediately clear what your partner is trying to point out to you, you may refer back to his gaze and look where he is looking. This can prompt him to look in his intended direction with more emphasis. 'Pointing' here is then interactionally achieved. Once you know what he is pointing out to you, let us assume for the sake of the example that the thing he points out is ambiguous. You can now start to make sense of the so established object of attention together. If he was pointing to someone's gestures for instance, you may come

⁴ This does not mean that there is no receptivity in cognition. Receptivity is part of cognition, but it is not passive. It is also an active engagement, think of attention, focus and openness for instance.

⁵ In fact, even this presupposes that we have experience, which on Sheets-Johnstone's account – which includes an original developmental illumination about self-movement and the wholeness of the experience of infants – we also can only gain through movement.

⁶ To clarify, I do not mean that sense-making is mapped onto certain movements in a bodylanguage kind of way where a certain body movement means a certain thing (e.g. arms crossed across chest = disagreement), but rather that sense-making is in movement, happens through movement. This has nothing to do with mapping.

⁷ I can only provide a brief statement of these ideas here. My task in this paper is not to argue for them, but I refer the reader to the relevant literature mentioned in the text for their argumentation.

to a mutual understanding of them based on an intricate to-and-fro of glances, utterances, gestures and so on – eventually reaching an understanding that was not there for either of you before your interaction.

In other words, we are no longer tied to thinking of intentions as individually formed and then brought into a social situation to be interpreted, simulated, or directly perceived. Instead, we now have the opportunity to gain an insight into *intersubjectivity* in the most literal sense of the word. Experimental and modeling tools in terms of dynamical systems analysis, as I mentioned, are available to study such interactive processes (Di Paolo et al., 2008; Ramseyer & Tschacher, 2006; Auvray et al., in press).

Sense-making in the social realm is then best characterised as *participatory sense-making*. As we have seen, this is a cognising that ranges from individual thoughts and feelings being influenced by an other's or by interacting with him, all the way to the truly intersubjective, joint transformation and generation of meaning (De Jaegher & Di Paolo, 2007).

5. Transparent interaction

Di Paolo and I have put forward this account as an alternative departure point for the investigation and understanding of intersubjectivity. I now explain its relevance for Gallagher's direct perception account. We have seen that Gallagher's proposal runs into difficulties when it comes to explaining how direct perception works. It could easily be reinterpreted as a cognitivist, individualist conception of cognition.

I propose that perception in the social realm acquires the richness and feels as direct as it does because of the *mediated immediacy* of social understanding through the interaction process.⁸ We are so proficient at social interaction that the process gains a transparency that makes perception in the social realm feel direct. By 'transparent' I do not mean to indicate that the interaction is like a window through which I see an other – that would lend the whole too much of a snapshot air. On the contrary, the interaction process is transparent while being also a *process*. Transparency can perhaps be better illustrated by the way a blind person's cane can be transparent: this instrument is only 'transparent' as long as it is being used. It is not a means of perception and the blind person does not see anything 'through' it when he simply holds it in his hand. It is only when he uses it, actively and skilfully explores his surroundings with it, that the kerb or a crack in the pavement become 'visible' 'through' it. And even this does not do justice to the intricacies of the social interaction process of course, because as a social agent involved in a social interaction, I am at once prodder and prodded, and so is my interaction partner. Each of us is capable both of perceiving the other and of changing through the interaction process. This proficiency builds on years of extensive experience throughout development, starting at the earliest stage with infant-caregiver interactions, and continuing throughout life in our daily encounters with others.⁹

But – and this is a crucial point – *it isn't always like this*. An other's intentions are not always transparent to me. Sometimes it is unclear what is going on with someone. In Gallagher's words: "the other person may in some circumstances be a real puzzle" (Gallagher, 2008a, p. 540). It is in his discussion of this issue that the question whether his proposal remains observational or not can be answered: it does. According to Gallagher, we gain our understanding from directly perceiving the other. We have already wondered how that could work and found that the nature of his argumentation lends itself too easily to cognitivist answers and makes us lose sight of the fact that we are dealing with the conundrum of *interpersonal* understanding. With regard to misunderstanding an other, cognitivist accounts of social cognition as well as perceptual accounts that are in danger of being annexed by cognitivist ones treat breakdown of social understanding as due to individual faults (hampered inferencing, faulty simulating, lack of perceptual grasp or ways to reduce ambiguity). Ways out of such breakdowns are also restricted to individual efforts. Gallagher himself seems quickly seduced again by the old solutions when he says that in such cases he "would not hesitate to say that we might then turn to other means (narrative or even theory or simulation)" (Gallagher, 2008a, p. 540).

