

Social Constraints on the Direct Perception of Emotions and Intentions

Shaun Gallagher · Somogy Varga

Published online: 8 October 2013
© Springer Science+Business Media Dordrecht 2013

Abstract In this paper, we first review recent arguments about the direct perception of the intentions and emotions of others, emphasizing the role of *embodied* interaction. We then consider a possible objection to the direct perception hypothesis from social psychology, related to phenomena like ‘dehumanization’ and ‘implicit racial bias’, which manifest themselves on a basic bodily level. On the background of such data, one might object that social perception cannot be direct since it depends on and can in fact be interrupted by a set of cultural beliefs. We argue, however, that far from threatening the idea of direct perception, these findings clearly contradict the idea of hardwired theory of mind modules. More generally, we suggest that in order to further the understanding of social cognition we must take seriously insights about in-group and out-group distinctions and related phenomena, all of which are currently neglected in the mainstream social cognition literature.

Keywords Direct perception · Theory theory · Simulation theory · Interaction theory · Embodiment · Dehumanization

1 Introduction

The problem of other minds, or of how we come to understand others is a central issue in philosophy and

psychology. The epistemological version of this problem involves the question of how to justify the possibility of gaining knowledge about entities (minds) that purportedly are not observable in the external world. One way of formulating this problem is in terms of an asymmetry between our direct access to and knowledge of our own experiences, and the lack of direct access in the case of the experiences of others. On the background of such a gap it might seem that a sceptic could challenge our entitlement to knowledge.

On most accounts, then, the problem of social cognition is framed as a problem of access to the other person’s mind. The supposition is precisely that the other person’s mental states are hidden away and are therefore not accessible to perception. Since psychological (cognitive, affective, motivational, experiential, etc.) states are imperceptible, we have access only to behavior or action, which is not meaningful unless supplemented by some kind of an additional inferential operation. This ‘principle of imperceptibility’ is nicely stated by Johnson:

Mental states, and the minds that possess them, are necessarily unobservable constructs that must be inferred by observers rather than perceived directly. (Johnson 2000, p. 22).¹

Accepting this principle motivates the question of whether our epistemic access to other minds works in a way that is

¹ This idea is frequently expressed in the literature on social cognition. For example, Dretske, states it clearly: “We cannot see other minds. They are unobservable. You can see the smile (at least the upturned mouth) but not the thought ‘behind’ it. You can see the perspiration, the flushed face, the wrinkled forehead, the squint, the jerky motion of the arms ... but you cannot see the fear, the embarrassment, the frustration, the desire ...” (Dretske 1973, 36). Likewise, Wellman (1990, p. 107), defending this view writes: “Mental states, such as beliefs and desires, are private, internal and not observable in others.”

S. Gallagher (✉)
Department of Philosophy, University of Memphis, Clement
Hall 331, Memphis, TN 38152, USA
e-mail: s.gallagher@memphis.edu

S. Varga
Department of Philosophy, University of Memphis, Clement
Hall 327, Memphis, TN 38152, USA

fundamentally different from, structurally similar to, or analogous to those processes by which we acquire knowledge about other domains (Stueber 2006). The two standard “theory of mind” (ToM) answers to the epistemological problem of other minds, simulation theory (ST) and theory theory (TT) generally accept the ‘principle of imperceptibility’ and posit some extra-perceptual cognitive step (inference or simulation) as necessary for “mindreading” the mental states of others. Hence, much of the discussion is concerned with identifying the mechanisms or rules by which we habitually infer or simulate our way to others’ mental states on the basis of their behavior. TT holds that the standards that guide our inferences come from an intuitive theory, which, although non-scientific, is comparable with bodies of scientifically structured knowledge about phenomena and their causes. In other words, TT argues that we have recourse to a set of folk-psychological rules that allow us to infer an explanation of the observed behaviour in terms of mental states and dispositions (e.g., Frith and Frith 2007). The competing account, ST, suggests that the standards that guide mindreading are not tied to theoretical knowledge, but come from ourselves: to understand what is going on in other minds, we rely on simulation routines that enable us to mirror and reproduce the other’s inner states in our own system, based on our own first-person experience (e.g., Goldman 2006). Despite such disagreements about the nature of the process, both TT and ST subscribe to the ‘principle of imperceptibility’ and hold that when we perceive others, we perceive mere bodily activity, patterns of mechanical movements that warrant the inference or suggest the correct simulation to the other’s intentional states. I may register what the other person does, but until I call forth some theory, or until I run through a simulation routine, I seem not to have any sense of what that person is up to or what the behavior means.

Yet, this principle sometimes appears to be paradoxical.

Mental states are unobservable, and have complex logical properties... if anything, we should have expected that mental state concepts should be bafflingly difficult to acquire, and yet even the most unremarkable child seems to understand them—without any explicit teaching (Baron-Cohen and Swettenham 1996, p. 158)

Indeed, somewhat surprising from the perspective of TT, infants as young as 13 months have been shown capable of understanding actions, intentions, and, purportedly, beliefs (Onishi and Baillargeon 2005). Perhaps for this reason some theorists have begun to suggest that the idea that some mental states may be perceptible is not inconsistent with TT. They stop short, however, of endorsing a direct perception view. Thus, for example, Lavelle (2012, 228) concludes: “The moral is that while theoretical entities need not be

unobservable, one requires a theory in order to observe them.” And, according to Carruthers (2013, 144n3): “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.”² Like many theory theorists, however, both Lavelle and Carruthers discount the phenomenology and place all the real action of social cognition in extra-perceptual processes at the subpersonal level where theory makes up for an impoverished perception (also see, e.g., Jacob 2011; Spaulding 2010). Carruthers, for example, adopts Leslie’s idea of a domain specific innate module that starts to function at around 6 months of age, and which provides infants “with the concepts and core knowledge necessary to represent the mental states of other agents” (2013, 142).³

Opposing this picture, philosophers, such as Wittgenstein,⁴ phenomenologists like Scheler (1954) and Merleau-Ponty,⁵ and philosophers of mind like Duddington (1918), Dretske (1973), and Cassam (2007), have, to different extents, argued against the principle of imperceptibility. In the contemporary landscape, interaction theory (IT) (De Jaegher et al. 2010; Gallagher 2001, 2004, 2005, 2012; Gallagher and Hutto 2008) defends the idea that social understanding is based primarily on embodied social interaction. On this account, in our everyday interactions with others, we are able to directly perceive their intentions and emotions; perception can grasp more than just surface behavior—or to put it precisely, it can grasp meaning—the

² We thank one of the journal referees for pointing out the statements made by Carruthers and Lavelle.

³ One finds stronger claims about the role of perception in the simulationists. Thus, noting the idea “that in some circumstances some mental states of others can be the objects of direct perception,” 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). Robert Gordon (2008, 221), too, makes room for direct perception: “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 ‘meanings of others’ facial expressions, gestures and intentional actions”. Within ST, however, there has been, so far as we know, no exploration of this issue which does not insist on the need for an extra step of simulation to supplement perception.

⁴ “Look into someone else’s face, and see the consciousness in it, and a particular shade of consciousness. You see on it, in it, joy, indifference, interest, excitement, torpor, and so on. ... Do you look into yourself in order to recognize the fury in *his* face?” (Wittgenstein 1967 §229) And again: “In general I do not surmise fear in him—I *see* it. I do not feel that I am deducing the probable existence of something inside from something outside; rather it is as if the human face were in a way translucent and that I were seeing it not in reflected light but rather in its own.” (Wittgenstein 1980, § 170).

⁵ “Anger, shame, hate, and love are not psychic facts hidden at the bottom of another’s consciousness: they are types of behavior or styles of conduct which are visible from the outside. They exist on this face or in those gestures, not hidden behind them.” (Merleau-Ponty 1964, 52–53).

intention in intentional behavior and the emotion in emotional expression. The view is that intentions and emotional states are in most cases not purely mental (“in-the-head”) episodes of which we acquire knowledge via inference from behavior. Rather, we perceive them directly in the bodily movement, expressions, and actions of others.

In this paper, our objectives are twofold. We first clarify how it is possible to directly perceive intentions and emotions. Second, we consider a possible objection to the direct perception hypothesis from social psychology that—to our knowledge—has not yet been considered in the social cognition literature.

In regard to the first objective, we address three questions: (1) What is an intention if it is something that can be perceived? (2) What is an emotion if it is something that can be perceived? And, (3) what is the nature of social perception if intentions and emotions can be perceived? The latter question is important since according to one sense of ‘perceiving another’s mental states’, X perceives Y’s mental states iff Y’s mental states figure in the content of X’s perceptual experience.⁶ One might then argue that neither TT nor ST should deny the phenomenological claim that we have perceptual experiences with such content. The question then becomes whether sub-personal processes that give rise to such perceptual experiences involve extra-perceptual processes. As we’ve already noted, TT and ST suggest that at the subpersonal level theory or simulation enter into the process. Accordingly, clarifying a sense in which perceptual processes at the subpersonal level do not involve either theoretical inference or simulation will be important. In order to elucidate what it is to directly perceive intentions and emotions, we borrow from action theory and emotion theory certain conceptions that are consistent with the idea that we can perceive such things directly, i.e., without having to infer or to simulate or to add any other cognitive process to perception. In this regard, we argue, mindreading, as construed by either TT or ST, is not our primary way of understanding others (although it may still be possible in special cases).

