UNIVERSITY OF COPENHAGEN

The interactive Turn in Social Cognition Research

A Critique

Overgaard, Søren; Michael, John Andrew

The Interactive Turn in Social Cognition Research: A Critique

Søren Overgaard and John Michael

Proponents of the so-called "interactive turn in social cognition research" maintain that mainstream research on social cognition has been fundamentally flawed by its neglect of social interaction, and that a new paradigm is needed in order to redress this shortcoming. We argue that proponents of the interactive turn ("interactionists") have failed to properly substantiate their criticisms of existing research on social cognition. Although it is sometimes unclear precisely what these criticisms of existing theories are supposed to target, we sketch two possibilities: interactionists can either accept the primary explanandum addressed by mainstream social cognition research – namely mindreading – and claim that interactionism contributes some hitherto neglected but necessary component of a successful explanans, or they can argue that mainstream research has focused on a misconceived explanandum. We argue that interactionist claims of both sorts are problematic.

1. Introduction

The so-called "interactive turn in social cognition research" (De Jaegher et al., 2010) is made up of a family of views (henceforth "interactionism"), which includes the "second-person approach" (Hutto, 2004; Reddy, 2008; cf. Ratcliffe 2007, ch. 6), "interaction theory" (e.g. Gallagher, 2001, 2004), and "the enactive approach" (e.g. De Jaegher and Di Paolo, 2007; De Jaegher, 2009a). These theories are unified by the claim that mainstream research on social cognition has been fundamentally flawed by its neglect of social interaction, and that a complete overhaul of research on social cognition is needed in order to redress this shortcoming.[1]

---

[1] Often, the emphasis on interaction is coupled with a commitment to enactive approaches to cognition. An evaluation of these approaches is beyond the scope of this paper, although we touch upon enactivism in section 4.3.

In this paper, we argue that the defenders of these views have failed to properly substantiate their criticisms of existing research on social cognition. We identify two forms of interactionist criticism: interactionists can either accept the primary *explanandum* addressed by mainstream ongoing research on social cognition – namely mindreading – and claim that interactionism offers an important corrective to existing *explanantia,* or they can argue that mainstream research has focused on a misconceived *explanandum*. After giving a brief characterization of the interactive turn (Section 2), we examine both of these options. In Section 3 we argue that interactionist claims of the first sort are all either implausible, or compatible with existing research, or not "interactionist" in any obvious sense, and in Section 4 we show that proposals of the second sort also fail to stand up to scrutiny.

2. Mindreading and The Interactive Turn

2.1 The Mindreading Debate

Most researchers working on social cognition have been concerned to explain the set of abilities that has been referred to variously as "theory of mind," "mentalizing," or most commonly nowadays, "mindreading." In a classic paper, two primatologists defined "theory of mind" as follows:

> In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others (either to conspecifics or to other species as well). (Premack & Woodruff 1978, p. 515)

Although the term "mindreading" has since tended to replace "theory of mind" as the preferred label for the ability or capacity in question – because the former avoids begging the question of whether the capacity is to be explained in terms of the possession of a *theory* – the general understanding of the capacity in question has not changed. Alvin Goldman (a prominent contributor to the debate) and Chandra Sripada give the following characterization of mindreading:

> Mindreading is the capacity to identify the mental states of others, for example, their beliefs, desires, intentions, goals, experiences, sensations and also emotion states. (2005, p. 193)

Shaun Nichols and Stephen Stich, two other prominent theorists, approach the topic of mindreading via a description of a character in Salman Rushdie's *Midnight Children*, who, besides being telepathic, also has the more mundane ability to understand others that the rest of us have. Of this latter ability, Nichols and Stich write:

> He detects *fury* in the face of one teacher and *amusement* and *perplexity* in the gestures of another. And he further surmises that even if he confessed his [telepathic] ability to his teachers, his teachers wouldn't *believe* him. It is this kind of mental state attribution, prediction, and explanation that we will explore …. In the burgeoning literature on the capacity of ordinary people to understand the mind, 'mindreading' has become a fashionable label for this ability. (Nichols & Stich, 2003, pp. 1-2)

'Mindreading', then, picks out our perfectly ordinary ability to understand others as having mental states of various sorts, including emotions, sensations, beliefs, and desires, as well as the sorts of predictions and explanations of their behavior that such an understanding makes possible.[2] The mindreading debate is about *how* we accomplish this – that is, about the nature of the strategies, routines, or underlying mechanisms and processes that are involved when we mindread.

There are two well-known families of positions in the mindreading debate, the basic ideas of which may be outlined as follows:

> *Theory Theory (TT):* We mindread by utilizing a rich body of information about mental states and how they are connected with other mental states, with observable behavior, and with the environment. This body of information is either:
>
>> *(a)* A "theory" that is formed on the basis of observation, testing, and learning more generally (e.g., Gopnik & Wellman, 1992); or:
>> *(b)* Contained in a "module" that is activated at some point in development (e.g., Leslie, 1994; cf. Baron-Cohen, 1995)

---

[2] Some theorists prefer more restrictive definitions of the term 'mindreading', which limit the term to, for example, cases involving the application of mental concepts (Hutto, forthcoming; for discussion, see Gordon, 2008), or inferences (Gallagher, 2008a). We think that such a move would pre-judge important questions, such as what concepts are, or to what extent mental state attributions must rely on inference (some deny that it must; see, e.g. Gordon, 1995), and we believe that a broad and neutral characterization of the *explanandum* minimizes the risk of unproductive terminological disputes. Moreover, as the quotes from prominent contributors to the mindreading debate illustrate, it seems to be this wide definition that at least some of the main players adopt. Note that mindreading, as Goldman, Stich, Nichols and others understand it – and as we understand it here – is not co-extensive with social cognition or understanding as such. There are plenty of ways of understanding and thinking about others that do not involve understanding or thinking about their *mental states*, a few examples of which will be discussed in section 4.

Simulation Theory (ST): We mindread by putting ourselves in other people's "shoes," using our own mind to work out what we would do, think or feel in their situation – and then attributing those intentions, thoughts, or emotions to those other people (e.g., Gordon, 1986; Goldman, 1989; Heal, 1995).

To see how TT and ST accounts of mindreading differ *in concreto*, consider the following simple example. You have been informed that some injustice has just happened to an acquaintance of yours. You run into her immediately after the event has occurred, and you notice that her face is flushed and she is frowning. You conclude that she is angry. How did you reach that conclusion? According to (a simplified version of) TT, you connected the information about the injustice and the visual information about her frowning with stored general information about people, of something like the following kind: *ceteris paribus*, people to whom an injustice has been done tend to be angry; *ceteris paribus*, people who frown tend to be angry. Together, the information about this particular case and the stored 'theoretical' information allow you to infer that your acquaintance is angry. A simplified version of ST, on the other hand, would maintain that you used the visual and other information about the particular case in a different way. Instead of connecting this with general assumptions about people and what makes them upset, you imagine how you would feel if someone had done the relevant sort of injustice to you (or how you would feel if you made a face like that) and then attribute the result of this 'simulation' to your acquaintance.

That both the example and the rival accounts of what goes on in it are highly simplified is obvious, but irrelevant. The point of outlining them is to make clear how the two sorts of reply work: how they generate a conclusion that attributes a mental state to another person either by applying stored general information about mental states and how they are connected with other

mental states, behaviour etc. (TT), or by modelling or simulating with our own mind what the target might be going through (ST).[3]

2.2. The Interactive Turn

The interactive turn that has recently been proclaimed can be traced back to Shaun Gallagher's (2001) proposal, which he labeled "interaction theory". Gallagher argued that TT and ST approaches neglect the interactive contexts in which social cognition is embedded, and thereby overlook embodied social processes that are engaged in interactions, and which are important components of social cognition. Drawing on a broad array of empirical and theoretical sources – including developmental psychology, dynamical systems research, phenomenology, and philosophy of mind – various authors have recently developed Gallagher's original criticisms further (and in some cases radicalized them). In a recent paper, Gallagher himself maintains that "interaction theory" is "posed as a challenge to both TT and ST ... by shifting the very framework, by questioning the very suppositions, that TT and ST assume to be in place when it comes to

---

[3] An anonymous reviewer suggests that TT and ST accounts dodge the real issue – namely how information about mental states get into the theory, module or simulation in the first place. We don't think this objection is on target. Modular theorists can argue that evolution has furnished us with a specialized module that biases us to treat others as having mental states if they exhibit certain properties, such as having faces or being self-propelled, and possibly applies core mental concepts in doing so (Carey, 2009). Child scientist TT, by contrast, maintains that the information is gathered much like scientists gather information about the world. And the core idea of ST is precisely the idea that we don't *need* any general information about mental states; we can simply use our own minds to model the target's states. Having said that, the objector has a point to the extent that all theories seem to presuppose that an infant must have some innate grasp of other people as special, as agents or "minded creatures" in some minimal sense, for the project of simulating them or theorizing about them to even make sense (see, e.g., Gopnik & Wellman, 1992, p. 150).

understanding others" (Gallagher, 2008b, p. 165). Matthew Ratcliffe (2007, p. 2) denies that his interactionist account is an "elaboration, revision or supplementation of the orthodox account of folk psychology. Instead, it departs from that account in just about every respect." Finally, De Jaegher maintains that "The interactional proposal … changes the problem space of social cognition research" (2009a, p. 541), indeed "changes the research landscape and turns the endeavour of understanding social cognition on its head" (2009b, p. 549).