Even though Gallagher is referring to exceptional situations where his proposed mechanisms would break down, examples of such situations prompt us to wonder how this account deals with failures in everyday social understanding. Failures in understanding an other's behaviour are not exceptional. On the contrary, they form part and parcel of the ongoing process of social understanding. More even, misunderstandings are the pivots around which the really interesting stuff of social understanding revolves. In these instances where coordination is lost, we have the potential to gain a lot of understanding. More often than not I feel myself disagreeing or questioning when my interaction partner has already reacted accordingly before he or I even formulate that disagreement or questioning for ourselves. Loss and recovery of coordination are mediated through the interaction itself – a coordination which, as we have seen above, is not a constant attunement, but rather a variable, sometimes even discontinuous process. It is when the flow of movement across me and my interaction partner grinds to a halt that my or our sense-making activity can and often does change direction. Where the ongoing flow of interacting and coordinating breaks, opportunities for redirecting sense-making open up. On such occasions, one of the things we can do is to attempt to repair our interaction, for instance through a re-attunement of movements and or utterances. But we also

⁸ The term 'mediated immediacy' comes from Plessner, H. (1981). *Die Stufen des Organischen und der Mensch. Gesammelte Schriften IV*. Frankfurt am Main: Suhrkamp.

⁹ This proficiency is interactional too: with some people we have it, with others we do not. Social skill is an interactional phenomenon as much as social understanding is. This also bears on social developmental disorders like autism (De Jaegher, H. (2006). *Social Interaction Rhythm and Participatory Sense-Making: An embodied, interactional approach to social understanding, with implications for autism*, Unpublished D.Phil. Thesis: University of Sussex, Brighton, UK).

have more ‘high-level’ possibilities at our disposal, for instance, we can ask for an explanation, e.g. “what did you mean?” And we can also employ our capacity for rational thinking or put ourselves in the other’s shoes in order to find out the reasons for their comportment. The latter may remind of TT or ST, but the kinds of skills that TT and ST refer to need to be rethought from an interactional, intersubjective framework. We do occasionally explicitly think about why an other did what she did, or think about what we would do in her situation, but the origins of such a capacity need to be re-imagined from an intersubjective perspective. They do not spring from individual heads, but have their basis in intersubjective engagement and participatory sense-making, developmentally as well as pragmatically, to borrow an argument from Gallagher (2001).

With regard to direct perception, breakdowns are the place where social perception can be ‘enriched’ the most – and ‘on the fly’: in breakdown the other is opaque to me, but its recovery can lead to a better, richer understanding of the other, and to a change in the ongoing course of interacting and participating in sense-making. Gallagher himself indeed advocates an account of social cognition based on interaction in several places (2004; Gallagher & Hutto, 2008) but without an in-depth account of the role of the interaction process in social cognition the opportunity for a real change of perspective in social cognition research would be hopelessly lost.

6. Conclusion

While traditional approaches conceive of understanding as something besides perception, Gallagher believes they are part of one and the same process. In the paper discussed here, he chose to argue for this by inserting understanding into perception as it were. This move of bringing perception and understanding together is artificial in light of their foundational unity. If cognition is the active engagement through which an organism brings forth its world, then social perception can only already be understanding (i.e. be direct perception) if it is based directly in the dynamic patterns of intersubjective engagement, and brought out of the individual in this way.

Gallagher is right to assert that there is no problem of other minds in the sense that others for us are puzzles to solve. This is indeed not the way understanding each other works. Others are not opaque to us. But they are also not continually transparent to us. We still work at understanding them. But rather than doing that from our strictly individual vantage point, we do it in interaction (and sometimes we collaborate towards this goal, and sometimes we don’t). Meaning is generated and transformed by people in interaction, it is not transported between them.

The interactional proposal presented here changes the problem space of social cognition research. That there will be open issues is to be expected. To list just a few of them: How do the autonomy of the interaction partners and that of the interaction pattern relate? What happens when interaction is significantly asymmetric (e.g. in mother-infant or teacher-student dyads)? How do more ‘removed’ forms of social understanding (e.g. understanding someone over the phone, through letters, or understanding a play or a film) develop? What does an account of social understanding look like that does justice both to the *subjectivity* and *autonomy* of those engaged in social encounters and to the fact that intentions evolve and change in and through their interaction? How can we move from an account of social cognition to one of intersubjectivity?

The present paper offers a way in which accounts such as Gallagher’s can be extended so as to definitively break away from the individualist terms of the debate set by cognitivist approaches. In doing so, it also advocates such a break as the path to follow for a properly intersubjective, not a merely spectatorial, approach to interpersonal understanding. One that takes seriously both the subjectivity and autonomy of those involved in social encounters, and the phenomenon that meaning is interactively generated and transformed.

Acknowledgments

This paper owes a lot to comments by Shaun Gallagher, Steve Torrance and an anonymous reviewer, and to discussions I’ve had with Sanneke de Haan, Ezequiel Di Paolo and Thomas Fuchs. Writing it was made possible by the Marie Curie Research Training Network DISCOS (Disorders and Coherence of the Embodied Self).