It is important to keep in mind that the account of direct perception outlined here is one part of a larger theory (IT), which emphasizes the role of embodied interaction and narrative, in addition to direct perception, in the constitution of social cognition. Direct perception plays an important role in IT, but an explanation of its role is not meant to be a full explanation of social cognition. Moreover, one can argue in favor of the direct perception thesis in several ways (Smith 2010; McNeill 2012; Krueger and Overgaard 2012) but endorsing it by no way commits one to IT.

Although social cognition does not reduce to direct perception, our intent is not to provide a full account of

social cognition and we will not rehearse all aspects of IT, or the research in developmental psychology, phenomenology and neuroscience that supports this approach (for good summaries of this research see Hobson 2002; Reddy 2008; Trevarthen 1979; Gallagher 2005). Rather, we keep the focus on direct perception.

In regard to the second objective, we consider a possible objection to the idea of direct perception from social psychology, related to empirical findings about phenomena like ‘dehumanization’ and ‘implicit’ racial bias. One might object, based on these findings, that social perception is not direct since it obviously depends on a body of cultural beliefs that operate along the lines of a culturally relative folk psychology. In addition to defending against this objection, we will suggest more generally that studies in social psychology help us to see that an adequate account of social cognition requires going beyond the relatively narrow realm that the current theorizing on this topic occupies. This may be surprising to many since the study of social cognition is held to be a prime example of interdisciplinary research, involving cognitive and developmental psychology, neuroscience and philosophy. However, we think there is something missing, namely insights from an understanding of social cognition that is embraced in social psychology and more generally in social philosophy. These insights involve in-group and out-group distinctions that manifest themselves in different ways, as ideological constructs, cultural narratives about otherness, class relations in societies, and so on. For purposes of this paper we focus only on the in-group/out-group distinction in general, without going deeper into the different ways in which it can be motivated. We claim that this is an issue that is currently neglected in the mainstream literature on social cognition that focuses on debates between TT, ST, or IT. To avoid misunderstandings, we should note that there is no shortage of research in social psychology on in-group and out-group relations (on general issues see Tajfel 1981; Tajfel et al. 1971; Leyens et al. 1994; Brewer and Silver 2000; Brewer et al. 1993. More specifically connected to our topic Haslam et al. 2005, 2007; Bain et al. 2009; Bastian and Haslam 2010; Fiske 2004, 1991; Likowski et al. 2008). Our point is that such results have not been taken into account in the debates between TT, ST and IT. Although in studies of social cognition these factors are considered to be little more than icing on the cake, we will suggest that they form an essential part of the recipe of social cognition.

2 Perceiving Intentions

We start with some familiar distinctions from action theory. Searle (1983), Bratman (1987), and Pacherie (2006,

⁶ We thank an anonymous reviewer for this formulation.

2008) make important distinctions among different types of intention. For our purposes, Pacherie's three-fold distinction will be the most convenient to consider.

- *D(istal) or F(uture) intention*: the product of a thoughtful deliberation—I prospectively form an intention to do something (e.g., Considering certain facts, I decide to buy a car next week).
- *P(roximate or Present) intention*: Searle's concept of intention-in-action—I specify my intention in the particular requirements of the action situation (e.g., I take a taxi to the car dealership and kick some tires).
- *M(otor) intention*: the intention intrinsic to the movements that make up my action. The M-intention specifies the precise movements necessary to accomplish the action (e.g., standing thus-and-so, or taking up a certain position to be able to kick the tire).

In the following, we will focus on P- and M-intentions, and we start by clarifying M-intentions. The notion of M-intention can be associated with the phenomenological discussion of “operative intentionality” (*fungierende Intentionalität*) in contrast to “act [mental-content] intentionality” (Merleau-Ponty 2002). Bodily or motor intentionality involves the aspects of meaningful motor behavior and expression that constitute what we call intentions. This idea attempts to capture the fact that the experiencing agent is intentionally engaged with the world through actions and projects that are not reducible to simple internal mental states, but involve an intentionality that is motoric and bodily. Actions have intentionality because they are directed at some goal or project, and this is something that we can see in the actions of others. As Merleau-Ponty indicates:

[O]perative intentionality is that which brings about the natural and prepredicative unity of the world and of our lives; it appears more clearly ... in our visual field than in objective knowledge (2002).

This account of M-intention is also consistent with pragmatist and neo-pragmatist views:

A founding idea of pragmatism is that the most fundamental kind of intentionality (in the sense of directedness towards objects) is the *practical* involvement with objects exhibited by a sentient creature dealing skillfully with its world. (Brandom 2008, 178).

Turning from phenomenology and pragmatism to science, there is good evidence that we can directly perceive M- and P-intentions since such intentions are actually present in the movements that we can see. Studies by Becchio et al. (2012) show that even in the absence of contextual information, intentions can be perceived in bodily movement. These

studies build on well-known work in kinematics showing that different action intentions specify different kinematic dynamics in movement (Ansuini et al. 2008; Marteniuk et al. 1987; Ansuini et al. 2006; Sartori et al. 2011a). The first point, then, is that intention shapes action kinematics (e.g., grasping): what you are going to do with an apple (eat it, offer it to someone, throw it) shows up in the dynamics of one's reach, and in variations in grasp. In this respect the intention is built into the movement of the action. Second, Becchio et al. show that perceivers are sensitive to these differences in kinematics and can see (with above 70 % accuracy) the intentions in these movements—they are able to discriminate between cooperative, competitive and individual-oriented actions (Sartori et al. 2011b). Furthermore, subjects are able to discriminate these differences even without specific contextual information—in the dark with point-lights on the wrist and fingers of the agent (Manera et al. 2011a, b, c; Becchio et al. 2012). Further evidence for the perception of intentions can be found in studies of adult bodily kinematics and the dynamics of social attention and interaction (Atkinson et al. 2007; Lindblom 2007).⁷

Someone might still object that these experiments do not rule out the idea that upon seeing what we see of the movement, we still must infer the intention. But this is to misunderstand the nature of the M-intention.⁸ It is not that the M-intention lies somewhere outside of the action movement or behind it, such that we need an inference to get to it; the action movements (the kinematic dynamics) constitute the M-intention. In many cases of intentional action there is no prior D-intention—no deliberative planning out; and there may be no or a minimal P-intention. For example, if I am sitting at my desk working hard to solve a philosophical problem, I may reach for my cup of coffee to take a drink, even as my attention remains on the problem. This is an intentional action, but I did not first form an intention to take a drink (although, of course, if asked, I could retrospectively formulate an intention or reason to explain my action). At best, the intention was formed in the movement itself, and there was nothing other than that M-intention—no intention other than the one you can see in the movement. Certainly in some cases there may be something like a P-intention formed along with the movement and we may call this an intention to have a

⁷ The study by Lindblom (2007), for example, can help to address a concern raised by one of the journal's referees. The experiments by Becchio and colleagues imposed an artificial constraint in asking subjects to identify one of three possible action intentions. The Lindblom study, however, involved field research rather than lab experiment, and the ecological situation that she studied did not involve artificial limitations in terms of a limited range of possible answers. The subjects studied were interacting with each other on a complex task in a work setting.

⁸ This also goes to the issue of what precisely the nature of perception is. We address this below in Sect. 4.

drink. But even in that case there is an M-intention and the intention to have a drink can be seen in the kinematic dynamics. So even if there is a P-intention formally distinguishable from the M-intention, one doesn't need to make an inference about it, since one can see the M-intention. To suggest that we need to go beyond what is just there in the movement in order to infer an intention located somewhere beyond the action, is to invent something (some hidden intention) that in some cases does not exist. In these kinds of intentional actions, the P-intention may be nothing over and above the M-intention—it's literally the intention-*in-action*.⁹ Furthermore, even in cases when the other person has formed a D-intention and we are attempting to grasp that D-intention, perhaps by processes that involve inference or simulation, we do so in many cases only by starting with the M-intentions that we take as (at least) expressing the D-intention. That is, it is only by directly perceiving the M-intention that we can even start to make an inference to the D-intention, if in fact that is necessary at all.

One question that frequently comes up in connection with this kind of direct perception view concerns deception. Surely we cannot know that another person intends to deceive us without some serious mindreading. But first, the possibility of erring about another's intention to deceive is nothing special (McNeill 2012). It is often the case that situations occur in which we err about basic features of objects, so the mere possibility of error, in no way undermines the possibility of direct perception. Second, it's important to note that if indeed another person's behavior motivates a suspicious mindreading about possible deception, it is likely something that we perceived in their behavior that was the motivating factor. Third, however, we may in fact recognize deception in the motor behavior of the other without having to take the further, mindreading, step. At the level of M-intentions and P-intentions, subjects are already able to detect attempts at deception. "If an actor pretends, say, that a suitcase he is carrying is heavier than it actually is, his movements will have a non-natural kinematics that can be detected by observers" (Pacherie 2005, p. 9 citing Runeson and Frykholm 1983).