What the various interactionist approaches have in common is their conviction that mainstream research in social cognition offers a fundamentally mistaken picture of how individuals understand each other, and that this is due to a failure to take interaction seriously. It is useful to distinguish two very different versions of this critique. Some interactionist criticisms seem to target the particular accounts of mindreading provided by TT and ST. Such a critique would involve accepting that our ability to mindread is the (or a) right *explanandum* while maintaining that existing accounts have failed to provide the right *explanans*. Other interactionist criticisms are more radical in that they seem to question whether mindreading is a legitimate explanandum. We discuss the former in the next section, and the latter in section 4.

3. Critiquing the explanantia

Although most interactionists are arguably not interested in mindreading (for reasons we will discuss in section 4), some interactionists have suggested ways in which their accounts challenge not the mindreading framework per se but TT and ST as *accounts* of mindreading. In this section, we consider two sorts of suggestions: interactionists might offer an account that *replaces* the existing accounts – a new and better account of the mechanisms involved in mindreading.

Alternatively, the suggestion could be that interactionism *supplements* one or more of the existing paradigms by adding a crucial, but hitherto neglected element to the account (or accounts) in question. We examine proposals of both sorts and argue that any plausible points interactionists have made about the relevance of interaction to mindreading are much harder to set firmly apart from the mainstream theories than interactionists often seem to think.

3.1 Interactionism as a Competitor to Theory Theory and Simulation Theory

In rough outline, what might an independent interactionist account of mindreading be? As far as we are aware, there are two potential candidates for an independent interactionist account of mindreading: Shaun Gallagher's theory of "direct social perception" (DP) and Daniel Hutto's "Narrative Practice Hypothesis" (NPH). We discuss them in turn.

DP is part of a complex of theories that Gallagher calls "interaction theory" (Gallagher, 2008b), other elements of which will be discussed in sections 3.2, 4.1 and 4.2. For present purposes, DP can be understood as making the following claim:

> *DP:* In some cases, the information that your perception (e.g., your visual perception) makes available to you *already* includes information about the other person's mental states.

Recall the example we used to illustrate TT and ST accounts of mindreading above: you run into an acquaintance of yours to whom an injustice has just happened, and you notice her flushed and frowning appearance. According to DP, to see your friend's flushed cheeks and frowning

expression might constitute *seeing* that she is angry and thus obviate the need for any theoretical inference or simulation exercise. What you can see you don't need to infer (Gallagher, 2008a).

Note that DP, no less than TT and ST, has an immediately graspable reply to the question of how we are able to mindread.[4] For it says, quite simply, that we sometimes *perceive* – in the same way that we perceive, say, that a car is parked in the driveway – that others are angry, sad, in pain, and so on. Gallagher calls this "smart perception" and illustrates it with an example taken from outside the domain of mindreading. Normally, when you look at your (red) car, it is not the case that you see "a certain unrecognized red mass with a specific shape", which you then need to interpret "in some non-visual, non-perceptual cognitive steps that go beyond perception itself" in order to work out that it is your car (Gallagher, 2008a, p. 536). Rather, you simply see your car. The case is similar with respect to your frowning acquaintance: you don't see her eyebrows as shaped in a particular way and *infer* from this that she is angry; rather, you immediately recognize the anger in her frowning expression. So far so good. Yet in order to evaluate the DP proposal we need to inquire whether DP really constitutes a fundamental alternative to TT and ST type accounts of mindreading. We believe there is good reason to think the answer is no.

First of all, DP only constitutes the alternative to existing accounts of mindreading that interactionists are looking for if the existing accounts are committed to denying what DP affirms. Gallagher is emphatic that "the standard versions of TT and ST" assume that "the other

---

[4] Gallagher (2008a) regards his proposal as a critical corrective to assumptions underpinning the mindreading debate, rather than an account of mindreading. This depends on his more narrow understanding of the mindreading *explanandum*. Given that Gallagher wants to claim that "the intentions and emotions of other persons are perceptually interpreted in movements, gestures, postures, facial expressions and contextualized behaviours" (2001, p. 103), he is, however, making a claim about mindreading as we understand it. And as noted above (note 2), we don't think our understanding of the mindreading *explanandum* is unduly broad. At any rate, the points we raise in this section also apply (*mutatis mutandis*) to DP construed as a critical response to the mindreading *explanandum* (see note 9).

person's mental states are hidden away and are therefore not accessible to perception" (Gallagher, 2008a, p. 536). Indeed, he suggests, this is precisely the reason "extra-perceptual cognitive elements seem to be required" (ibid.) in the form of simulation routines or theoretical inferences. It is, however, hard to find simulationists explicitly defending the view that all mental states are "hidden away"; in fact, prominent advocates of ST gesture at the very opposite view. Discussing the claim "that in some circumstances some mental states of others can be the objects of direct perception", Jane Heal remarks: "Nothing I have said is meant to rule out this idea; and exploration of its connections with simulationism might be of interest" (Heal, 1995, p. 50). Similarly, Robert Gordon avers that "There should be no conflict between ST ... and Gallagher's [view] ... that our primary and pervasive way of engaging with others rests on 'direct', non-mentalizing perception of the 'meaning's of others' facial expressions, gestures and intentional actions" (Gordon, 2008, p. 221; cf. Gordon, 1986, p. 169).

The link between TT and the idea of the mental as being "unobservable" may seem more secure, yet some advocates of modular TT have said things that seem compatible with DP, including Brian Scholl[5] and Peter Carruthers. In his recent book *The Opacity of Mind* the latter states that "we often *see* the person as pleased at a compliment, for example, or *hear* someone as expressing a judgment" (Carruthers, 2011, p. 70). Carruthers makes a similar point in a forthcoming paper: "the phenomenology of much everyday mindreading is that we just *see* someone as being about to act in some specific way in pursuit of a presumed goal, or *hear* the intent behind what they say".[6] This certainly rules out conscious, personal-level inferences as being the normal everyday route to an awareness of another's intention, and seems consistent with the idea of

---

[5] See Scholl & Tremoulet, 2000, a paper to which Gallagher repeatedly refers for scientific backing for DP.

[6] Carruthers, Mindreading in Infancy, revised draft, p. 5. Available from:
http://www.philosophy.umd.edu/Faculty/pcarruthers/ (accessed June 5th, 2012).

"smart perception" being at work. It has even been argued that child scientist TT need not maintain that other's mental states are hidden, since the construction of an adequate theory can render hitherto unobservable entities perceptible (Lavelle, 2012).