References

- Auvray, M., Lenay, C., & Stewart, J. (in press). Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, doi:10.1016/j.newideapsych.2007.12.002.
- Bavelas, J. B., Coates, L., & Johnson, T. (2002). Listener responses as a collaborative process: The role of gaze. *Journal of Communication*, 52, 566–580.
- Condon, W. S., & Ogston, W. D. (1971). Speech and body motion synchrony of the speaker-hearer. In D. L. Horton & J. J. Jenkins (Eds.), *The perception of language*. Columbus, Ohio: Charles E. Merrill Publishing Co.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4), 485–507.
- De Jaegher, H., & Di Paolo, E. (2008). Making sense in participation. An enactive approach to social cognition. In F. Morganti, A. Carassa, & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective to the study of interactions*. Amsterdam: IOS Press.
- Di Paolo, E. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 97–125.
- Di Paolo, E., Rohde, M., & Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing. *New Ideas in Psychology*, 26(2), 278–294.
- Di Paolo, E. (in press). Extended life. *Topoi*.
- Di Paolo, E., Rohde, M., & De Jaegher, H. (in press). Horizons for the enactive mind: Values, social interaction, and play. In J. Stewart, O. Gapenne & E. Di Paolo (Eds.), *Enaction: Towards a new paradigm for cognitive science*. Cambridge, MA: MIT Press.
- Gallagher, S. (2001). The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies*, 8(5–7), 83–108.

- Gallagher, S. (2004). Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry and Psychology*, 11(3), 199–217.
- Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University Press.
- Gallagher, S. (2008a). Direct perception in the intersubjective context. *Consciousness and Cognition*, 17(2), 535–543.
- Gallagher, S. (2008b). Inference or interaction: Social cognition without precursors. *Philosophical Explorations*, 11(3), 163–174.
- Gallagher, S., & Hutto, D. D. (2008). Understanding others through primary interaction and narrative practice. In J. Zlatev, T. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity*. Amsterdam: John Benjamins.
- Gallagher, S., & Zahavi, D. (2007). *The phenomenological mind: An introduction to philosophy of mind and cognitive science*. London: Routledge.
- Gill, S. P., Kawamori, M., Katagiri, Y., & Shimojima, A. (2000). Role of body moves in dialogue. *International Journal of Language and Communication*, 12.
- Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology*. New York: Irvington Publishers.
- Goodwin, C. (1984). Notes on story structure and the organization of participation. In M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 225–246). Cambridge: Cambridge University Press.
- Granic, I. (2000). The self-organization of parent-child relations: Beyond bidirectional models. In M. D. Lewis & I. Granic (Eds.), *Emotion development and self-organization. Dynamic systems approaches to emotional development* (pp. 267–297). Cambridge: Cambridge University Press.
- Hobson, R. P. (2002). *The cradle of thought*. London: Macmillan.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). *Rhythms of dialogue in infancy: Coordinated timing in development*. Vol. 66. Oxford: Blackwell.
- Kendon, A. (1982). Coordination of action and framing in face-to-face interaction. In M. Davis (Ed.), *Interaction rhythms* (pp. 351–363). New York: Human Sciences Press.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and Cognition. The Realization of the Living*. Dordrecht: Reidel.
- Ramseyer, F., & Tschacher, W. (2006). Synchrony: A core concept for a constructivist approach to psychotherapy. *Constructivism in the Human Sciences*, 11(1), 150–171.
- Ratcliffe, M. (2007). *Rethinking commonsense psychology: A critique of folk psychology. Theory of Mind and Simulation*. Hampshire/New York: Palgrave Macmillan.
- Sheets-Johnstone, M. (1999a). Emotion and movement: A beginning empirical-phenomenological analysis of their relationship. *Journal of Consciousness Studies*, 6(11–12), 259–277.
- Sheets-Johnstone, M. (1999b). *The Primacy of movement*. Amsterdam: John Benjamins.
- Sheets-Johnstone, M. (in press). Body and movement: Basic dynamic principles. In S. Gallagher & D. Schmicking (Eds.), *Handbook of Phenomenology and Cognitive Science*. Dordrecht: Springer.
- Thompson, E. (2007). *Mind in life: Biology phenomenology and the sciences of mind*. Harvard University Press.
- Trevarthen, C. (2005). Action and emotion in development of cultural intelligence: Why infants have feelings like ours. In J. Nadel & D. Muir (Eds.), *Emotional development* (pp. 61–91). Oxford: Oxford University Press.
- Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry*, 42(1), 3–48.
- Varela, F. J. (1991) (pp. 79–107). Dordrecht: Kluwer.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience* (6th ed.). Cambridge, MA: MIT Press.
- Warner, R. M. (1996). Coordinated cycles in behavior and physiology during face-to-face social interaction. In J. Watt & C. A. VanLear (Eds.), *Cycles and dynamic patterns in communication processes*. Newbury Park, CA: Sage.
- Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2), 97–125.