⁹ John McDowell (2011) makes a similar point: "adapting Brian O'Shaughnessy's dual aspect conception of the will, I propose that when one intentionally engages in bodily action, the action's intentional character is an aspect of something that is also bodily through and through. The result stands in contrast with familiar philosophical pictures of the relation between mind and body" (p. 1). He proposes that "If rationality can be in bodily activity as opposed to behind it, we have a vivid contrast with a familiar picture according to which a person's mind occupies a more or less mysterious inner realm, concealed from the view of others. If physical activity can be rationality in action, as opposed to a mere result of exercises of rationality, we have a vivid contrast with the tendency to distance a person's body" (p. 17). For the notion of bodily rationality, see Gallagher 2013b and in press).

Subjects can discern whether activity is intended or not in staged social actions, even when watching point-light displays of the agents' movements (Good 1985). Importantly, these capabilities start to take shape in infancy. Seven- to nine-month-old infants perceive certain ambiguous acts like offering and withdrawing object as playful intentions—with different goals and outcomes than when the same intentions are interpreted literally (Legerstee 2005, p. 124; Reddy 1991, 2008).

Intentional bodily movements therefore have very distinctive properties; they are simultaneously constrained by the agent's goal, by the attributes of the situation and by a set of kinematic and biomechanical rules that jointly shape their dynamics (Pacherie 2005). The intentional aspects of bodily movements are not extrinsic to those movements—they are intrinsic and are reflected in the organization of movement. Intentional actions have observable characteristics that distinguish them from non-intentional behaviors. Intentional kinematics reflect, not only a distinctive dynamics contingent on the agent's goal, but also specific features of *the situation*. Thus, an important aspect of both P- and M-intentions concerns the fact that intentional actions are not carried out in thin air—they are always situated in physical and social environments. How I will carry out my D-intention to buy a car will depend on the various circumstances of who, what, when, where ... and specific environmental and bodily conditions that may facilitate or hinder my action. Intentions involve feedback-governed processes that extend into the world, and which exhibit, as Brandom puts it,

a complexity [that] cannot in principle be specified without reference to the changes in the world that are both produced by the system's responses and responded to [Such practices] are 'thick', in the sense of essentially involving objects, events, and worldly states of affairs. Bits of the world are *incorporated* in such practices" (Brandom 2008, 178)

According to IT and the direct perception hypothesis, social perception is enactive.¹⁰ That is, my perception of your action is already formed in terms of how I might respond to your action. I see your action, not as a fact that needs to be interpreted in terms of your mental states, but as a situated opportunity or affordance for my own action in response. The intentions that I can see in your movements appear to me as logically or semantically continuous with my own, or discontinuous, in support or in opposition to my task, as encouraging or discouraging, as having

¹⁰ IT adopts an enactivist view of perception. See, e.g., De Jaegher et al. (2010). It is not the case, however, that all enactivists or all proponents of direct perception adopt IT, or that all proponents of direct perception are enactivists.

potential for (further) interaction or as something I want to turn and walk away from. As Merleau-Ponty put it, my experience of movement is not as a meaningless mechanical event, but is a ‘praktognosia’ (2002, 162). My own perceptually informed bodily responses to the world or to another person are ways of encountering the other that cannot be reduced (or inflated) to a form of mindreading. The perceiver is enactively engaged in perceiving the intentions of others, in such a way that her own motor intentionality contributes to perception.

We know from research on mirror neurons (MNs) that they are activated for intended actions but not for unintended movements; activation depends on the action intentions of the perceived person, as well as pragmatic context (Fogassi et al. 2005; Iacoboni et al. 2005; Kaplan and Iacoboni 2006). In contrast to an internalist/simulationist interpretation of MN activation, the enactivist view conceives of MN activation not as subserving an act of mindreading, but as something that is intrinsic to the structure of perception—my perception being shaped by my own action possibilities—what I *can do* in response to the other. As we interact with others we can perceive their meanings and their (M- and P-) intentions in their bodily movements, gestures, facial expressions, in what they are looking at, and what they are doing in the rich pragmatic and social contexts of everyday life. Even if, to some degree, action movements by themselves are underdetermined, pragmatic and social contexts add specification. On the enactive view, one doesn’t need to go to the level of mental states (propositional attitudes, beliefs, desires, inside the head); rather, on both (or all) sides of social interaction, intentions are in the movement, in the action, in the environmentally attuned responses. In such contexts, we normally perceive another’s intentionality in terms of its appropriateness, its pragmatic and/or emotional value in the particular situation, or in terms of our own possible responses, rather than as reflecting inner mental states, or as constituting explanatory reasons for her further thoughts and actions.

Is this a form of behaviorism? No. The idea of “thick” behavior involves rejecting the view that takes “behavior to be just bodily movement and so strips it of intentionality, relocating all that is alive and intelligent in the hidden mind” (Leudar and Costall 2004, 603).¹¹ Movement, behavior, gesture, expression, and action are infused with

intentionality—not only because they are expressive of or specified by M- and P-intentions, which may reflect D-intentions, but also because they are situated in meaningful contexts. What is out there to be seen is more than thin behavior understood as a series of mere movements; rather we can perceive a rich mixture of physical and social contexts, intentions, and meanings.

3 Perceiving Emotions

We should be clear, the claim is not that we can directly perceive all or all kinds of mental states. We may see contextualized behavior that suggests that a person believes some particular fact, or is thinking in a certain way. But we do not claim that we can see his belief or his thought. In contrast, however, we do think that we can directly perceive some emotions. According to the thin ToMistic view, emotions are mental states that need to be inferred in the light of other mental states, e.g., beliefs and desires (Harris et al. 1989; Nguyen and Frye 1999; Wellman and Banerjee 1991). They may be expressed in bodily ways, but to perceive bodily expressions/behaviors is not to perceive the emotion itself. We require inferences to move from bodily expressions to an understanding of actual emotions. So on the direct perception view, how is it possible to perceive emotions?

The claim that we directly perceive emotions is not by way of a Jamesian move that might reduce emotions (at least in part) to observable bodily expressions; the claim is not that to perceive an emotion is reducible to the idea that I perceive the gestures of the other’s body *simpliciter*. Nor is it the idea that I perceive the visible expressions and *apperceive* the hidden sides of those expressions (see Joel Smith’s (2010) appeal to a Husserlian idea; see Krueger 2012 for critique). Rather, we want to say that emotions are often perceivable because of their embodied and complex nature. If we think of emotions as complex patterns of experiences and behaviors—and as such as “individuated in patterns of characteristic features” (Newen et al., under review)—features that may include bodily expressions, behaviors, action expressions, etc., then emotion perception can be considered a form of pattern recognition (Izard 1972; Izard et al. 2000; Newen et al., under review).

On this view, particular expressions and expressive actions may be constitutive features of a specific emotion but not necessary components of all instances of the emotion—in this respect we don’t always see a person’s emotion. Emotion, accordingly, is a cluster concept, characterized by a sufficient number of characteristic features, although none of them are necessary to every instance. What we do perceive when we perceive an emotion is a package (a gestalt, or what Green 2010 calls “an interrelated set of phenomena” or a “systematically related set of

¹¹ “If we stop thinking of behaviour as something that must be described in ‘thin’ terms, and recognize that it can also be described in ‘thick’ terms, then the illusion that the line between ‘the observable’ and ‘the unobservable’ is to be drawn along the line of thin descriptions will evaporate, and one will stop thinking that ‘the mental’ is unobservable, obscured from view by bodily movements and accessible only as a matter of inference” (Coulter and Sharrock 2009, 77).

components”) that includes a number of different constitutive aspects of the emotion pattern—not necessarily all aspects—but enough of its significant constituent features to count.¹²

Jacob (2011) has objected that the direct perception view leads to a crude behaviorism. The dilemma that Jacob puts forward is that if the direct perception hypothesis argues that bodily expressions are constitutive of emotional or cognitive states, if they can be identified with patterns of observable behavior then direct perception advocates must embrace an unattractive behaviorist position. This, however, does not follow. It is possible to maintain that some bodily actions are expressive of and partly constitute mental phenomena (in the sense that they actually make up their proper parts), without *reducing* psychological states to expressive behavior (Krueger and Overgaard 2012). The claim is simply that embodied mental states are only partly constituted by perceptible behaviors. As Green (2007) puts it, if we accept that we sometimes perceive objects by perceiving their parts, then it is also acceptable that we can perceive intentions and emotions although they entail other components that are not fully perceptible (for a differing view see McNeill 2012). Furthermore, the perceptual aspects of the complex pattern of an emotion are not reducible to purely bodily expressions. We also need to add (consistent with Dewey’s critique of James) a situational account—where the fact that emotional experiences and behaviors are situated in specific ways is part of the pattern (Mendoça 2012). This is to take seriously the phenomenological point that emotions involve intentionality, something that helps to disambiguate emotional expressions. Including situational aspects as part of the perceptual pattern of emotions also suggests that one can perceive complex, and not just basic emotions. Certain postures and gestures and the style of certain glances may be perceived as jealousy only when enough of the context is also perceived.¹³

¹² We note that this kind of emotion pattern perception fails in autism and schizophrenia where subjects have a propensity to view the face as an array of unrelated details; they miss the pattern/gestalt and fail to recognize the emotion. “While most people perceive the face or body of another as a familiar whole imbued with life, subjectivity, and expression, schizophrenia patients will sometimes focus on individual parts or the purely material aspect of the person before them” (Sass and Pienkos 2013).