This is obviously not to say that there are no defenders of TT and ST who deny what DP affirms. Nor does it follow from what we have said that defenders of TT and ST would agree with Gallagher about the scope of DP. Whereas Gallagher thinks "that for the most part … direct perception delivers sufficient information for understanding others" (Gallagher, 2008a, p. 540), other theorists might hold that social perception only works for *some* mental states (e.g., basic emotions) and not others (e.g., beliefs) and thus does not deliver all that we need for understanding others. But note that granting this difference already implies that the idea of some binary opposition between TT and ST, on the one hand, and DP, on the other, is wrong. Rather, we seem to have a continuum of positions, some of which (say, Carruthers' version of TT and Gordon's ST) are closer to DP than others.[7]

Irrespective of what individual defenders of ST and TT may or may not have said, however, there is a more general reason to be sceptical of the idea that DP constitutes a genuine alternative to TT and ST. As Gregory Currie (2008, p. 215) has suggested, TT and ST are best understood as research programmes, rather than specific theories. This means that "they are somewhat unspecific proposals about how to explain intersubjective understanding, and they make specific predictions only in conjunction with auxiliary hypotheses" (ibid.). TT, as it is sometimes put, attributes "information-rich" strategies to mindreaders, strategies that crucially rely on a rich

---

[7] Besides, depending on how Gallagher's proposal is interpreted, the question is whether it does not overstate the case. Many mental states, including beliefs and desires, have propositional content that can be very complex. It seems highly unlikely that I might see your belief that "Noah wasn't the only person who could have survived the flood" (Apperly and Butterfill, 2009, p. 960).

body of general information about the mind. ST, by contrast, attributes "information-poor" strategies to mindreaders (see Nichols and Stich, 2003). Thus, what is distinctive of ST is that it claims mindreaders use their own emotional, motivational, and other resources to work out what others are thinking and feeling. The important point for present purposes is that what we have referred to as "strategies" can be understood to refer both to personal-level routines and strategies *and* to sub-personal mechanisms and processes. If so, however, there is no reason in principle why the sub-personal processes underlying a perceptual experience could not be of either the information rich or the simulation variety (Herschbach, 2008).[8] And if so, then DP would turn out be a specific theory belonging to one of the existing research programmes, rather than an alternative to them.[9]

DP, then, does not constitute a fundamental alternative to TT and ST, because it remains unclear to what extent the picture DP overturns – the picture of the mental as "hidden" and imperceptible – is one that informs ST and TT across the board. Having said that, it should not be forgotten that Gallagher proposes DP as part of a more comprehensive theory – "interaction theory" – that "gives interaction a central role" (Gallagher, 2009, p. 547). Thus, it will also be necessary to consider whether DP *in conjunction with the emphasis on interaction* (i.e. as Gallagher

---

[8] Interactionists might object that to apply personal-level concepts to subpersonal processes is to commit a category mistake. But there is nothing particularly personal-level about the notion of "information-rich" cognitive processes.

[9] Note that if DP is construed as an attack on the mindreading *explanandum* – rather than on the TT and ST explanations of mindreading – our criticisms still hold. *Pace* Gallagher, if TT and ST are understood as broad research programmes in the way indicated, there is no reason to view them – or the debate between them – as being premised on the assumption "that the other person's mental states are hidden away and are therefore not accessible to perception" (Gallagher, 2008a, p. 536). Obviously, the fact that several main contributors to the debate explicitly *deny* that assumption only confirms our point.

intends his proposal to be taken) states something that cannot be accommodated by existing theories. We will return to this below (in section 3.2), and again our conclusion will be negative.

Unlike DP, the Narrative Practice Hypothesis (NPH) has explicitly been offered as a straightforward interactionist competitor to TT and ST. The NPH, writes Hutto, "competes directly with empirical variants of TT and ST that attempt to explain the basis of our [folk psychological] competence" (Hutto, 2008, p. 186). However, whereas TT and ST, as we understand them here, are rival accounts of how we mindread others, the NPH is introduced as a theory about what drives the *development* of mindreading abilities: "Distinctive kinds of narrative encounters are what first allow us to develop our folk psychological competence" (Gallagher and Hutto, 2008, p. 28; cf. Hutto, 2008, p. 177). This developmental story in itself may not be incompatible with TT and ST accounts of mindreading. But the NPH comprises a further claim. It "denies that in developing this ability [i.e., the ability to mindread], we are somehow acquiring a guiding theory or set of principles that are then stored in our minds", or an ability to simulate others (Hutto, 2008, p. 178). In other words, if it is through "narrative encounters" that we acquire our core mindreading abilities, and if this acquisition does not involve assembling a rich body of information of the TT sort, nor triggering a module containing the relevant information, nor developing simulation routines, then TT and ST accounts of how we mindread must be false. It cannot be the case that we draw on a body of general information about mental states and how they interconnect, for example, if acquiring the ability has nothing to do with making such information available to the child. Thus, it seems we have the rough outline of a truly interactionist account of mindreading that is a direct competitor to the existing theories.

Indeed, in addition to an account of the development of mindreading, the NPH also offers an account of *how* we mindread. The guiding idea here seems to be that mindreading is something like a practical ability: "after appropriate training, children become *skilled in the practice*

of making sense of actions in FP [i.e., folk-psychological] terms" (Hutto, 2008, p. 178; emphasis added). As Gallagher and Hutto put it, children "develop an implicit practical understanding of *how* to make sense of persons as those who act for reasons" (Gallagher & Hutto, 2008, p. 29). In line with the last quote, philosophers often highlight the distinction between practical abilities and theoretical knowledge by distinguishing "knowing how (to)" and "knowing that". Couched in these terms, the NPH maintains that developing mindreading competence consists in acquiring knowledge of how to make sense of others in mentalistic terms. And it also maintains – *contra* ST – that the relevant "knowing how" is not, and does not crucially involve, knowing how to simulate others.

It seems to us, however, that the NPH is in danger of collapsing into a version of TT. For it relies upon the irreducibility of (practical) knowing-how to (theoretical) knowing-that; otherwise practical competency with narratives may be explained by an abstract (theoretical) understanding of mental states and the relations among them. But some recent work in philosophy calls the distinction between knowing-how and knowing-that into question across the board (Stanley & Williamson, 2001; Snowdon, 2004). Moreover, even if such criticism is misguided and there are genuine examples of knowing-how that do not reduce to, or crucially involve, knowing-that, the question is whether what the NPH describes may plausibly constitute such an example. Some of what Hutto himself writes strongly suggest that it does not. For he seems to link the mindreading competence that children acquire via engagement in "folk-psychological narratives" with an understanding of how mental states are connected with other mental states and with contextual factors. Such narratives, he says, "put on show how these attitudes can integrate with one another (and also how they fit with other mental states and stand with respect to other contextual factors)" (Hutto, 2008, p. 178). By engaging in such narrative practices, including "discussing what

the story characters know, feel or want", "children learn how these states of mind behave in relation to each other and other terms in the psychological family" (Gallagher & Hutto, 2008, p. 29).

As plausible as it is that engagement with a certain kind of narratives can teach children these things and thereby help to develop their mindreading abilities, this explanation seems to cast serious doubt on Hutto's claim that the NPH is a real alternative to the existing theoretical approaches. For the sort of understanding children are said to gain is not plausibly construed exclusively in terms of knowing-how. Rather, it also seems to be a matter of knowing *that* mental states are connected in such-and-such ways with other mental states and contextual factors. Indeed, if children did not abstract general principles from the various narratives they become acquainted with, it is difficult to see how they could learn to combine and modify them, etc., and thereby to bring them to bear upon novel situations. For they would be limited to memorizing specific narratives, which would hardly enable them to achieve understanding in real-life social situations that unavoidably differ significantly in the details. In other words, it seems to us that the NPH must also appeal to precisely the sort of general knowledge of the mental to which TT appeals in attempting to explain our ability to mindread. Thus, in order for the NPH to be established as an alternative to TT, its advocates must address the challenge of explaining how an ability to understand an open-ended range of varying real-life situations can be developed on the basis of engaging in narrative practices, if no abstraction of general principles is involved.

We suspect this conclusion points to a more general lesson. Emphasizing the importance of interaction – including the sort of interaction involved in engaging in narrative practices – is not sufficient to give an account of mindreading that competes directly with TT and ST. For an account is needed of *how* interaction contributes to mindreading. And here notions such as theory-formation or the honing of simulation skills might slip back into the picture. Furthermore,

as we will go on to show in the next section, some versions of TT and ST are emphatic about the importance of social interaction.

3.2 Interaction as Enabling Mindreading

If interactionism is not a direct competitor of TT and ST, then one obvious suggestion would be that social interaction might play some foundational, enabling role for mindreading. Some theorists seem to hold that it is our ability to mindread that enables us to negotiate our social environment. For example, this is the view expressed by Currie and Sterelny (2000, p. 145): "Mind-reading and the capacity to negotiate the social world are not the same thing, but the former seems to be necessary for the latter." The interactionist view, on the current reading of it, would be that this gets things exactly backwards. It is mindreading that "builds on," "emerges from," or has its "basis" in social interaction (e.g., Fuchs & De Jaegher, 2009; De Jaegher, 2009a), not the other way around. Though mindreading might be an enabling condition for certain, perhaps rather rare, sorts of interaction, there is a much more fundamental and ubiquitous sense in which social interaction is what enables us to mindread.