¹³ This is just a sketch of the pattern theory of emotion (see Newen et al. under review, for a more detailed account). Goldman and Sripada (2005) provide a simulationist account of face-based emotion recognition, emphasizing that non-perceptual processes might be necessary “to elaborate purely perceptual information.” This need not contradict the idea that we directly perceive emotions since the latter does not deny that ‘elaborating’ on perceptual information can involve cognitive (non-perceptual) processes. The issue considered here, in any case, is not one of elaboration.

4 Social Perception

In much philosophy of mind, direct perception is suspect because it has traditionally been associated with the idea that it cannot be mistaken. If there is no representation that mediates perception, then we cannot account for error or illusion. Malcolm (1953), for example, considers G.E. Moore’s claim that one can have a direct perception of an after-image (in a way that one cannot have a direct perception of an environmental object) and ends up thinking that “‘impossibility of error’ is the main feature of the philosophical conception of direct perception.” But taking an after-image or a visual illusion as an example of something we can directly perceive, and concerning which we cannot make a mistake, is, we think, a mistake. Dealing with this issue would take us too far afield, however, so let’s put things differently.

By direct perception we mean perception that does not involve a certain kind of inference, but can still involve error. Perception, however, can involve inference in two ways: in either intra-perceptual processing or extra-perceptual processing. We argue against a view that would posit the latter type of processing and that would suggest:

1. We perceive (or sense) X, but X is meaningless unless we add something to perception, and.
2. What gets added to perception is an inference—a very fast inference (or some other cognitive process like simulation) to make sense of X.

Going back to the idea that we can directly perceive M- and P-intentions, and the evidence to this effect provided in the experiments by Becchio et al., one might object that from the claim that a perceiver understands that p is the case solely on the basis of perceptual *stimuli* (e.g., the M-intention in the action), it is not legitimate to conclude that the perceiver understands that p is the case based solely on perceptual *processes*, i.e., that the subpersonal processes are entirely perceptual. Rather, the objection might go, perception of an intention is underpinned by both perceptual and extra-perceptual subpersonal processes. On some versions of TT and ST something like an extra-perceptual inference (or simulation) is added to the perception because perception by itself is characterized as an impoverished form of observation, detached from action (or interaction). On this view there is a disconnection between my perception and anything that might involve my own action. On such models, if I were to remain with only what I literally perceive of your apparent behavior I would seemingly be in the dark, or totally perplexed, or at least puzzled by it.

In contrast to this view, the enactive approach argues that perception, without any extra-perceptual processes, can grasp more than just surface behavior—or to put it

precisely, it can grasp behavior as meaningful—it's a kind of smart perception (Gallagher 2008a, b). In the case of a not-so-smart social perception I open my eyes and I see a body moving in a meaningless way, flailing her arms for example, and I have to make sense of it in some non-perceptual way. My eyes are working fine; my visual cortex is processing all of the visual information, but what vision delivers is relatively meaningless, “thin” behavior, which I then have to interpret in some further cognitive steps that involve inference. In contrast, in the case of smart perception, in the very same situation, when I open my eyes I see a person engaged in an exercise routine at her gym. I do not see meaningless behavior and then infer that it is a form of exercise—and I don't have to call on inference unless she is doing something out of context or something that consists of weird or inappropriate movements.

My perception is obviously informed by my prior experience, so if I never encountered yoga before then I might start to wonder and to make inferences when I see the other person in a certain yoga position. Likely I may even have to ask someone what she is doing.

On the smart perception view, there is no denial that subpersonal processes in the brain contribute to perception. Even Gibson's notion of direct perception does not deny that subpersonal brain processes are involved in our ability to see affordances in the environment. In the case of smart social perception, the brain actively contributes—more precisely, the organism, including the brain, is engaged, and has something to contribute to the shaping of perception. Perception involves complex, dynamic processes at a subpersonal, sensory-motor level—but these processes are part of an enactive engagement or response of the whole organism, rather than additional, extra-perceptual, inferential or simulative processes.

For example, the fusiform “face area” of the brain is activated, not only for face perception, but also when we look at the front (grill, headlights) of cars (Gauthier et al. 2000; Xu 2005). The significance of this is that our neural processes are plastic and can be tuned by (social and cultural) experience. This activation (part of what constitutes the perception I have of my car if in fact I am looking at its front end) is not the underpinning of some additional inferential cognitive act. I do not perceive and then go through some other process that correlates to the activation of the fusiform face area; rather, fusiform activation helps to constitute the way that I perceive the car, or the other person's face, etc.¹⁴ Importantly, perception of another's

face activates not just the face recognition area and ventral stream, but the dorsal visual pathway—suggesting that we perceive affordances in the face of the other (Debruille et al. 2012). Faced with the face of a real person, the perceiving subject, at a minimum, makes eye contact with very subtle eye movements. Accordingly, face perception presents not just objective patterns that we might recognize as emotions. It involves complex interactive behavioural and response patterns arising out of an active engagement with the other's face—not a simple recognition of facial features—but an interactive perception that constitutes the recognition of emotions.

Meaningful perception of any sort may rely on activation of association brain areas outside of very early perceptual processing areas in, for example, V1 in the visual cortex. Recent research shows that even neuronal activity in the earliest of perceptual processing areas, such as V1, reflects more than simple feature detection. For example, V1 neurons are activated in ways that anticipate reward if they have been tuned by prior experience (Shuler and Bear 2006). This is not perception first, followed by an additional neural or cognitive function that registers the possibility of reward.

It's just here that followers of Helmholtz (1867) will argue that these subpersonal (intra-perceptual) processes that constitute perception are composed of inferences. And in terms of social perception, the theory theorist might be tempted to say that this subpersonal inference just is the theoretical inference that allows us to mindread (e.g., Lavelle 2012). But the theory theorist cannot claim that Helmholtzian inferences that are underpinning perception are the ones underpinning mindreading. First, Helmholtzian inferences (if there are such things) are characterized as very basic processes involving, for example, the visual perception of edges, colors, shapes, and so forth, and are meant to answer very basic questions about how we perceive anything as a visual object. It's not clear how such processes would be related to folk psychology.

Footnote 14 continued

others, as if intentions and emotions were reducible to simple sensory properties like the redness of blushing. Perceiving my red car (not its sensory properties) and perceiving another person's intentions or emotions may be similar insofar as both perceptions are direct and enactive, and in that sense “smart.” That there are important differences, both neurologically and phenomenologically, between object perception and person perception, even in regard to what ‘enactive’ means, has been explicated (even in Gallagher 2008a, but also in Gallagher 2005, 2008b; and Gallagher and Zahavi 2008), for example, in terms of the differences between instrumental and social affordances, between mere observation and engaged perception, and in terms of the dynamics of interaction. Bohl and Gangopadhyay are after a more detailed explanation of “the cognitive-perceptual mechanisms” of direct social perception and are clearly not satisfied with an embodied, dynamical approach that includes an enactive interpretation of the mirror system.

¹⁴ In discussing the concept of direct perception, Gallagher (2008a) made use of the example of perceiving his red car. This has misled some critics (e.g., Bohl and Gangopadhyay 2013) into thinking that he was suggesting there is no difference between perceiving the sensory properties of things and perceiving the intentions or emotions of

Helmholtzian inferences, at least in the classic sense, are not rich enough to underpin mindreading. Second, if TT did make this claim, it would be tantamount to the claim that mindreading just is the *perception* of mental states. The Helmholtzian idea that perception involves subpersonal inferences may or may not be correct (see Bennett and Hacker 2008, 9–10; Orlandi 2011; Hutto and Myin 2013). Even if the theory theorist thinks that it is correct, however, strictly Helmholtzian inferences will not give TT what it needs for social cognition, since according to TT such inferences would, at best, deliver only a perception of behavior—e.g., I see the agent reaching for the cup. One would need to add to perception some other kind of subpersonal, extra-perceptual, extra-Helmholtzian mindreading inferences. That would bring us back to not-so-smart perception plus some other cognitive process.

Alternatively, the theory theorist might try to make the extra-perceptual inferential process part of perception itself. This is precisely Jane Lavelle's (2012) proposal in her TT critique of direct perception. Objecting to the claim that MNs are able to register the action intention or goal of a perceived action she proposes an inferential process in a classic syllogistic set of brain processes that are to be integrated, somehow, into perceptual processes.

Premise 1 is generated in the representation of a motor action in the parietal mirror area.

Premise 2 is equivalent to “knowledge” representations about cultural practices or folk psychology formulated [propositionally] in some other brain module.

Conclusion: the brain computes across these representations to infer the best explanation in terms of the other's mental states.

Cultural knowledge or folk psychological platitudes need to be provided by some mechanism (a ToMM, for example) other than the primary activation of perceptual and mirror areas. Accordingly, Lavelle rejects Gallese's idea that low-level processes are sufficient for understanding actions.

[W]e don't need to suppose an over-arching top-down influence in order to have a neural mechanism that maps the goal. We already have it in the pre-motor [or parietal] system. We don't need to imply a further mechanism that maps the goal. (Gallese 2006, p. 15)

Lavelle's rejection of this proposal ignores the possibility that sensory-motor areas have undergone plastic modifications in prior experience. This kind of discounting of brain plasticity is a retreat to a standard computational model of the mind (see Fodor 1983; Strawson 1994, and the critique of hyperintellectualist models in Hutto and Myin 2013). On

this view, perception by itself is impoverished, and meaning would be added, top-down, piled onto the perceptual vehicle forming a new representation.

We question whether it is best to think of social and cultural factors in terms of theory laden perception, as if the way our experience is (in)formed by social and cultural factors translates into the possession of a theory (folk psychology) that needs to be added to perception to formulate an extra inferential step in understanding others. Rather than adding extra-perceptual inferential processing (generated in a ToMM, or a folk-psychological module, for example) to perception, there is good evidence that perceptual processes at the subpersonal level are already shaped, via mechanisms of plasticity, by bodily (enactive) and environmental (including social and cultural) factors and prior experience. For example, consider the now well-known difference between the way Westerners and Asians perceive and attend to visual objects and contexts (Goh and Park 2009). One also finds, for example, not only brain processes that are different relative to the use of different cultural tools and practices, but also cultural variations in brain mechanisms specifically underlying person perception and emotion regulation (Kitayama and Park 2010). For example, relative to European Americans, Asians show different neural processing in response to images of faces that represent a social-evaluative threat (Park and Kitayama 2012). In very specific ways, social and cultural factors have a physical effect on brain processes that shape basic perceptual experience and emotional responses.

To summarize, M- and P-intentions are not hidden, purely mental events; they are visible in situated, embodied actions. We can perceive M- and P-intentions without the need for extra-perceptual cognitive inferences. Emotions are constituted by patterns of bodily-experiential-expressive aspects, some significant parts of which can be perceived and understood without the need for extra-perceptual cognitive inferences. In most cases, the subpersonal processes of perception that contribute to understanding others, even if they involve intra-perceptual Helmholtzian inferences, do not require extra-perceptual cognitive inferences to do the job. Finally, in the larger system of brain-body-environment, brain plasticity plays an important role in building social and cultural factors into the way perception works.

5 Some Concerns from Social Psychology

Although we have argued that in many if not most cases we directly perceive intentions and emotions in others, because they are visible (or audible) in situated, embodied actions, evidence from social psychology may seem to put the idea of direct perception into question. It is well known, for

example, that individuals are more accurate at recognizing the intentions and the emotions of members of their own culture versus those of other cultures (Elfenbein and Ambady 2002a, b; Matsumoto 2002). In itself, this phenomenon does not constitute a challenge for the view of direct perception proposed here. We have said that emotions are best thought of as complex patterns of experiences and behaviors and that emotion perception can be considered a form of pattern recognition. In that case, it makes sense that the cultural differences between these patterns might make it harder for individuals to recognize the emotions of individuals from other cultures. There are subtle differences in emotional ‘dialects’ across cultures, which reduce cross-cultural emotion recognition (Elfenbein et al. 2007). Research also shows that the in-group advantage in emotion recognition is largely independent of biological or ethnic factors. It seems that individuals make best sense of emotions expressed by a member’s own cultural group, regardless of race and ethnicity (Elfenbein and Ambady 2003). Something similar might be said about intentions insofar as there are culturally typical ways of doing things, and culturally typical things to do.

However, research also shows that independent of ‘dialects’, *beliefs*, and most strikingly negative beliefs about out-group members can interfere with one’s ability to recognize emotions (Gutsell and Inzlicht 2010). In these cases making sense of the emotions of others is not constrained by the differences in emotion patterns, but by specific beliefs about the outgroup member. At least, it seems that whether X is able to recognize the emotions and intentions of Y, is crucially dependent on X’s beliefs about the racial or ethnic group to which Y belongs. On some conceptions it is not just a matter of ‘having a belief’ but of having a set of beliefs or a set of platitudes about the out group that constitute part of folk psychology, or, in effect, a theory. On this view, the kind of subpersonal syllogism suggested by Lavelle’s (2012) seems more feasible.

The phenomenon of *dehumanization* shows that being able to experience the other as a human being, and to grasp her intentions and emotions, are to a large degree contingent socio-cultural factors. Dehumanization, a phenomenon often found in war and genocide contexts, refers to processes in which individuals or groups are simply understood as somehow lacking full humanity. Others are understood as lacking characteristics that in-group members take to be characteristically human (a sense of morality, civility, higher cognitive abilities, emotional warmth etc.). In-group members occasionally perceive people of a certain ethnicity as animal-like (animalistic dehumanization) or as automata (mechanistic dehumanization). In extreme cases, such out-group members are met with disgust and perceived as somehow non-humans or sub-humans, as beings *without an inner life* (Harris and Fiske 2006; Haslam 2006). It seems

that in such cases perception completely fails to grasp the other, and basic empathy, the grasp of another as a fellow human, is missing. In situations of extreme conflict this helps overcome revulsion against killing; but moderate versions of this phenomenon are present in subtle everyday processes (Haslam et al. 2005, 2008a and b; Haslam 2006; Haslam and Bain 2007; Bain et al. 2009; Bastian and Haslam 2010; Fiske 2004, 1991; Goffman 1986). Thus, dehumanization is a matter of degree, not of kind.

Again, this seems to present a challenge to the direct perception hypothesis. Indeed, dehumanization is manifest on the level of bodily interaction usually connected with ‘primary intersubjectivity’: non-conscious processes of automatic mimicry of others’ expressions, gestures, and body postures are less frequent for dehumanized out-group members (Likowski et al. 2008). Also, what are considered innate motor-resonance mechanisms that supposedly allow us to directly perceive intentions and emotions are modulated by cultural factors and inextricably bound to group membership. A study by Xu et al. (2009) dramatically demonstrates the neural effects of implicit racial bias and shows that empathic neural responses to the other person’s pain are modulated by the racial in-group/out-group relationship. fMRI brain imaging showed significant decreased activation in the anterior cingulate cortex (ACC), an area thought to correlate with empathic response, when subjects (Caucasians or Chinese) viewed racial out-group members (Chinese or Caucasian respectively) undergoing painful stimulations (needle penetration) to the face, compared to ACC activation when they viewed the same stimulations applied to racial in-group members.

We are simply less responsive to out-group members and display significantly less motor cortex activity when observing out-group members (Molnar-Szakacs et al. 2007). Most strikingly, in-group members fail to understand out-group member actions, and this is particularly prominent for disliked and dehumanized out-groups. The more dehumanized the out-group is, the less intuitive the grasp of out-group member intentions and actions (Gutsell and Inzlicht 2010).

The evidence from studies of dehumanization and implicit racial bias thus seems inconsistent with the direct perception hypothesis, and shows that mechanisms of basic empathy are constitutively dependent upon historical-cultural situatedness and group membership. While the argument for direct perception draws on empirical findings concerning primary intersubjectivity and enactive interpretations of resonance processes, it seems that social psychology and cultural neuroscience raise questions about exactly such phenomena. Recent studies on embodied primary intersubjectivity and mirror neuron activity deliver evidence for our basic understanding of others being constitutively dependent on culturally sanctioned beliefs. In

light of these findings, we may ask: should such cultural beliefs that enable and disable social cognition not be seen as a form of theory? And, if the recognition of emotions and intentions depends on such a theory would this then not contradict direct perception and support theory-theory?

Our answer is no. First of all, recall that the idea that we have direct perceptual awareness of the other's intentions and emotions is part of a larger *interaction theory* of social cognition (IT) which draws on evidence that our basic empathic understanding of others is enabled by innate or very early developing embodied capabilities and by interaction itself (see De Jaegher et al. 2010). The term 'innate' here signifies those capabilities that have developed prenatally as a combination of genetic and prenatal experiential factors. The newborn comes already prepared for interaction with others, as evidence on neonate imitation and primary intersubjectivity suggests (Meltzoff and Moore 1977, 1994; Trevarthen and Aitken 2001; including cross-cultural studies, Meltzoff and Moore 1989). To disrupt a common metaphor, however, this does not mean that the infant comes "hardwired." Rather, it means that the newborn infant has some circuits already working, but even these circuits are open to plastic reorganization; they are either reinforced or they deteriorate depending on subsequent experience¹⁵; generally speaking, they are reshaped by social and cultural experiences.