As a claim about the evolution of mindreading, this is very likely to be true. It was undoubtedly the highly social nature of our ancestors' lives that created the evolutionary pressure to develop mindreading skills. But hardly anyone, it seems to us, would suppose otherwise. As Tooby and Cosmides (two evolutionary psychologists with TT sympathies) state, "Humans evolved this ability [to mindread] because, as members of an intensively social, cooperative, and competitive species, our ancestors' lives depended on how well they could infer what was on one another's minds" (1995, p. xvii).

It is more plausible to interpret the interactionist claim as pertaining to development (cf.

Carpendale & Lewis, 2004). The claim would then be that without intense social interaction a child is unlikely to develop a normal capacity for mindreading. Again, there would seem to be much truth in this suggestion; but, as before, it is unclear that this is a claim that advocates of TT and ST are committed to denying. The "child-scientist" version of TT would surely affirm the claim enthusiastically. Thus, Andrew Meltzoff and Alison Gopnik have proposed that "If we want to find the origins of common-sense psychology a good place to look might be in infant interactions with and understanding of persons" (Meltzoff & Gopnik, 1993, p. 336), drawing particular attention to the importance of mutual imitation games for the development of mindreading abilities.[10] Along similar lines, Gopnik and Repacholi have suggested that toddlers' ability to understand other people's desires may in large part depend on "emotionally charged interactions" with other people (Repacholi & Gopnik, 1997, p. 19). Studies by Perner, Ruffman and colleagues on the so-called "sibling effect" indicate "the importance of social interactions for facilitating belief understanding" (Ruffmann et al., 1998, p. 172), which on their view spells trouble for at least some versions of modular TT, while being compatible with "child scientist" (and ST) views (Perner, Ruffman, & Leekam, 1994; Ruffman et al. 1998). According to this version of TT, children actively experiment with their social environment; thus, without interaction with others, the child would lose its most important source of social information to develop a fully representational understanding of beliefs and other mental states (see Meltzoff, Gopnik & Repacholi, 1999, p. 36).

"Modular" TT, on the other hand, would suggest that without social interaction of the right sort or in the right degree, the Theory-of-Mind (ToM) module might be prevented from coming "online" at the normal stage of development. According to Scholl and Leslie (1999, p. 139), "the relevant sort of environmental input (i.e. social interaction) [is] necessary to trigger and tune the

---

[10] Although Meltzoff and Gopnik (1993, p. 338) acknowledge that imitation games constitute only "a subset of interactive games".

maturation of the ToM module." Simulationists, too, are unified in attributing a crucial role to social interaction in development. Goldman (2006, p. 196) cites evidence that joint play facilitates the refinement of mindreading skills in childhood (i.e. more than solitary pretence) to document the importance of a "vivid deployment of simulational exercises." Thus, it could be in social interaction that children learn to correct for differences between themselves and others when running simulations of others (cf. Harris, 1996, pp. 215-16) – so that, for example, children eventually becomes able to set aside their own preference for cookies over broccoli when predicting what an adult might prefer.[11]

In order to state something that goes beyond what the existing theories are able to accommodate, then, interactionists must make a stronger claim. One possibility is to maintain that *current* social interaction is needed in order to enable or facilitate mindreading. Clearly, this suggestion can be interpreted in various ways. The *very* strong version of the claim maintains that current interaction is necessary for mindreading (or even social cognition) quite generally. As far as we know, no one has defended this view – which is fortunate, since the view seems implausible. To see this, note that a common, and quite reasonable, interactionist complaint against the classic version of the false-belief task is that it puts children in the unnatural position of passive onlookers rather than interacting agents (Gallagher, 2001, p. 99; Hutto, 2009, p. 225; Ratcliffe, 2007, p. 54). Yet, as the many years of research in this field demonstrate conclusively, most children of five or more years nevertheless pass the test (Wellman et al., 2001; Doherty, 2009). Clearly, then, it is possible to demonstrate social understanding in the absence of current interaction.

---

[11] An anonymous reviewer remarks that ST and TT accounts generally maintain that the development of mindreading abilities is something that happens early in ontogeny and then is in place once and for all, whereas interactionists believe our abilities keep evolving and developing throughout life. This point seems to indicate some limitations in the way the mainstream literature has tended to study mindreading. But it deserves to be mentioned that recent research has explicitly aimed to redress these limitations (see, e.g., Apperly 2011).

A weaker claim would be that current interaction is necessary for *some* mindreading processes. Interactionists sometimes suggest that the ability to "directly perceive" others' mental states (cf. DP) depends on interaction with others. De Jaegher, for example, writes: "we may experience an other's feelings and intentions directly, but direct perception builds on something, namely on skillful interaction with others" (De Jaegher 2009a, p. 538). Although this is not De Jaegher's point,[12] it could be maintained that the ability skillfully to interact with others is an insufficient foundation for direct social perception: what is needed is current interaction. We suspect that this claim, too, would be implausibly strong. If there is a sense in which we may directly perceive others' emotions when we are interacting with them, then surely we may also do this when we are merely observing them. Theatre, film, television, and even sports events would be a lot less interesting if we did not have the ability to "perceive" the feelings and emotions of people with whom we are in no way interacting. If you can perceptually detect, say, from a person's facial expression that she is feeling sad when you are talking to her, then surely you can do the same when she is talking with someone else (and you are just observing). At the very least, whoever wants to deny this has to bear the burden of proof and show why these two situations are so different that we must think of them as involving two very different mindreading processes.

If, however, interactionists wanted to make the claim that current interaction often *facilitates* mindreading,[13] then again they would have a good case. For example, there is evidence that young children perform better on interactive versions of the false belief task (Chandler & Hala, 1994; Wellman et al. 2001, pp. 666-667; Buttelmann et al., 2009). Yet two points must be

---

[12] The "basis" De Jaegher refers to seems to be a developmental basis, and thus her point is one that was addressed in the previous sub-section. Here we are exploring *possible* argumentative moves that an interactionist could make; we are not attributing any of these moves to De Jaegher or to anyone else.

[13] Gallagher (2008b, p. 168), for example, claims that "our sense of the other person's behavior is helped along through our continued interactions".

emphasized here. First, we see no reason why TT or ST should be committed to denying this weak interpretation of the interactionists' claim. If children mainly develop and employ their mindreading skills in interaction with others, as many defenders of TT and ST maintain, then it is not surprising that they would be better at employing those skills when interacting with others than when passively observing. Second, interactionists would do well to resist very strong versions of the claim about interaction facilitating mindreading (or social cognition more generally). For example, it seems plausible that in some cases (a heated discussion might be an example) it may be easier for a detached onlooker to detect one party's sadness, say, than for the other party to the interaction to do so. We all know how being too involved can blind us to things that are obvious to a bystander. Thus, it is unlikely that current social interaction *always* facilitates "direct" emotion detection (or indeed social cognition more generally). In fact, it is not even clear that it does so as a general rule. For example, there is evidence to suggest that whether this is the case may depend on the sort of emotion a subject is detecting. Subjects more readily perceive so-called approach-oriented emotions (joy and anger) when the target's gaze is directed at them – i.e., when the subject is not just an onlooker, but also someone looked at – than when it is averted. However, the perception of so-called avoidance-related emotions (fear and sadness) seems enhanced if the target's gaze is averted (Adams & Kleck, 2005). Again, then, there seems little room for an interactionist proposal that is neither implausible nor fully compatible with existing proposals.

4. Rejecting the *Explanandum*

Perhaps the contribution that the interactive turn makes to social cognition research is not a contribution to the mindreading debate but, rather, a critical corrective to fundamental assumptions

underpinning that debate. In this section, we examine the most common interactionist criticisms of the mindreading debate, and argue that none of them are successful.

4.1 The Scope of Mindreading

The first criticism centers on the idea that mindreading may be a lot less pervasive a feature of our social lives than contributors to the mindreading debate typically think. Some seem to construe mindreading as our exclusive way of understanding others.[14] Baron-Cohen, for example, has claimed that "it is hard for us to make sense of behavior in any other way than via the mentalistic ... framework. We just can't help doing it this way" (Baron-Cohen, 1995, p. 3). According to many interactionists, this is not just false: it turns the real state of affairs completely on its head. We have plenty of other ways of making sense of other people, interactionists maintain, and we only resort to mindreading in rare cases where those other, more basic means of understanding give out (Gallagher, 2008a, p. 540; 2008b, pp. 165, 170).[15]

Some of our social interactions – perhaps driving on a busy road might be an example – seem to a large extent governed by social norms (Ratcliffe, 2007, pp. 86-94). We expect that

---

[14] But this assumption is not shared by all participants to the mindreading debate. Goldman (2006, p. 210) for example, suggests that "there are many varied forms of social cognition, not all of which involve understanding mental states." Apperly (2011, p. 4) makes the same point.