Although interaction theorists, in their critique of TT and ST, focus on primary embodied processes, they also grant that social and cultural contexts are important for a full understanding of the other. IT maintains that we are not only action oriented in our pragmatic dealings with the world, we are also, from the very beginning, interaction oriented in our encounters with others. Thus, beyond the embodied capacities of primary intersubjectivity, IT has acknowledged the importance of secondary intersubjectivity (starting with joint attention in the first year of life, and including the pragmatic understanding of others in highly contextualized situations) and of communicative and narrative practices (Gallagher and Hutto 2008). The stories that we listen to as children, or that we see enacted (in various media), or play-acted, and even the stories we are exposed to as adults—parables, plays, myths, novels, films, television, etc.¹⁶—are not neutral with respect to

how we perceive the world. Cultural narratives, as well as our own culturally situated experiences with others, bias our expectations in regard to their actions and, as the science shows, can bias perception itself. While it was once thought that such biases were automatic and more or less immune from change, it is now accepted that the manipulation of the social context can moderate in-group racial bias, down to the level of perceptual processes (Barden et al. 2004; Blair 2002; Bargh 1999). Thus, IT can and does acknowledge that social and cultural forces play an essential or constituting role in social perception and particularly in the understanding of emotions and intentions.

Moreover, this kind of evidence puts into question accounts of social cognition that assume we are hardwired to intuitively grasp others as "fellow human beings" by means of innate, modularistic ToM mechanisms or pre-programmed mirror systems operating in an automatic and context-independent fashion, yielding capacities of the sort described by Scholl and Leslie (1999, 136–137).

One hallmark of the development of a modular cognitive capacity is that the end-state of the capacity is often strikingly uniform across individuals. Although the particulars of environmental interaction may affect the precise time-table with which the modular capacity manifests itself, what is eventually manifested is largely identical for all individuals. As the modular account thus predicts, the acquisition of ToM is largely uniform across both individuals and cultures. The essential character of ToM a person develops does not seem to depend on the character of their environment at all. It is at least plausible, *prima facie*, that we all have the same basic ToM! (...) The point is that the development of beliefs about beliefs seems remarkably uniform and stable.¹⁷

Others like Segal (1996) maintain that the pattern of ToM development is identical across the species, which is in marked contrast to the uneven and culturally dependent development of many other capacities. Evidence from

Footnote 16 continued

changes involve brain plasticity; they are not induced because players change their theories or improve their inferences.

¹⁷ Scholl and Leslie (1999, 140) add that "It is certainly the case that these basic ToM abilities may eventually be recruited by higher cognitive processes for more complex tasks, and the resulting higher-order ToM activities may well interact (in a non-modular way) with other cognitive processes, and may not be uniform across individuals or cultures." On this view, higher-order cognitive elaboration does not affect perceptual or basic ToM processes. On the enactive view proposed here, cultural practices, including communicative and narrative practices, go deeper and affect embodied perceptual processes and action possibilities. The material we draw on suggests that cultural factors can also lead to the loss or diminishment of the ability to understand others.

¹⁵ Evidence along this line can be found in what is sometimes considered induced autism as the result of extreme social deprivation in orphanages (see Hobson 2002).

¹⁶ Rorty (1989, xvi) suggests "the novel, the movie and TV program have, gradually but steadily, replaced the sermon and the treatise as the principle vehicles of moral change and progress" (1989, xvi). We should add video games to this list. As research shows, prolonged play of such games induces changes in various cognitive functions, including hand-eye coordination, spatial visualization, visual anticipation, reaction time, and attention (Latham et al. 2013). Such

studies of dehumanization, however, is inconsistent with these expectations, and shows that mechanisms of social cognition are constitutively dependent upon historical-cultural situatedness and group membership. This suggests that the fundamental perceptual level of understanding others as persons is essentially context dependent—an aspect that any theory of social cognition must account for.

To deny that cultural factors have such effects on perception would only make sense if one were to accept the thesis of the ‘cognitive impenetrability of perception’ (Pylyshyn 1999) and hold on to the distinction between ‘seeing’ and ‘seeing as’. However, many now acknowledge that perception is cognitively penetrable (Siegel 2011). The frequent example in discussions of cognitive penetrability involves beliefs. When you know that bananas are yellow, this knowledge affects what color you see bananas to be, so that an achromatic banana will appear to be yellow (Gegenfurtner et al. 2006). This leads too quickly to the idea that perceptions are “theory laden,” a concept borrowed loosely from philosophy of science. But moods, traits, practices and skills also can modulate perception. For example, to the newly trained reader of Russian, a sheet of cyrillic script looks different than it looked to her before she could read it; to a vain performer, the faces in the audience never look disapproving, while to a performer who lacks confidence, the same audience may look displeased (Siegel 2011). In a kind of circular way, and as Siegel points out, in a way that can be epistemically pernicious, penetrated perceptions are confirmatory of the belief, mood, trait, etc. In the case of cultural biases, they can also be neurologically pernicious since they can reinforce neuronal firing patterns and result in the plastic changes discussed above. More generally, they can reinforce embodied practices and postures, behavioral habits, and intersubjective interactions. None of this, however, counts against the idea that my perception of another’s intentions and emotions are direct, requiring no extra-perceptual inference that would take us beyond what we perceive. All such changes, pernicious or not, are not additions to perception, an added-on set of inferences; rather, they transform the perceptual process itself. In the case of dehumanization, for example, one is not trained to make bad inferences; one is conditioned to directly perceive others as non-persons.

6 Conclusion

We started by noting that the problem of social cognition is usually framed as a problem of access to the other person’s mind and that it is usually supposed that the other person’s mental states are not accessible to perception. Buying into the ‘principle of imperceptibility’ it is assumed that we have access only to behavior or action, which is not

meaningful unless supplemented by some kind of an additional inferential operation. In this framework the central discussion is concerned with the nature of the standards that guide our inferences, and whether they come from a folk-psychological theory (TT), or from our own experiential resources (ST). In opposition, IT denies the ‘principle of imperceptibility’ and argues that perception can directly grasp meaning in intentional behavior and emotional expression.

In this paper, we drew from action theory and emotion theory to elucidate how it is possible to directly perceive intentions and emotions. We then considered a possible objection that could be launched against the direct perception hypothesis by drawing on social psychology studies concerning ‘dehumanization’ and ‘implicit’ racial bias. Such findings could be interpreted to prove that perception is not direct, but rather depends on cultural beliefs that might be seen as a form of theory. We argued, however, that far from threatening the idea of direct perception within IT, these findings seem to clearly contradict the idea of innate or hard-wired ToM modules.

In conclusion, we want to suggest that the science of social cognition needs to take into account the role of ideological constructs, cultural narratives about otherness, phenomena concerning in-group and out-group dynamics, and, we would add, class and power relations in societies. These are topics that are currently neglected in the mainstream social cognition literature found in philosophy of mind, cognitive psychology, and neuroscience. Too often, in this mainstream literature, social cognition is portrayed as dependent on internal mechanisms that belong to a neutral observer of another person’s behaviour, *simpliciter*, without taking into consideration that social interaction processes are shaped by forces external to the individual, and by social and institutional practices that impact intersubjective understanding to the extent that they form and sometimes deform perception (Gallagher 2013a), as well as any further cognitive processes involved in our understanding of others.

Acknowledgments S.G.’s research for this paper was supported in part by the Marie Curie Actions ITN project 264828, *Towards an embodied science of intersubjectivity* (TESIS), and by the Humboldt Foundation’s Anneliese Maier Research Fellowship.

References

- Ansuini C, Santello M, Massaccesi S, Castiello U (2006) Effects of end-goal on hand shaping. *J Neurophysiol* 95:2456–2465
- Ansuini C, Giosa L, Turella L, Altoè GM, Castiello U (2008) An object for an action, the same object for other actions: effects on hand shaping. *Exp Brain Res* 185:111–119
- Atkinson AP, Tunstall ML, Dittrich WH (2007) Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition* 104:59–72

- Bain P, Park J, Kwok C, Haslam N (2009) Attributing Human Uniqueness and Human Nature to Cultural Groups: distinct forms of subtle dehumanization. *Group Process Intergroup Relat* 12:789
- Barden J, Maddux WW, Petty RE, Brewer MB (2004) Contextual moderation of racial bias: the impact of social roles on controlled and automatically activated attitudes. *J Pers Soc Psychol* 87:5
- Bargh JA (1999) The cognitive monster: the case against the controllability of automatic stereotype effects. In: Chaiken S, Trope Y (eds) *Dual-process theories in social psychology*. Guilford Press, New York, pp 361–382
- Baron-Cohen S, Swettenham J (1996) The relationship between SAM and ToMM: two hypotheses. In: Carruthers P, Smith PK (eds) *Theories of theories of mind*. Cambridge University Press, Cambridge, pp 158–168
- Bastian B, Haslam N (2010) Excluded from humanity: the dehumanizing effects of social ostracism. *J Exp Soc Psychol* 46:107–113
- Becchio C, Manera V, Sartori L, Cavallo A, Castiello U (2012) Grasping intentions: from thought experiments to empirical evidence. *Front Hum Neurosci* 6:117
- Bennett MR, Hacker PMS (2008) *History of cognitive neuroscience*. Wiley-Blackwell, Oxford
- Blair IV (2002) The malleability of automatic stereotypes and prejudice. *Pers Soc Psychol Rev* 6:242–261
- Brandom R (2008) *Between saying and doing: towards an analytic pragmatism*. Oxford University Press, Oxford
- Bohl V, Gangopadhyay N (2013) Theory of mind and the unobservability of other minds. *Phil Explor*. doi:10.1080/13869795.2013.821515
- Bratman M (1987) *Intention, plans, and practical reason*. Harvard University Press, Cambridge, MA
- Brewer MB, Silver MD (2000) Group distinctiveness, social identification, and collective mobilization. In: Stryker S, Owens TJ, White RW (eds) *Self, identity, and social movements*. University of Minnesota Press, Minneapolis, MN, pp 153–171
- Brewer MB, Manzi JM, Shaw JS (1993) In-group identification as a function of depersonalization, distinctiveness, and status. *Psychol Sci* 4(2):88–92
- Carruthers P (2013) Mindreading in infancy. *Mind and Language* 28(2):141–172
- Cassam K (2007) *The possibility of knowledge*. Clarendon Press, Oxford
- Coulter J, Sharrock W (2009) Theory of mind. In: Leudar I, Costall A (eds) *Against theory of mind*. Palgrave Macmillan, Basingstoke, pp 56–88
- De Jaeger H, Di Paolo E, Gallagher S (2010) Can social interaction constitute social cognition? *Trends Cogn Sci* 14:441–447
- Debruille JB, Brodeur MB et al (2012) N300 and social affordances: a study with a real person and a dummy as stimuli. *PLoS One* 7(10):e47922
- Dretske F (1973) Perception and other minds. *Nous* 7:34–44
- Duddington NA (1918) Our knowledge of other minds. *Proc Aristot Soc* 19:147–178
- Elfenbein HA, Ambady N (2002a) Is there an in-group advantage in emotion recognition? *Psychol Bull* 128:243–249
- Elfenbein HA, Ambady N (2002b) On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychol Bull* 128:203–235
- Elfenbein HA, Ambady N (2003) When familiarity breeds accuracy: cultural exposure and facial emotion recognition. *J Pers Soc Psychol* 85:276–290
- Elfenbein HA, Beaupré M, Le'vesque M, Hess U (2007) Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions. *Emotion* 7:131–146
- Fiske AP (1991) *Structures of social life: the four elementary forms of human relations*. Free Press, New York
- Fiske AP (2004) Four modes of constituting relationships: Consubstantial assimilation; space, magnitude, time, and force; concrete procedures; abstract symbolism. In: Haslam N (ed) *Relational models theory: a contemporary overview*. Lawrence Erlbaum Associates, Inc, Mahwah, NJ, pp 61–146
- Fodor Jerry A (1983) *Modularity of mind: an essay on faculty psychology*. MIT Press, Cambridge, MA
- Fogassi L, Ferrari PF, Gesierich B, Rozzi S, Chersi F, Rizzolatti G (2005) Parietal lobe: from action organization to intention understanding. *Science* 308:662–667
- Frith CD, Frith U (2007) Social cognition in humans. *Curr Biol* 17:R724–R732
- Gallagher S (2001) The practice of mind: theory, simulation, or interaction? *J Conscious Stud* 8(5–7):83–107
- Gallagher S (2004) Understanding interpersonal problems in autism: interaction theory as an alternative to theory of mind. *Philos Psychiatry Psychol* 11(3):199–217
- Gallagher S (2005) *How the body shapes the mind*. Oxford University Press, Oxford
- Gallagher S (2008a) Direct perception in the intersubjective context. *Conscious Cogn* 17:535–543
- Gallagher S (2008b) Intersubjectivity in perception. *Cont Philos Rev* 41(2):163–178
- Gallagher S (2012) In defense of phenomenological approaches to social cognition: interacting with the critics. *Rev Philos Psychol* 3(2):187–212
- Gallagher S (2013a) The socially extended mind. *Cogn Syst Res* 25–26:4–12
- Gallagher S (2013b) Enactive hands. In: Radman Z (ed) *The hand: an organ of the mind*. MIT Press, Cambridge, MA, pp 209–225
- Gallagher S (in press) To follow a rule: lessons from baby logic. In: Weinstock D, Levy J (eds) *Essays in honor of Charles Taylor*. McGill-Queen's University Press, Montreal
- Gallagher S, Hutto D (2008) Primary interaction and narrative practice. In: Zlatev J, Racine T, Sinha C, Itkonen E (eds) *The shared mind: perspectives on intersubjectivity*. John Benjamins, Amsterdam, pp 17–38
- Gallagher S, Zahavi D (2008) *The phenomenological mind*. Routledge, London
- Gallese V (2006) Intentional attunement: a neurophysiological perspective on social cognition and its disruption in autism. *Brain Res. Cog. Brain Res.* 1079:15–24
- Gauthier I, Skudlarski P, Gore JC, Anderson AW (2000) Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* 3:191–197
- Gegenfurtner K, Olkkonen M, Walter S (2006) Memory modulates color experience. *Nat Neurosci* 9(11):1367–1368
- Goffman E (1986) *Stigma: notes on the management of spoiled identity*. Simon & Schuster, New York
- Goh JO, Park DC (2009) Culture sculpts the perceptual brain. *Prog Brain Res* 178:95–111
- Goldman A (2006) *Simulating minds: the philosophy, psychology and neuroscience of mindreading*. Oxford University Press, Oxford
- Goldman AI, Sripada CS (2005) Simulationist models of face-based emotion recognition. *Cognition* 94(2005):193–213
- Good JMM (1985) The perception of social actions from point light displays: an exploratory study. Presented at the third international conference on event perception and action, Trieste, Italy
- Gordon R (2008) Beyond mindreading. *Phil Explor* 11:219–222
- Green M (2007) *Self-expression*. Oxford University Press, Oxford
- Green M (2010) Perceiving emotions. *Aristot Soc Suppl Vol* 84:45–61
- Gutsell JN, Inzlicht M (2010) Empathy constrained: prejudice predicts reduced mental simulation of actions during observation of outgroups. *J Exp Soc Psychol* 46:841–845