[15] Gallagher, of course, associates mindreading with inferring unobservable mental states. Thus, it is not clear that in suggesting that we only rarely mindread, he intends to reject the idea that we routinely understand others in terms of their mental states. Note, though, that Gallagher sometimes seems to gesture towards the latter view. He writes, e.g., that "In most situations, we are not trying to mindread the other person; we *are not concerned about the other person's mental states*" (2012, p. 194; emphasis added). And: "What we call social cognition is often nothing more than … social interaction" (2008a, p. 540).

people will do certain things, such as drive on a particular side of the road, avoid bumping into other vehicles, signal before turning, etc., not – it could be maintained – because of what (we assume) they are thinking or feeling, but because people are supposed to act like that. Some such norms are institutionalized legal norms such as the Highway Code, others may be more implicit, and no doubt a great many of them vary across cultures (e.g., how close you stand to another person when conversing with them). Rather than ponder what individual drivers might be thinking, interactionists could maintain, by far the most efficient way of figuring out on which side of the road they will drive is to assume that they will do what one is supposed to do (in the country in question). Nor does it seem to us, in situations such as this, that we do consider what the other drivers are thinking or feeling.

In other cases, we seem to understand others largely or even exclusively in terms of the social roles they occupy. As José Luis Bermúdez (2003) argues, for example, our ability to interact smoothly with waiters, butchers, and bus drivers depends not on our attributing beliefs, desires, or emotions to them, but on our identifying them *as* waiters, butchers, etc. Once their social role is clear we know very well what to expect from them and how to explain their actions – that is, as long as they continue to act as waiters and butchers are supposed to act in relation to customers. Thus, in many routine social interactions involving clearly defined social roles, we need not – and, according to Bermúdez, do not – mindread at all.

The claim that the scope of mindreading is more restricted than often assumed has recently been met with substantive criticism (Spaulding, 2010). We shall not repeat these criticisms here. In fact, for the sake of argument, we are prepared to grant that the interactionists have a valid point. In a case such as the one highlighted by Bermúdez, not only does it not seem to us that we understand the waiter's behavior – when he approaches with the menu, say – in terms of what he is thinking or feeling, it is far from clear that we need to do so. Once the social role is defined, I can

interact perfectly well with the waiter without considering his beliefs, desires, or other mental states.

What is important, however, is that it cannot be an attractive option for proponents of the interactive turn to restrict the scope of mindreading too much. The examples discussed above are of cases where, as it seems, there is no reason for us to be interested in other people's mental states. And indeed, such cases abound in everyday life. Yet there are also situations where others' thoughts, desires and feelings are extremely important to us, and in such cases it seems implausible to claim that we generally do not resort to mindreading. To take a clear example, think about being on a first date with someone. In this sort of situation you carefully monitor your partner's expressions, interpret their utterances, etc., in an attempt to decipher their thoughts and feelings. Your are constantly on the lookout for signs that they are getting bored, or upset by something you said, as well as, more positively, indications that they are having a good time, that they like you, feel comfortable etc.

Note two things about this example. First, while it is sometimes suggested that even when we are interested in the thoughts and feelings of others, we rarely mindread, since a much more common reaction is simply to ask them how they feel (Gallagher & Zahavi, 2008, p. 193), this strategy seems neither common nor very wise in the sort of example we are considering.[16] Constantly asking your date how they are feeling, or what they are thinking, is hardly the thing to do in the circumstances. Indeed, even outside the context of dating, it would not infrequently be socially awkward to ask such questions as "What are you feeling?" out of the blue. And when we

---

[16] Though we will not argue this point in detail here, we suspect this suggestion is problematic for other reasons as well. On many accounts, verbal communication requires the ability to interpret others as having communicative intentions, and this is a (rather sophisticated) mindreading ability.

do ask people about their feelings, it is often because we have reason to suspect that they are feeling something in particular – i.e., when we have already "mindread" them in the sense discussed here.

Secondly, in the dating situation we do not resort to mindreading because of the breakdown of some more immediate and original way of relating to others. When we are dating someone for the first time, we are simply extremely concerned about their thoughts, emotions and the rest, and that is why we mindread them – not because some smooth, non-mentalizing interaction with them has been called to a halt, thereby rendering them opaque to us (*pace* Gallagher, 2008a, p. 540). Indeed, they may be anything but opaque, and we may be able to (mind)read them like an open book: it may just be obvious to you that your partner is bored, for example. The crucial point is that we *do* mindread, and we do so because dating is all about what the other person is thinking, feeling, and desiring.

The question, of course, is just how unusual the dating situation is. On this point, people's intuitions are likely to diverge considerably. We think that it is just a somewhat extreme case of something very common, and that we routinely mindread in many of our interactions with colleagues, friends, and family, whose mental life generally is of some concern to us. Others will disagree with this, and it is hard to see how the question could be settled conclusively. Yet for our purposes, it is sufficient that there should be common examples of mindreading that are not the result of the disruption of some more basic form of social cognition, and a first date is one very clear such example. Thus, even if mindreading is not as ubiquitous as Baron-Cohen and others suppose, it cannot plausibly be reduced to a compensatory strategy adopted only in rare cases when more primary forms of social understanding and engagement give out. Hence, this first interactionist criticism of the mindreading debate fails.

4.2 Detached and Spectatorial

Another criticism that interactionists frequently make of both major sides of the mindreading debate is that the latter suppose that people primarily adopt a "detached," "spectatorial," or "third-personal" attitude in their relations with others (Fuchs & De Jaegher, 2009, p. 468; Hutto, 2004, p. 549; Ratcliffe, 2007, p. 86; Reddy, 2008, p. 7). Interactionists deny precisely this. In the words of Gallagher, interactionism "rejects the spectatorial supposition that we are primarily spectators or observers of others' behaviors. Our normal everyday stance toward the other person is not third-person, detached observation; it is second-person interaction" (Gallagher, 2008b, p. 164).[17]

Our reply to this criticism is twofold. First, it strikes us as doubtful whether we have *one* "normal everyday stance" toward other people. Surely, in the course of any one day, we not only interact with others in various ways – ranging from personal conversations with good friends to navigating through crowds of strangers on busy sidewalks – but we also, and not infrequently, simply observe people. In fact, if the "spectatorial" attitude were so unnatural for us to adopt, it is hard to understand why people would watch so much television – in particular so-called "reality shows," where you simply watch other people engage in sometimes quite unremarkable everyday activities. Furthermore, when we interact with people it is far from clear that we are not *also*, often at least, observing them, monitoring their reactions to what we are saying, trying to decipher their body language, etc. Think, again, of a couple on their first date. While there is a lot of interacting going on, one also monitors the reactions of one's partner closely, thinks about what he or she might be thinking, and so on.

Something similar can be said with respect to the second-person/third-person dichotomy. If a "second-person" stance is exemplified by the stance we adopt when we address

---

[17] Again, "second-person interaction is *the* primary and ordinary way of encountering the other person" (Gallagher, 2001, p. 99; emphasis added).

someone in a conversation ("you"), then presumably a third-person stance is what we adopt when we talk *about* a person ("he," "she"). But as soon as this is made explicit, we see that the idea of a single everyday stance has to be wrong. If there is one thing that people like to talk about it is other people – especially absent people. Not only is there no reason to think that our everyday social lives are built around the adoption of a single attitude, but in fact we very often combine observational/third-personal attitudes effortlessly with interactional/second-personal ones.

Our second point targets the notion that defenders of ST or TT are committed to viewing our most fundamental stance with regard to others as spectatorial and third-personal. Interactionist critics appear simply to assume that this is the case. After caricaturing the conception of psychological understanding that TT offers by likening it to a theory that "one of Star Trek's famously unemotional Vulcans" would construct if given the task of understanding human psychology, Reddy (2008, p. 24) asserts that "any theory-theory is necessarily wedded to this type of detached understanding." Nor does ST fare much better, Reddy maintains. Both theories "run the risk of positing an observer and an observed, of thinking of 'mind-reading' primarily as a 'spectatorial' process" (Reddy 2008, p. 25).