- Harris LT, Fiske ST (2006) Dehumanizing the lowest of the low: neuroimaging responses to extreme outgroups. *Psychol Sci* 17:847–853
- Harris PL, Johnson CN, Hutton D, Andrews G, Cooke T (1989) Young children's theory of mind and emotion. *Cogn Emot* 3:379–400
- Haslam N (2006) Dehumanization: an integrative review. *Pers Soc Psychol Rev* 10:252–264
- Haslam N, Bain P (2007) Humanizing the self: moderators of the attribution of lesser humanness to others. *Pers Soc Psychol Bull* 33:57–68
- Haslam N, Bain P, Douge L, Lee M, Bastian B (2005) More human than you: attributing humanness to self and others. *J Pers Soc Psychol* 89:937–950
- Haslam N, Kashima Y, Loughnan S, Shi J, Suitner C (2008a) Subhuman, inhuman, and superhuman: contrasting humans with nonhumans in three cultures. *Soc Cogn* 26:248–258
- Haslam N, Loughnan S, Kashima Y, Bain P (2008b) Attributing and denying humanness to others. *Eur Rev Soc Psychol* 19:55–85
- Heal J (1995) How to think about thinking. In: Davies M, Stone T (eds) *Mental simulation*. Oxford, Blackwell, pp 33–52
- Helmholtz H (1867) *Handbuch der physiologischen Optik*. Leopold Voss, Leipzig
- Hobson P (2002) *The cradle of thought*. Macmillan, London
- Hutto D, Myin E (2013) *Radicalizing enactivism: basic minds without content*. MIT Press, Cambridge, MA
- Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, Rizzolatti G (2005) Grasping the intentions of others with one's own mirror neuron system. *PLoS Biol* 3(3):529–535
- Izard CE (1972) *Patterns of emotions: a new analysis of anxiety and depression*. Academic Press, New York
- Izard CE, Ackerman BP, Schoff KM, Fine SE (2000) Self-organization of discrete emotions, emotion patterns, and emotion cognition relations. In: Lewis MD, Granic I (eds) *Emotion, development, and self-organization*. Cambridge University Press, Cambridge, pp 15–36
- Jacob P (2011) The direct perception model of empathy: a critique. *Rev Philos Psychol* 2:519–540
- Johnson SC (2000) The recognition of mentalistic agents in infancy. *Trends Cogn Sci* 4:22–28
- Kaplan JT, Iacoboni M (2006) Getting a grip on other minds: mirror neurons, intention understanding, and cognitive empathy. *Soc Neurosci* 1(3–4):175–183
- Kitayama S, Park J (2010) Cultural neuroscience of the self: understanding the social grounding of the brain. *Soc Cogn Affect Neurosci* 5(2–3):111–129
- Krueger J (2012) Seeing mind in action. *Phenomenol Cognitive Sci* 11(2):149–173
- Krueger J, Overgaard S (2012) Seeing subjectivity: defending a perceptual account of other minds. *ProtoSoc: Conscious Subj* 47:239–262
- Latham AJ, Patston LL, Tippett LJ (2013) The virtual brain: 30 years of video-game play and cognitive abilities. *Front Psychol* 4(629). doi:10.3389/fpsyg.2013.00629
- Lavelle JS (2012) Theory-theory and the direct perception of mental states. *Rev Philos Psychol* 3(2):213–230
- Legerstee M (2005) *Infants' Sense of People: Precursors to a Theory of Mind*. Cambridge University Press, Cambridge
- Leudar I, Costall A (2004) On the persistence of the 'problem of other minds' in psychology: chomsky, grice and 'theory of mind'. *Theory Psychol* 14:601–621
- Leyens Jacques-Philippe, Yzerbyt Vincent, Schadron Georges (1994) *Stereotypes and social cognition*. Sage Publications, London
- Likowski KU, Mühlberger A, Seibt B, Pauli P, Weyers P (2008) Modulation of facial mimicry by attitudes. *J Exp Soc Psychol* 44:1065–1072
- Lindblom J (2007) *Minding the body: interacting socially through embodied action*. Linköping: linköping studies in science and technology, Dissertation No. 1112, Department of Computer and Information Science, Linköpings Universitet, Sweden
- Malcolm N (1953) Direct perception. *The Philosophical Quarterly* 3(13):301–316
- Manera V, Becchio C, Cavallo A, Sartori L, Castiello U (2011a) Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Exp Brain Res* 211:547–556
- Manera V, Becchio C, Schouten B, Bara BG, Verfaillie K (2011b) Communicative interactions improve visual detection of biological motion. *PLoS One* 6:e145–e194
- Manera V, Del Giudice M, Bara BG, Verfaillie K, Becchio C (2011c) The second-agent effect: communicative gestures increase the likelihood of perceiving a second agent. *PLoS One* 6:e226–e250
- Marteniuk RG, MacKenzie CL, Jeannerod M, Athenes S, Dugas C (1987) Constraints on human arm movement trajectories. *Can J Psychol* 41:365–378
- Matsumoto D (2002) Methodological requirements to test a possible in-group advantage in judging emotions across cultures. *Psychol Bull* 128:236–242
- McDowell J (2011) Some remarks on intention in action. *Amherst Lect Philos* 6:1–18
- McNeill W (2012) Embodiment and the perceptual hypothesis. *Philos Quart* 62(248):569–591
- Meltzoff AN, Moore MK (1989) Imitation in newborn infants: exploring the range of gestures imitated and the underlying mechanisms. *Dev Psychol* 25(6):954–962
- Meltzoff A, Moore MK (1977) Imitation of facial and manual gestures by human neonates. *Science* 198:75–78
- Meltzoff A, Moore MK (1994) Imitation, memory, and the representation of persons. *Infant Behav Dev* 17:83–99
- Mendoça D (2012) Pattern of sentiment: following a Deweyan suggestion. *Trans Charles S Peirce Soc* 48(2):209–227
- Merleau-Ponty M (1964) *Sense and non-sense*. Northwestern University Press, Evanston
- Merleau-Ponty M (2002) *The phenomenology of perception*. Routledge, London
- Molnar-Szakacs I, Wu AD, Robles FJ, Iacoboni M (2007) Do you see what I mean? Corticospinal excitability during observation of culture-specific gestures. *PLoS One* 2:e626
- Newen A, Welpinghus A, Jukel G (under review) Emotion recognition as pattern recognition: the relevance of perception
- Nguyen L, Frye D (1999) Children's theory of mind: understanding of desire, belief and emotion with social referents. *Soc Dev* 8(1):70–92
- Onishi K, Baillargeon R (2005) Do 15-month-olds understand false beliefs? *Science* 308:255–258
- Orlandi N (2011) *Embedded seeing: Vision in the natural world*. Noûs. doi:10.1111/j.1468-0068.2011.00845.x
- Pacherie E (2005) Perceiving intentions. In: Sâágua J (ed) *A Explicação da Interpretação Humana*. Edições Colibri, Lisbon, pp 401–414
- Pacherie E (2006) Towards a dynamic theory of intentions. In: Pockett S, Banks WP, Gallagher S (eds) *Does Consciousness Cause Behavior? An Investigation of the Nature of Volition*. MIT Press, Cambridge, MA, pp 145–167
- Pacherie E (2008) The phenomenology of action: a conceptual framework. *Cognition* 107:179–217
- Park J, Kitayama S (2012) Interdependent selves show face-induced facilitation of error processing: cultural neuroscience of self-threat. *Soc Cogn Affect Neurosci*. doi:10.1093/scan/nss125
- Pylshyn Z (1999) Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behav Brain Sci* 22(3):341–365

- Reddy V (1991) Playing with others' expectations: teasing and mucking about in the first year. In: Whiten A (ed) *Natural Theories of Mind*. Blackwell, Oxford, pp 143–158
- Reddy V (2008) *How Infants Know Minds*. Harvard University Press, Cambridge
- Rorty R (1989) *Contingency, irony, and solidarity*. Cambridge University Press, Cambridge
- Runeson S, Frykholm G (1983) Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender recognition, and deceptive intention. *J Exp Psychol* 112:585–615
- Sartori L, Becchio C, Castiello U (2011a) Cues to intention: the role of movement information. *Cognition* 119:242–252
- Sartori L, Straulino E, Castiello U (2011b) How objects are grasped: the interplay between affordances and end-goals. *PLoS One* 6(9):e25203. doi:10.1371/journal.pone.0025203
- Sass L, Pienkos E (2013) Faces of intersubjectivity: interpersonal experience in melancholia, mania, and schizophrenia. *J Conscious Stud* 20(7–8):103–130
- Scheler M (1954) *The nature of sympathy*. London: Routledge. Kegan Paul
- Scholl BJ, Leslie AM (1999) Modularity, development, and 'theory of mind'. *Mind Lang* 14:131–153
- Searle J (1983) *Intentionality: an essay in the philosophy of mind*. Cambridge University Press, Cambridge
- Segal G (1996) The modularity of theory of mind. In: Carruthers P, Smith P (eds) *Theories of theories of mind*. Cambridge University Press, Cambridge, pp 141–157
- Shuler MG, Bear MF (2006) Neurons in the primary visual cortex respond differently to a flash of light after it has been paired with a reward, unexpectedly showing that cognitive information is coded at this level in the cortex. *Science* 311:1606–1609
- Siegel S (2011) Cognitive penetrability and perceptual justification. *Nous* 46(2):201–222
- Smith J (2010) Seeing other people. *Philoso Phenomenol Res* 81(3):731–748
- Spaulding S (2010) Embodied cognition and mindreading. *Mind and Lang* 25:119–140
- Strawson G (1994) *Mental reality*. MIT Press, Cambridge, MA
- Stueber KR (2006) *Rediscovering empathy: agency, folk psychology, and the human sciences*. MIT Press, Cambridge, MA
- Tajfel H (1981) *Human groups and social categories*. Cambridge University Press, Cambridge
- Tajfel H, Billig MG, Bundy RP, Flament C (1971) Social categorization and intergroup behaviour. *Eur J Soc Psychol* 1(2):149–178
- Trevarthen CB (1979) Communication and cooperation in early infancy: a description of primary intersubjectivity. In: Bullowa M (ed) *Before speech*. Cambridge University Press, Cambridge
- Trevarthen C, Aitken KJ (2001) Infant intersubjectivity: research, theory, and clinical applications. *Annual Research Review. The J Child Psychol Psychiatry* 42(1):3–48
- Wellman HM (1990) *The child's theory of mind*. MIT Press, Cambridge, MA
- Wellman H, Banerjee M (1991) Mind and emotion: children's understanding of the emotional consequences of beliefs and desires. *Br J Dev Psychol* 9:191–214
- Wittgenstein L (1967) *Zettel*. In: Anscombe GEM, von Wright GH (eds) (G. E. M. Anscombe, Trans.). University of California Press, Berkeley
- Wittgenstein L (1980) *Remarks on the philosophy of psychology*. In: von Wright GH, Nyman H (eds) (Vol. II) (C. G. Luckhardt & M. A. E. Aue, Trans.). Blackwell, Oxford
- Xu Y (2005) Revisiting the role of the fusiform face area in visual expertise. *Cereb Cortex* 15(8):1234–1242
- Xu Y, Zuo X, Wang X, Han S (2009) Do you feel my pain? Racial group membership modulates empathic neural responses. *J Neurosci* 29(26):8525–8529