But why should defenders of ST or TT accept this? We do not see any reason why these theoretical paradigms must imply the (surely false) claim that human beings are primarily observers of other people. Nobody denies that we interact with others in all sorts of ways (cf. Carruthers, 2011, p. 231). Theory theory of the "child-scientist" stripe is probably the *prima facie* most "spectatorial"-sounding of the available alternatives, but even advocates of that view are clear that they do not wish to cast doubt on the fundamental nature of the second-person perspective. Alison Gopnik, one of the most consistent advocates of the child-scientist theory, writes, for example:

The new research in developmental psychology tells us that quite literally from the moment we first see other people, we see them *as* people. … To see someone as a person is to see a face, not a mask; a "thou," not an "it." (Gopnik, Meltzoff, & Kuhl, 2001, p. 24)

Nor does Gopnik mean to deny or downplay the importance of interaction for the development of mindreading skills, as is illustrated by her suggestions that children's understanding of others' desires are dependent upon "emotionally charged interactions", and that infants' interactions with others is the place to look for the origins of our mindreading abilities. Children, on this account, are neither detached, nor are they simply passive onlookers. "Development in social cognition depends on two-way traffic between self and other", as Meltzoff, Gopnik and Repacholi (1999, p. 19) wrote more than a decade ago. The same conclusion emerges from theory theorists' work on the sibling effect (mentioned in section 3 above) and related effects, such as parenting style. Ruffman, Perner and Parkin's study of the latter, for example, adds, as they put it, "to a growing body of evidence which shows that children's understanding of mental states develops in interactions with others" (Ruffman et al., 1999, p. 409).

Here one should also note the emergence of more ecological, interactive experimental designs in social cognition research. Meltzoff and Gopnik's (1993, p. 338) "mutual-imitation games", which they argued played an important role in the development of social cognition abilities, were obviously interactive. Repacholi and Gopnik's (1997) experiments employed an interactive experimental design encouraging children to offer food to an experimenter. As already mentioned, more interactive versions of the false-belief test have also been developed (e.g. Avis & Harris, 1991; Chandler & Hala, 1994). In their meta-analysis of the literature on false-belief tasks, Wellman, Cross and Watson (2001, p. 666) admit that "Often children are essentially passive

onlookers; for example, they watch as someone transfers Maxi's chocolate from one place to another"; and they go on to note that actively involving the child significantly improves performance (ibid., p. 667). Even researchers studying the neural mechanisms underlying social interaction have taken steps to make experimental designs more interactive. EEG lends itself especially well to use in interactive contexts (e.g. Tognoli et al., 2007), but interactive fMRI designs (i.e. "hyperscanning") are also becoming increasingly prevalent despite the technical challenges resulting from the need for subjects to remain isolated and relatively motionless (e.g. King-Casas et al., 2005; Schilbach et al., 2009; Schippers et al., 2010). As Ralph Adolphs (2006, p. 32) writes in a passage about the importance of ecological experimental designs, "there is no need to dwell on this issue because it is universally acknowledged, and because it is in fact now being surmounted."

The reason we mention these developments in experimental design is that they illustrate just how wide of the mark the accusation that mainstream approaches conceive of social cognizers as detached really is. Perhaps, as Wellman et al. suggest, traditional experimental designs did put subjects in the position of unengaged observers. But the fact that the importance of addressing this shortcoming is "universally acknowledged" shows that it must be universally acknowledged, too, that mindreaders, and social cognizers more generally, are not mainly spectators. For anyone who held the view that social cognizers are primarily detached observers should maintain that interactive paradigms are precisely *not* more ecologically valid. Hence, it is difficult to see the interactionist criticism of the "spectatorial" nature of the existing approaches as anything but a straw man.

4.3 Individualism

Some interactionists are after even bigger game: they want to challenge the *individualism* that they take to be at the heart of all accounts of mindreading. De Jaegher thus complains that there is a

sense in which DP is as problematic as ST and TT "because it does not dispel the idea that social cognition is something done in the individual head" (De Jaegher, 2009a, p. 535). We can distinguish a strong and a weak version of anti-individualism. The strong version – which we do *not* attribute to De Jaegher – would maintain that social cognition, quite generally, is *wholly* a matter of processes outside the individual. This strong position, however, seems inconsistent with a wealth of empirical studies, such as lesion studies showing that localized brain damage can selectively impair subjects' performance on social cognition tasks (e.g. Goldman & Sripada, 2005).

The weak version maintains that social cognition must not be reduced to a question solely about what is going on inside the cognizing individual. So on this reading, the claim is merely that one must also take into account matters in the individual's environment – in particular, the various sorts of interaction the individual has with others. As we saw above, when we discussed interactionism in the context of the mindreading debate, there is no question that most participants to that debate accept this weak form of anti-individualism. Surely, mindreading is not only a question of what goes on inside the mindreader, but also very much a question of the sorts of information the social environment makes available to the mindreader, and interaction is a crucially important way to gather such information (cf., e.g. Adolphs, 2006). Is there any scope, then, for an anti-individualist position that is neither implausible nor true-but-uncontroversial?

In a series of recent papers De Jaegher and colleagues (e.g. De Jaegher, 2009a; De Jaegher & Di Paolo, 2007; De Jaegher, Di Paolo & Gallagher, 2010), draw upon enactivist approaches in cognitive science to develop such a position. According to enactivist approaches, cognition occurs when autonomous systems actively regulate their interactions with the external world in such a way that the conditions for their own continued existence are maintained (Varela et al., 1991; Thompson, 2007; Di Paolo et al., 2008). Applying enactivism to social cognition, De Jaegher and colleagues argue that social interactions sometimes satisfy these criteria and thus count

as autonomous systems that perform (social) cognition. Further, they argue that there is a range of cases in which the social cognition performed by such emergent systems (i.e. by interactions) "replaces individual mechanisms" such as mindreading (De Jaegher et al., 2010, p. 441) and thus cannot be accommodated by individualist accounts.[18] In order to substantiate this position, De Jaegher and co-authors refer to an experiment by Auvray et al. (2009) that they believe documents a case where social interaction "constitutes" social cognition. Something constitutes – or is a "constitutive element" of – a case of social cognition if it is "part of the processes that produce" the social cognition, De Jaegher et al. (2010, p. 443) explain. They continue: "A constitutive element is part of the phenomenon (it must be present in the same time frame as the phenomenon). The set of all constitutive elements is the phenomenon itself" (ibid.).

Let us look more closely at the supposed example of interaction constituting social cognition. In the experiment, two blindfolded participants each move a sensor along a shared virtual 1-D line using a computer mouse. Unbeknownst to the participants, a "shadow" is connected with their sensor and copies the latter's movements at a fixed distance. Whenever participants encounter an object – the other's sensor, the "shadow," or a stationary object – they receive a tap on the finger. The participants' task is to click on the mouse whenever they judge that they are in contact with the other participant.

It turns out that participants consistently find each other, and a clear majority of their clicks occur when their sensors are actually in contact with each other. De Jaegher and co-authors stress that the results cannot be explained by appealing to individual capabilities to detect the contingency of stimuli. Participants are as likely to click when encountering the other participant's shadow as when encountering the other's sensor (De Jaegher et al., 2010, p. 445). What explains the

---

[18] Similarly, Gallagher claims that there are cases where the interaction "makes social cognition what it is" (2012, p. 189).

larger number of clicks when sensor encounters sensor is simply the higher frequency of sensor-sensor encounters. This, in turn, is explained by "the stability of the coupling" when sensor encounters sensor.

The experiment has to do with social cognition insofar as participants judge when they have encountered a social stimulus, i.e. the other participant.[19] However, the experiment fails to provide an example of interaction "making social cognition what it is", thereby "replacing individual mechanisms." The crucial point is a simple one: the actual (sensor-sensor) interaction has no influence on, and thus certainly is not part of, the processes that lead participants to form their social judgments. As De Jaegher, Di Paolo, and Gallagher themselves emphasize, the judgments are equally likely to occur when participants encounter the other's shadow – i.e., when there is no real interaction. Rather, the process of forming judgments would be best explained by supposing that participants rely on a simple heuristic to the effect that if an object moves, it is the other participant. This would explain why they click indiscriminately when they encounter the other participant's shadow and when they encounter his or her sensor. What the interaction does – i.e., the encounter with the other participant's sensor – is simply to yield many more opportunities for the participants to employ the theory and make the judgment (and, of course, thereby judge correctly).

How might interactionists respond to this? We want to end by sketching two possible replies. One of these, we will suggest, offers a plausible as well as non-trivial articulation of the

---

[19] Might interactionists deny that it is in virtue of those social judgements that the experiment has to do with social cognition? Perhaps, but then they would need to explain what *does* make the experiment have to do with social cognition, and we cannot imagine what that might be. The only remotely plausible candidate would be the fact that two people are interacting. But this would be to understand social interaction as sufficient for social cognition. And then the claim that interaction might be a constitutive element in some social cognition processes would be true by definition. This move is obviously question-begging; and it renders the introduction of the perceptual crossing experiment – which was supposed to *show* that interaction can constitute social cognition – pointless.

interactionist idea that social interaction can constitute social cognition, but comes at the price of abandoning the more radical anti-individualist claims. The other reply holds out the promise of rescuing the latter claims by questioning the cogency of the argument we have just offered; but we argue that this reply is defective. We start by considering the latter.

We have denied that social interaction was involved in the processes that lead participants to judge that they have encountered the other player and on this basis rejected the claim that the social interaction dynamics could be a constitutive element of participants' social cognition. But surely, what matters when assessing whether something constitutes social cognition is not merely whether it plays a crucial role in the subjects' formation of (just any) social judgments as such, but whether it plays a role in ensuring that they reach the *right* judgment. And in the perceptual crossing experiment, the "collective dynamics" precisely plays a crucial role in explaining subjects' successful performance – it uniquely explains the high number of clicks in the sensor-sensor situation.[20] Thus, one could reach the conclusion that the experiment *is* a case where social interaction constitutes social cognition and thereby replaces individual mechanisms.

This reply does not work, however Ensuring a successful outcome does not entail constituting social cognition. For if participants received electric shocks whenever they moved away from the other participant's sensor (outside a certain zone of proximity, say) this would probably ensure a high proportion of sensor-sensor clicks too. But it seems absurd to suggest that the electric shocks or the mechanism administering them might "constitute" social cognition. It seems reasonable to demand that only something that plays a central role in the participants' formation of their social judgments – including, of course, factors that plays such a role in their formation of *correct* judgments – could be said to "constitute" their social cognition. Factors that contribute to performance but bypass subjects' judgment formation – such as the somewhat cruel

---

[20] We are grateful for an anonymous referee for pressing this objection.

example we have imagined – are too "external" to be part of the cognitive process.

Note that in saying this, we are not taking an "internalist" view for granted. In particular, we are not ruling out that interaction *could* play a constitutive role in some social-cognitive processes. In fact, we think recent data might indicate that in some cases it does. These data indicate the second, more promising, line of response that interactionists can give to our argument.

In the follow-up study, the three objects are given unique auditory tags. Thus, the feedback participants receive when crossing an object is no longer indistinguishable from one object to the next. As a result, participants become able to distinguish among the three objects, and this removes their tendency to click when encountering the "shadow". In this new experiment, then, individuals are able to detect not only movement but also contingency – and, as the authors are at pains to emphasize, the latter achievement is not the result of a contingency detection modules being triggered, but of a learning process (Lenay & Stewart, 2012).

We agree that, in this new setup, the interaction does seem to be a constitutive element in a successful process of contingency detection, but, crucially, not in the sense that the interaction replaces individual processes. Rather, it is only with the help of the auditory tags the different objects can be distinguished, thereby making it possible to perceive distinct patterns in their behaviors. Thus, we still need an account of how the individual player processes the auditory information and uses it in arriving at the right judgments. So interaction, *in tandem with individual processes,* seems to be doing the work. This seems an important enough conclusion, which underscores the importance of interaction for social cognition in a range of cases[21]. But, in contrast to the apparently more far-reaching claims of some enactivists, the results still do not show that

---

[21] Indeed Froese and Di Paolo (2010) can be read as suggesting this: they claim that interaction plays a causal role in individual learning processes.

interaction replaces individual cognitive processes.[22] Thus, the notion that social cognition is at least partly done in individual heads seems secure, and interactionists have again failed to indicate a serious flaw in the mindreading explanandum.

5. Conclusion

We have identified two sorts of interactionist criticisms of mainstream social cognition research – criticisms of standard explanations of mindreading and criticisms targeting the explanandum as such – and argued that neither sort has proved successful so far. This does not mean that social interaction is not crucially important to social cognition, or that the latter should not be studied in ecological, interactive situations; for as we have seen, there is nearly universal agreement on both these points. Rather, it means that proponents of interactionism face a choice: either they revise and refine their criticisms in such a way as to address the points we have raised in this paper, or they abandon their revolutionary rhetoric vis-à-vis mainstream social cognition research.

References

Adams, R. B. and Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, 5, 3-11.

---

[22] Of course, as an anonymous reviewer correctly points out, this only means that the experiments on perceptual crossing so far conducted fail to support radical anti-individualism. Thus, we have not shown that the enactivist version of anti-individualism is false. And what we have established certainly does not warrant any general conclusions about enactivism, a comprehensive evaluation of which is beyond the scope of this paper. For a critique of the interactionists' use of the perceptual crossing experiment that does include some general conclusions about enactivism, see Herschbach (2011).

Adolphs, R. (2006). How do we know the minds of others? Domain-specificity, simulation, and

enactive social cognition. *Brain Research,* 1079, 25-35.

Apperly, I. (2011). *Mindreaders: The Cognitive Basis of "Theory of Mind".* Hove, UK:

Psychology Press.

Apperly, I. and Butterfill, S. Do Humans Have Two Systems to Track Beliefs and Belief-like

States? *Psychological Review*, 116 (4), 953-970.

Auvray, M., Lenay, C., and Stewart, J. (2009). Perceptual interactions in a minimalist virtual

environment. *New Ideas in Psychology*, 27(1), 32-47.

Avis, J. and Harris, P. H. (1991). Belief-Desire Reasoning among Baka Children. *Child

Development*, 62, 460-467.

Bahrami, B., Olsen, K., Latham, P., Roepstorff, A, Rees, G. and Frith, C. (2010). Optimally

interacting minds. *Science*, 329, 1081-5.

Baron-Cohen, S. (1995). *Mind-Blindness: An Essay on Autism and Theory of Mind*. Cambridge,

MA: The MIT Press.

Bermúdez, J. L. (2003). The domain of folk psychology. In A. O'Hear (ed.), *Minds and Persons*.

Cambridge: Cambridge University Press.

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false

belief understanding in an active helping paradigm. *Cognition*, 112, 37–342.

Carey, S. (2009). *The Origin of Concepts*. New York: Oxford University Press.

Carpendale, J. I. M. and Lewis, C. (2004). Constructing and understanding of mind: The

development of children's social understanding within social interaction. *Behavioral and

Brain Sciences*, 27, 79-96.

Carruthers, P. (2011). *The Opacity of Mind.* Oxford: Oxford University Press.

Chandler, M. and Hala, S. (1994). The Role of Personal Involvement in the Assessment of early

    False Belief Skills. In C. Lewis and P. Mitchell (eds.), *Children's Early Understanding of*

    Mind: Origins and Development. Hove, UK: Psychology Press

Csibra G., Gergely, G., Szilvia Biro, S., Kiro, O., Brockbank M. (1999). Goal attribution without

    agency cues: the perception of 'pure reason' in infancy. *Cognition*, 72, 237–267.

Currie, G and Sterelny, K. (2000). How to think of the modularity of mind-reading. *The*

    Philosophical Quarterly, 50, 145-160.

De Jaegher, H. (2009a). Social understanding through direct perception? Yes, by interacting.

    Consciousness and Cognition, 18, 535-542.

De Jaegher, H. (2009b). What made me want the cheese? A reply to Shaun Gallagher and Dan

    Hutto. *Consciousness and Cognition*, 18, 549-550.

De Jaegher, H. and Di Paolo, E. (2007). Participatory sense-making: An enactive approach to

    social cognition. *Phenomenology and the Cognitive Sciences*, 6, 485-507.

De Jaegher, H., Di Paolo, E. and Gallagher, S. (2010). Can social interaction constitute social

    cognition? *Trends in Cognitive Sciences*, 14, 441–447.

Doherty, M. J. (2009). *Theory of Mind: How Children Understand Others' Thoughts and Feelings*.

    Hove, UK: Psychology Press.

Froese, T., and Di Paolo, E. (2010). Modeling social interaction as perceptual crossing: An

    investigation into the dynamics of the interaction process. *Connection Science*, 22 (1), 43-68.

Fuchs, T. and De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and

    mutual incorporation. *Phenomenology and the Cognitive Sciences*, 8, 465-486.

Gallagher, S. (2001). The practice of mind: Theory, simulation or primary interaction? *Journal of*

    Consciousness Studies, 8, 83-108.

Gallagher, S. (2004). Understanding Interpersonal Problems in Autism: Interaction Theory as An
Alternative to Theory of Mind, *Philosophy, Psychiatry, & Psychology*, 11 (3), 199-217.

Gallagher, S. (2008a). Direct perception in the intersubjective context. *Consciousness and
Cognition*, 17, 535-543.

Gallagher, S. (2008b). Inference or interaction: social cognition without precursors. *Philosophical
Explorations*, 11, 163-174.

Gallagher, S. (2009). Deep and dynamic interaction: Response to Hanne De Jaegher.
Consciousness and Cognition, 18, 547-548.

Gallagher, S. (2012). In defense of phenomenological approaches to social cognition: Interacting
with the critics. *Review of Philosophy and Psychology*, 3 (2), 187-212.

Gallagher, S. and Hutto, D. D. (2008). Understanding others through primary interaction and
narrative practice. In J. Zlatev, T. P. Racine, C. Sinha, and E. Itkonen (eds.), *The Shared
Mind*. Amsterdam: John Benjamins.

Gallagher, S. and Zahavi, D. (2008). *The Phenomenological Mind: An Introduction to Philosophy
of Mind and Cognitive Science*. London: Routledge.

German, T. P. and Leslie, A. M. (2004). No (social) construction without (meta-)representation:
Modular mechanisms as a *basis* for the capacity to acquire an understanding of mind.
Behavioral and Brain Sciences, 27, 106-107.

Goldman, A. I (1989). Interpretation psychologized. *Mind and Language*, 4, 161–185.

Goldman, A. I. (2006). *Simulating Minds*. New York: Oxford University Press.

Goldman, A. I. and Sripada, C. S. (2005). Simulationist models of face-based emotion recognition.
Cognition, 94, 193-213.

Gopnik, A. (1996). The scientist as child. *Philosophy of Science,* 63 (4), 485-514.

Gopnik, A., Meltzoff, A. and Kuhl, P. (2001). *How Babies Think*. London: Phoenix.

Gopnik, A. and Wellman, H. (1992). Why the child's theory of mind really *is* a theory. *Mind & Language*, 7 (1-2), 145-171.

Gordon, R. (1986). Folk psychology as simulation. *Mind and Language*, 1 (2), 158-171.

Gordon, R. (1995). Simulation Without Introspection or Inference from Me to You. In M. Davies and T. Stone (eds.), *Mental Simulation*. Oxford: Blackwell.

Gordon, R. (2008). Beyond mindreading. *Philosophical explorations*, 11, 219-222.

Harris, P. (1996). Desires, beliefs, and language. In P. Carruthers and P. K. Smith (eds.), *Theories of Theories of Mind*. Cambridge: Cambridge University Press.

Heal, J. (1995). How to Think About Thinking. In M. Davies and T. Stone (eds.), *Mental Simulation*. Oxford: Blackwell.

Herschbach, M. (2008). Folk psychological and phenomenological accounts of social perception. *Philosophical Explorations*, 11, 223-235.

Herschbach, M. (2011). On the role of social interaction in social cognition: a mechanistic alternative to enactivism. *Phenomenology and the Cognitive Sciences*. DOI: 10.1007/s11097-011-9209-z

Hutto, D. D. (2004). The limits of spectatorial folk psychology. *Mind & Language*, 19, 548-573.

Hutto, D. D. (2008). The Narrative Practice Hypothesis: clarifications and implications. *Philosophical Explorations*, 11 (3), 175-192.

Hutto, D. D. (forthcoming). Elementary mind minding, enactivist-style, In A. Seemann (ed)., *Joint Attention: New Developments in Philosophy, Psychology, and Neuroscience.* Cambridge, MA: MIT Press.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S. R. and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science*, 308, 78–83.

Konvalinka, I., Vuust, P, Roepstorff, A. and Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, 4, 1-11.

Lavelle, J. S. (2012). Theory-theory and the direct perception of mental states. *Review of Philosophy and Psychology*. DOI 10.1007/s13164-012-0094-3

Lenay, C. and Stewart, J. (2012). Minimalist approach to perceptual interactions. *Frontiers in Human Neuroscience,* 6 (98). DOI: 10.3389/fnhum.2012.00098

Leslie, A. M. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. A. Hirschfeld and S. A. Gelman (eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119-148). Cambridge: Cambridge University Press.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31, 838–50.

Meltzoff, A. N. and Gopnik, A. (1993). The role of imitation in understanding persons and developing a theory of mind. In S. Baron-Cohen, H. Tager-Flusberg and D. J. Cohen (eds.), *Understanding Other Minds: Perspectives from Autism* (pp. 335-366). Oxford: Oxford University Press.

Meltzoff, A. N., Gopnik, A. and Repacholi, B. (1999). Toddlers' understanding of intentions, desires, and emotions: Explorations of the dark ages. In P. D. Zelazo, J. W. Astington and D. R. Olson (eds.), *Developing Theories of Intention* (pp. 17-41). Mahwah, NJ: Lawrence Erlbaum Associates.

Meltzoff, A. N. and Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75–8.

Nichols, S. and Stich, S. (2003). *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford: Oxford University Press.

Perner, J., Ruffman, T. and Leekam, S. (1994). Theory of Mind is Contagious: You Catch It from Your Sibs. *Child Development*, 65, 1228-1238.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4, 515-526.

Ratcliffe, M. (2007). *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*. Basingstoke: Palgrave Macmillan.

Reddy, V. (2008). *How Infants Know Minds*. Cambridge, MA: Harvard University Press.

Reddy, V. and Morris, P. (2009). Participants don't need theories: Knowing minds in engagement. In I. Leudar and A. Costall (eds.), *Against Theory of Mind* (pp. 91-107). Basingstoke: Palgrave Macmillan.

Repacholi, B. and Gopnik, A. (1997). Early Reasoning About Desires: Evidence from 14- and 18-Month-Olds. *Developmental Psychology*, 33, 12-21.

Ruffman, T. (2004). Children's understanding of mind: Constructivist but theory-like. *Behavioral and Brain Sciences*, 27, 120-121.

Ruffmann, T., Perner, J., Naito, M., Parkin, L. and Clements, W. A. (1998). Older (but Not Younger) Siblings facilitate False Belief Understanding. *Developmental Psychology*, 34, 161-174.

Ruffman, T., Perner, J. and Parkin, L. (1999). How Parenting Style Affects False Belief Understanding. *Social Development*, 8, 395-411.

Schilbach, L., Wilms, M., Eickhoff, S., Romanzetti, S., Tepest, R., Bente, G., Shah, N., Fink, G. and Vogeley, K. (2009). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, 22, 2702–2715

Schippers, M., Roebroeck, A., Renken, R. Nanetti, L. and Keysers, C. (2010). Mapping the information flow from one brain to another during gestural communication. *Proceedings of*

the National Academy of the Sciences, 107, 9388-9393.

Scholl, B. and Leslie, A. M. (1999). Modularity, development, and 'theory of mind'. *Mind & Language*, 14, 131-153.

Scholl, B. and Tremoulet, P. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4 (8), 299-309.

Snowdon, P. (2004). Knowing how and knowing that: A distinction reconsidered. *Proceedings of the Aristotelian Society*, 104, 1-29.

Spaulding, S. (2010). Embodied cognition and mindreading. *Mind & Language*, 25, 119-140.

Stanley, J. and Williamson, T. (2001). Knowing How. *The Journal of Philosophy*, 98 (8), 411-444.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press.

Tognoli, E., Lagarde, J., De Guzman, G. and Kelso, J. (2007). The phi complex as a neuromarker of human social coordination. *Proceedings of the National Academy of the Sciences*, 104, 8190–8195.

Tooby, J. and Cosmides, L. (1995). Foreword. In S. Baron-Cohen, *Mindblindness*. Cambridge, MA: The MIT Press.

Varela, F. J., Thompson, E. and Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Wellman, H. M., Cross, D. and Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655-684.